5 Development and Implementation of the Screening System

The aim of this work was the development of a peptide-based screening system for the systematic evaluation of the interaction properties of fluoroalkyl-substituted amino acids in native polypeptide environments. As described in section 4.3, an α-helical *coiled coil* peptide or protein is a perfect structural basis for a model polypeptide to be applied. Such a model polypeptide had to be designed and tested for its applicability for the desired investigations. Systematically altered, side chain-modified and $C^{\alpha,\alpha}$-dialkylated fluoroalkyl-substituted amino acids had to be incorporated into the hydrophobic and the charged interaction domain, respectively, within a specifically designed *coiled coil* protein. Thermostability measurements and replication experiments with the substituted *coiled coil* proteins were supposed to provide information about the properties and interaction behavior of fluoroalkyl groups in native polypeptide environments. Furthermore, a phage display strategy that is based on the *coiled coil* model polypeptide was to be invented for the selection of the best interaction partners for the fluoroalkyl-substituted building blocks.

5.1 The *Coiled Coil*-Based Screening System

The *coiled coil*-based model polypeptide had to be designed so that it would be applicable for both parts of the planned investigations, the evaluation of interaction properties of fluoroalkyl-substituted amino acids in native polypeptides as well as phage library screenings. Therefore, the non-natural building blocks had to be incorporated into both of the specific *coiled coil* interaction domains. To ensure that the fluoroalkyl side chains interact with native hydrophobic or charged amino acids, respectively, a homodimeric system was designed with an antiparallel orientation of both helical peptide strands. For phage display, the part of the *coiled coil* protein containing the specific interaction partners of the substitution sites constitutes the library peptide and has to be displayed on phage surface. The complementary part, containing the substitution sites, acts as ligand for library screening.

5.1.1 *Coiled Coil* Design for the Evaluation of the Interaction Properties

As described in section 4.2, one important requirement for a model polypeptide that can be used for the intended investigations is a tertiary structure that is not disturbed by the amino acid substitutions to be made. Therefore, an antiparallel *coiled coil* homodimer was designed

that consists of 41 amino acids per helix strand comprising almost six complete heptad repeats (Figure 5.1).

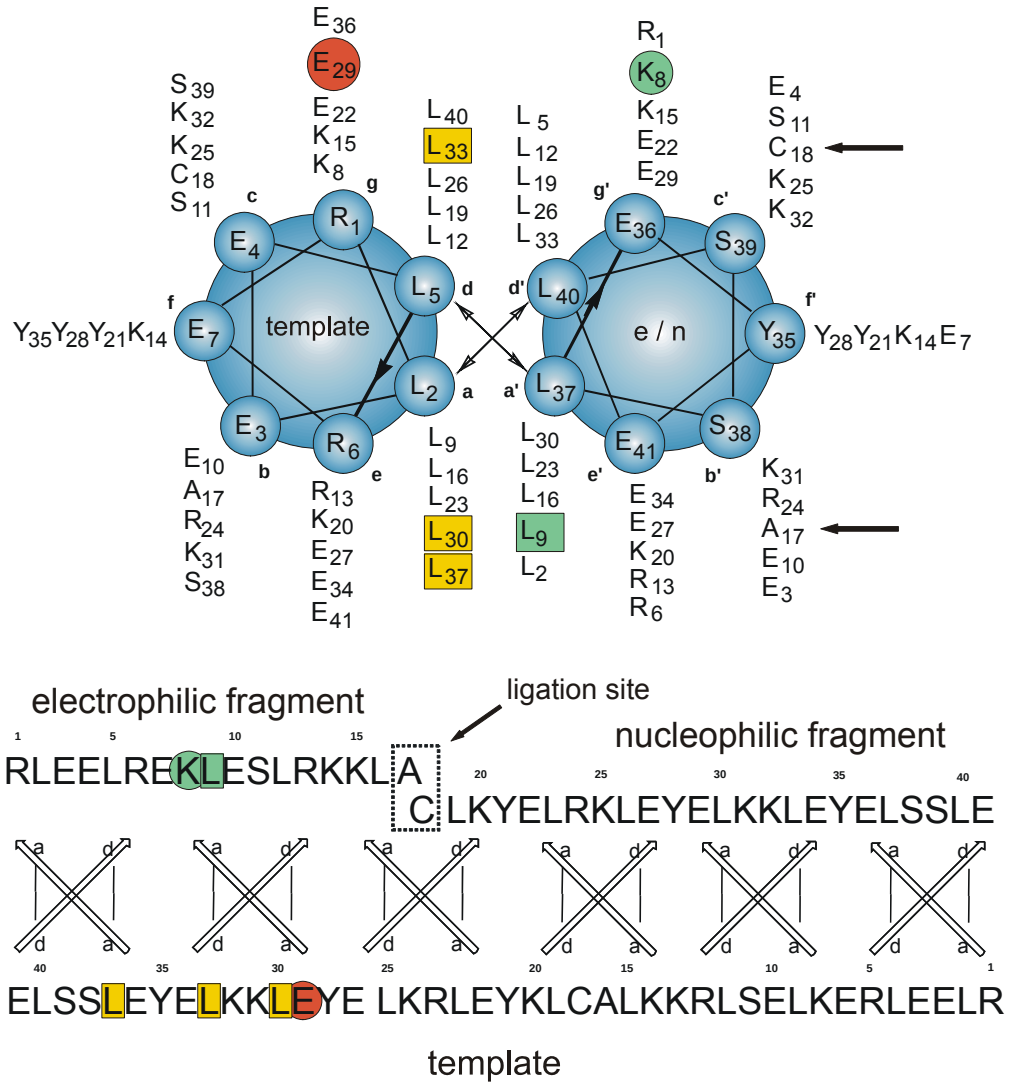


Figure 5.1: *Helical-wheel and sequence representation of the α-helical coiled coil homodimer. Substitution positions within the hydrophobic interface are highlighted with green squares and within the charged interface with green circles. Their interaction partners are highlighted with yellow squares and red circles, respectively. The ligation site of the electrophilic and nucleophilic fragments is marked with arrows.*

The hydrophobic core of the α-helical *coiled coil* is exclusively composed of leucine residues, which favors the dimerization over the formation of higher order oligomers[189] as well as an antiparallel over an parallel orientation of the homodimer.[190] An even stronger dictation for the relative orientation of both helix strands is provided by the design of the positions **e** and **g** within the charged interaction domain. The antiparallel orientation is characterized by twelve attractive Glu-Lys or Glu-Arg electrostatic interactions between positions **g**-**g**' and **e**-**e**',

respectively. In most cases, the formation of such salt bridges stabilizes a *coiled coil* structure.[191] Consequently, an parallel arrangement would be highly destabilized due to exclusively repulsive charged interactions. Three **f** positions, Tyr21, Tyr28, and Tyr35 contain tyrosine residues that serve as fluorescence labels for replication experiments. The occupation of the remaining amino acid positions is either based on the sequence of hepatitis δ antigen that forms stable dimeric *coiled coils* with an antiparallel orientation of both helix strands[192] or was configured in respect of the net charge of the protein and the *coiled coil*-forming propensities of the amino acids.[193] The design concept of the newly designed antiparallel *coiled coil* homodimer could successfully be proven by our group using FRET studies, NMR investigations and molecular dynamic calculations.[194]

Two amino acid positions within the helical peptides serve as substitution positions for the fluoroalkylated amino acids, Lys8 (**g** position) within the charged interaction domain and Leu9 (**a** position) within the hydrophobic core. The defined charged interaction partner for Lys8 on the complementary helix strand in the peptide dimer is Glu29. Leu9 forms a hydrophobic interaction cluster with Leu30 and Leu37 in **a** positions as well as Leu33 in **d** position of the dimer partner. Consequently, the substitution of one of these two native amino acids in a helix strand by a fluorinated building block results in two sites within the *coiled coil* dimer, where fluoroalkyl groups interact with native amino acids.

Since *coiled coil* self-replication was used to investigate the influence of fluoroalkyl-substitution of amino acid side chains on protein folding (see section 4.3.2), the *coiled coil* peptides were synthesized in two parts. The electrophilic peptide fragments that contain both of the substitution sites consist of amino acids 1-17 and are synthesized as C-terminal benzylthioesters. Since bigger side chains of the carboxy-terminal residue can interfere with the ligation reaction, alanine is present in position 17. The nucleophilic fragment comprises amino acids 18-41 and contains a cysteine residue at the amino-terminal end. Since the nucleophilic fragment contains the amino acid positions that constitute the interaction partners of the fluorinated side chains and bears no substitution site, it is identical for all substituted *coiled coils*. Consequently, only one nucleophilic fragment has to be synthesized, while amino acid alterations are all made within the electrophilic peptide fragments, which have to be synthesized seperately. The respective template strand, which constitutes the full-length monomer for *coiled coil* formation can then be produced by native chemical ligation of both the native or substituted electrophilic with the nucleophilic peptide fragments.

The described *coiled coil*-based model polypeptide can be used for the evaluation of the properties and molecular interactions of fluoroalkyl-substituted side chains with native charged residues (Glu29) and hydrophobic side chains (Leu30, Leu33, and Leu37), respectively, by applying two screening methods. A first screen measures the thermostabilities of the *coiled coil* dimers that bear amino acid substitutions in positions Lys8

or Leu9. In a first step, altered electrophilic fragments have to be ligated to the nucleophilic fragment in a quantitative scale (Figure 5.2).

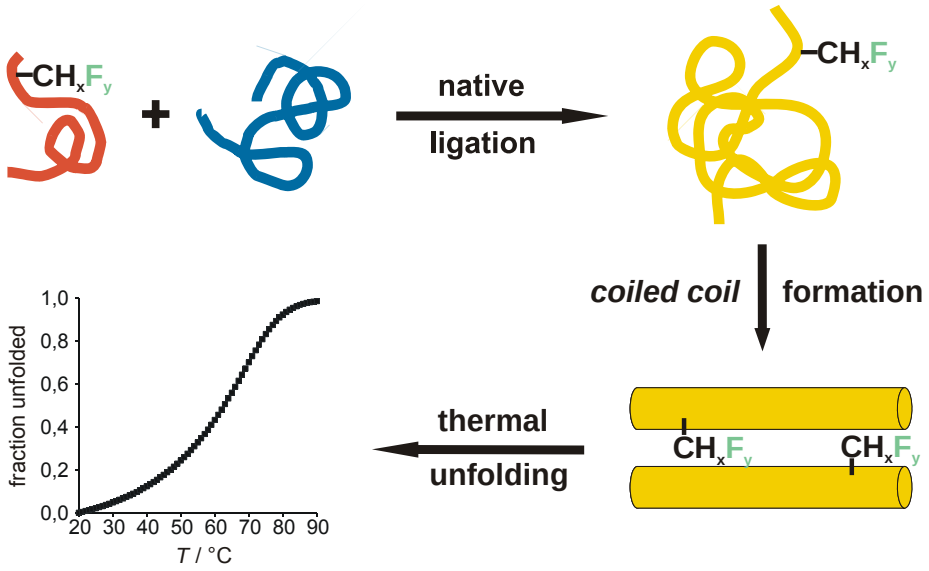


Figure 5.2: *Schematic representation of the thermal unfolding screen. Substituted electrophilic peptide fragments are shown in red, the nucleophilic peptide fragment in blue, and the ligation product in yellow. Amino acid side chain substitutions are marked with fluoroalkylgroups.*

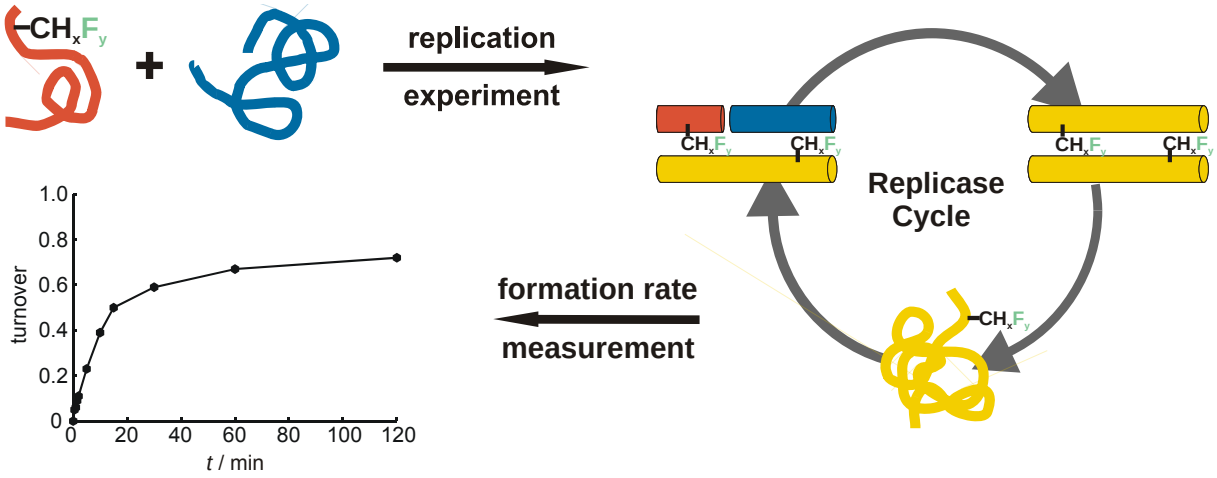


Figure 5.3: *Schematic representation of the self-replication screen. Substituted electrophilic peptide fragments are shown in red, the nucleophilic peptide fragment in blue, and the ligation product in yellow. Amino acid side chain substitutions are marked with fluoroalkylgroups.*

The purified ligation products will form stable antiparallel *coiled coil* dimers. The influence of the alkyl fluorination on amino acid interactions within both of the specific *coiled coil*

interaction domains will affect the thermostabilities of the full-length dimers and, thus, can be detected via recording thermal unfolding profiles of the subtituted proteins.
A second screen measures the product formation rates of self-replication experiments with substituted electrophilic peptide fragments and the nucleophilic peptide fragment (Figure 5.3).

The influence of fluoroalkyl-substitution within the amino acid positions Lys8 and Leu9 on the folding and unfolding properties of native proteins can thus be studied.

5.1.2 *Coiled Coil* Design for the Selection of Preferred Interaction Partners

The *coiled coil* model polypeptide that should be applied for the selection of the preferred interaction partners for fluoroalky-substituted amino acids via library screening is principally based on the antiparallel dimer that has been decribed above. Since phage display is used for the library screens, one part of the *coiled coil* is produced biosynthetically by bacteria and displayed on the phage surface, while the other constitutes the screening ligand, which is synthesized chemically. A protein fragment, the so-called *stem loop*, has been designed for being fused to the phage coat, which contains the full-length template and the nucleophilic fragment. Both parts are connected via a peptide linker (Figure 5.4).

In contrast to the *coiled coil* used for thermostability and self-replication screens, the *stem loop* fragment misses the three tyrosine residues in **f** positions, since they are not needed for analytical purposes. Furthermore, the amino acids required for native ligation, Ala17 and Cys18, are not included in the template part of the *stem loop*. In both cases, residues have been chosen instead that are found in the analogous positions within the native α-helical *coiled coil* dimer of hepatitis δ antigen. The amino-terminal end of the *stem loop* is occupied by a cysteine residue for native ligation. Comparable to the *coiled coil* used for the self-replication screen, electrophilic peptides that contain the substituted fluorinated amino acids, bind to the *stem loop* in a *coiled coil* manner and, thus, constitute the ligands for library screens (Figure 5.5). Native ligation results in a covalent linkage between the electrophilic peptide fragments and the fusion protein on phage. For capture of bound phage particles, electrophilic fragments contain an amino-terminal biotin.
The interaction partners of the fluorinated amino acids, which are located in the template part of the *stem loop*, constitute the randomization positions for the library. In addition to amino acids in positions 30, 33, and 37, which represent the interaction partners for the substitution position Leu9 in a hydrophobic core-packing, the side chain of the amino acid in position 34

can interact with the fluoroalkyl-substituted residue in position 9 and should, therefore, additionally be randomized (Figure 5.6).

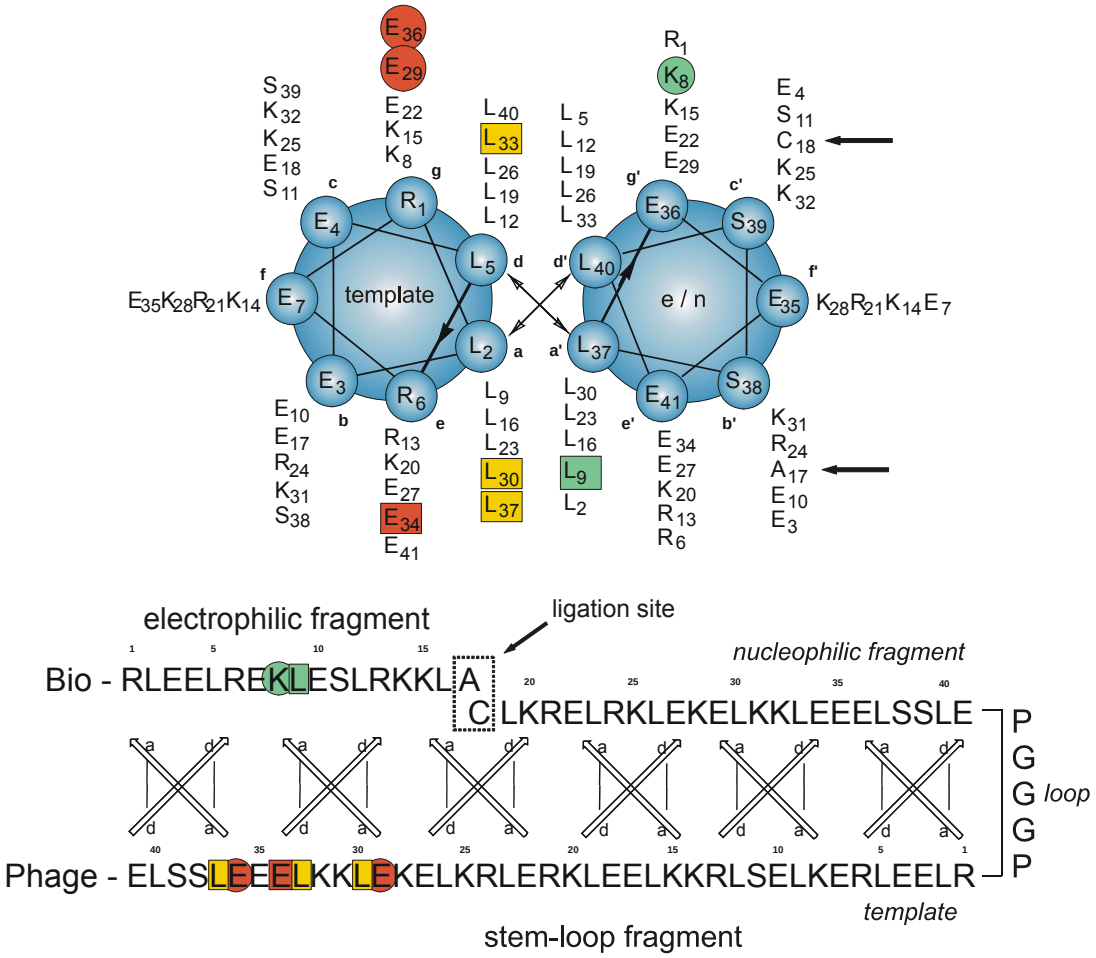


Figure 5.4: *Helical-wheel and sequence representation of the α-helical coiled coil dimer used for phage display. Substitution positions within the specific interaction domains in the screening ligand are highlighted with green squares and circles. The randomization positions are highlighted with yellow or red squares and circles, respectively. The ligation site of the electrophilic fragments and the stem loop is marked with arrows.*

This **e** position is naturally occupied by glutamic acid. Furthermore, the amino acid position Glu36 has to be randomized in the library, since it can compete with Glu29 for polar interactions with the fluorinated amino acid in substitution position Lys8. Thus, six amino acid positions in the template part of the *stem loop* have to be randomized on phage, resulting in an overall library diversity of 20^6 (6.4 x 10^7). Once the library containing *stem loop* protein is displayed on phage surface, biotinylated electrophilic fragments, which are substituted with fluorinated amino acids, are used to select the best binders out of the 20^6 library members.

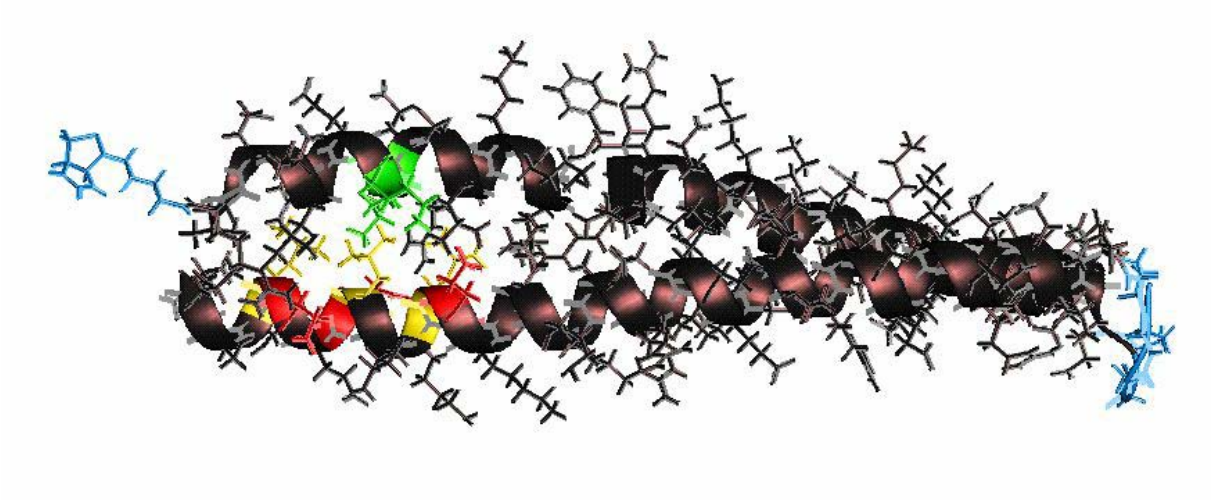


Figure 5.5: *Model of the stem loop protein to be displayed on phage surface with bound biotinylated electrophilic fragment. Biotin and the peptide linker connecting template and nucleophilic fragments are highlighted in blue. Substitution positions are colored in green and randomization positions in yellow (Leu residues) and red (Glu residues), respectively.*

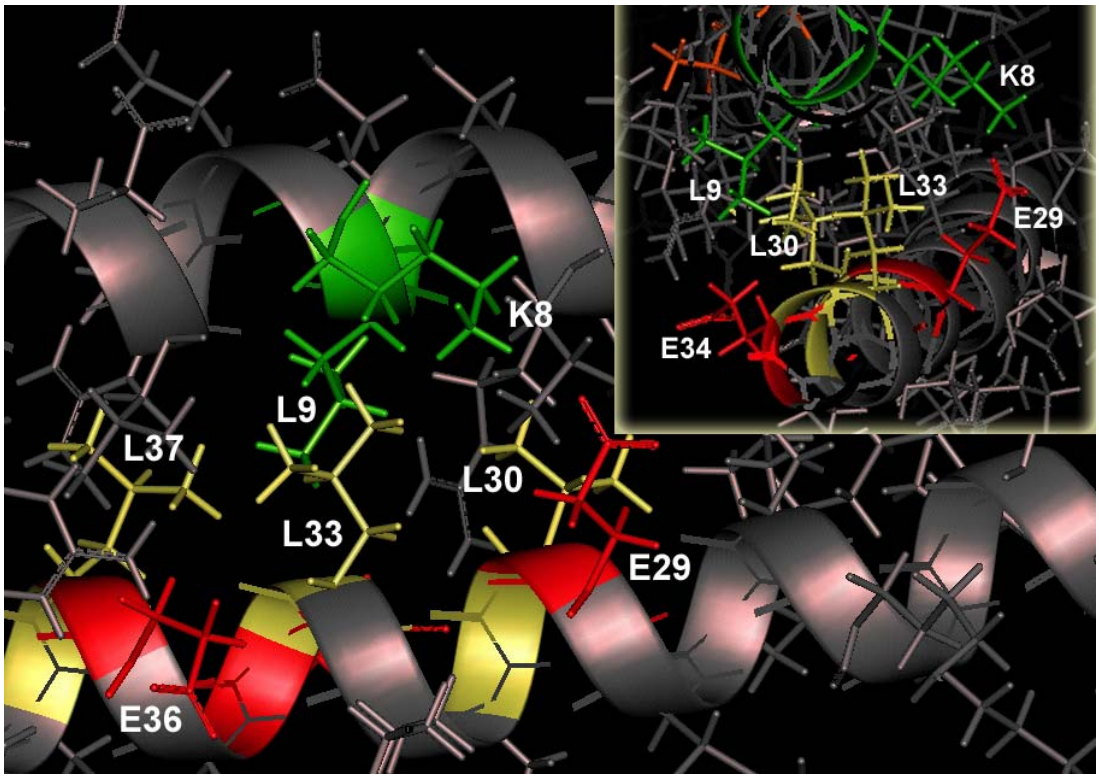


Figure 5.6: *Details of a model of the protein to be displayed on phage surface. The large image represents a side view of both the electrophilic and the stem loop fragments. The upper right inset image provides a view along the helix axis. Substitution positions Lys8 and Leu9 are colored in green. Their possible interaction partners in the library are colored in yellow (leucine residues in the hydrophobic core) and red (glutamic acids in the charged interaction domain), respectively.*

After several selection rounds, the DNA of the enriched phage clones is sequenced, and the sequence of the displayed proteins is decoded (Figure 5.7). The desired result is the determination of the optimal binding partners for the systematically altered fluoroalkyl groups. This information allows conclusions about the nature of contacts, e.g., hydrogen bonds, polar or hydrophobic interactions, in which different fluoralkyl moieties preferentially participate.

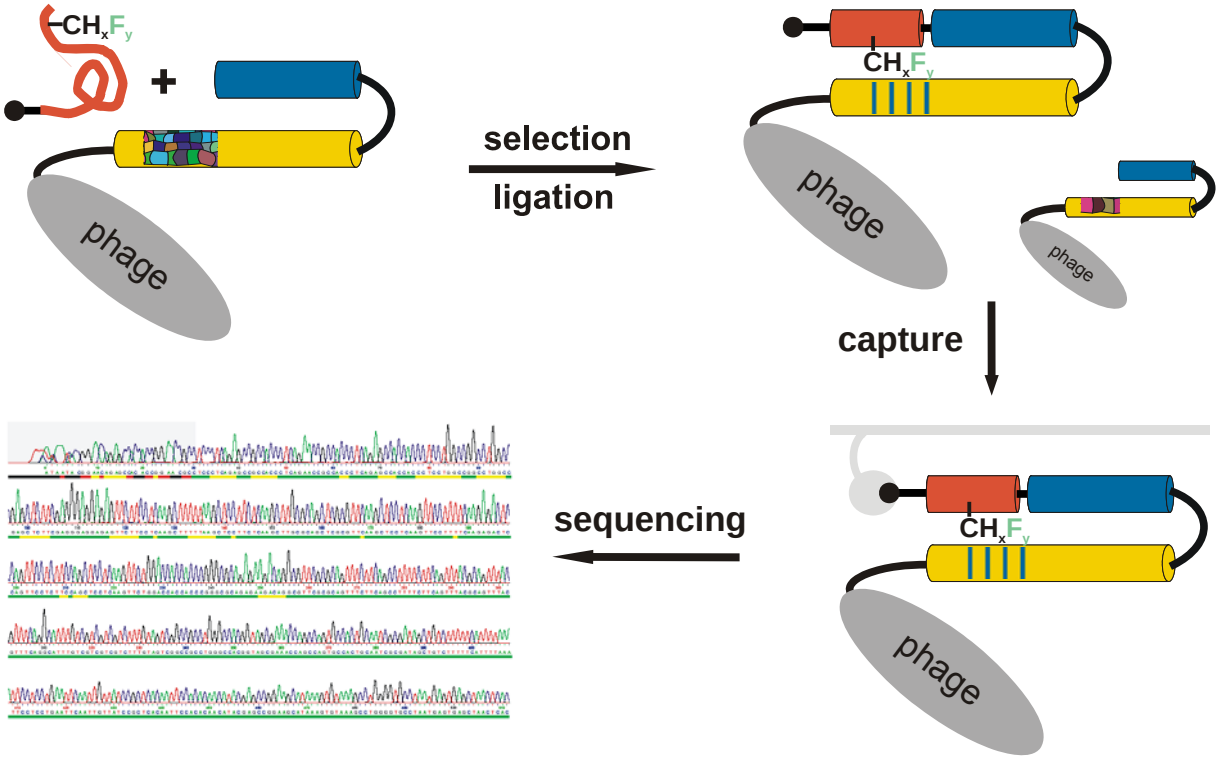


Figure 5.7: *Schematic representation of the library screen. Substituted electrophilic peptide fragments are shown in red, the nucleophilic part and the template part of the stem loop in blue and yellow, respectively. Amino acid side chain substitutions are marked with fluoroalkyl groups and the biotin anchor with a closed circle. The randomized part of the library is represented by multicolored squares, the selection by blue stripes.*

Since the amino-terminal biotin label of the electrophilic peptide fragments would not interfere with the thermostability and self-replication screens, the substituted peptides only have to be synthesized once and can also be used for all of the described investigations.

5.2 Investigations on the Interaction Properties of Fluorinated Amino Acids

The designed model system for the evaluation of the interaction properties of fluoroalkyl groups with native amino acids (Section 5.1.1) had to be proven regarding the sensitivity of

the thermostability and self-replication screens. The structural variance of the systematically modified fluoroalkyl side chains must result in different melting profiles and self-replication rates, respectively. However, the proof of the design concept requires good reproducibility of the obtained results. Therefore, both the thermostability and the self-replication experiments had to be optimized. At first, nucleophilic (N) and biotinylated electrophilic peptide fragments containing the original residues Lys8 and Leu9 in the substitution positions (E_o) were synthesized using appropriate solid phase peptide synthesis strategies.

5.2.1 The Optimization of the Screening Methods

The fluoroalkyl-substituted amino acids are incorporated into one of the two substitution positions Lys8 and Leu9, respectively, within the electrophilic peptide fragments. Consequently, only two of these non-natural building blocks are present in the full-length *coiled coil* dimer that consists of 82 amino acids. Thus, only minor differences in melting behavior and rate of self-replication between the peptides that contain the substituted amino acids may be detectable. The correct interpretation of these experiments implies minimal errors of the obtained results.

5.2.1.1 The Optimization of Thermal Denaturation Experiments

Since fluoroalkyl-substituted amino acids are incorporated into the electrophilic *coiled coil* fragments and thermal denaturation experiments are performed with the full-length helical dimer, the screening method consists of two experimental parts: the native ligation of the electrophilic peptide with the nucleophilic fragment and the recording of the melting profile of the dimeric ligation product.

Preparative Native Ligation

The native ligation of both peptide fragments via the carboxy-terminal benzylthioester of the electrophilic peptide and the amino-terminal cysteine residue of the nucleophilic peptide is catalyzed by *coiled coil* formation (Section 4.3.2). After the reaction, the ligation product has to be purified, because side products and remaining peptide fragments and reactants would influence the thermal denaturation of the *coiled coil* dimer. The isolation of the full-length *coiled coil* peptide from the reaction mixture proved to be problematic. Absorbance detection at 220 nm was not sufficient to select the ligation product in an HPLC run, because the chromatograms showed high complexity and the peptides eluted very closely to each other,

even when an extremely flat eluent gradient was used. Fluorescence detection using exitation and emission wavelengths, which are specific for tyrosine residues,[195] could be proven to be more effective, because only nucleophilic fragment, intermediate, and end products are shown in the chromatogram. Furthermore, this method was sensitive enough to detect even small amounts of the peptides. HPLC with fluorescence detection could thus be used for monitoring the reactions as well as for purifying the ligation products (Figure 5.8).

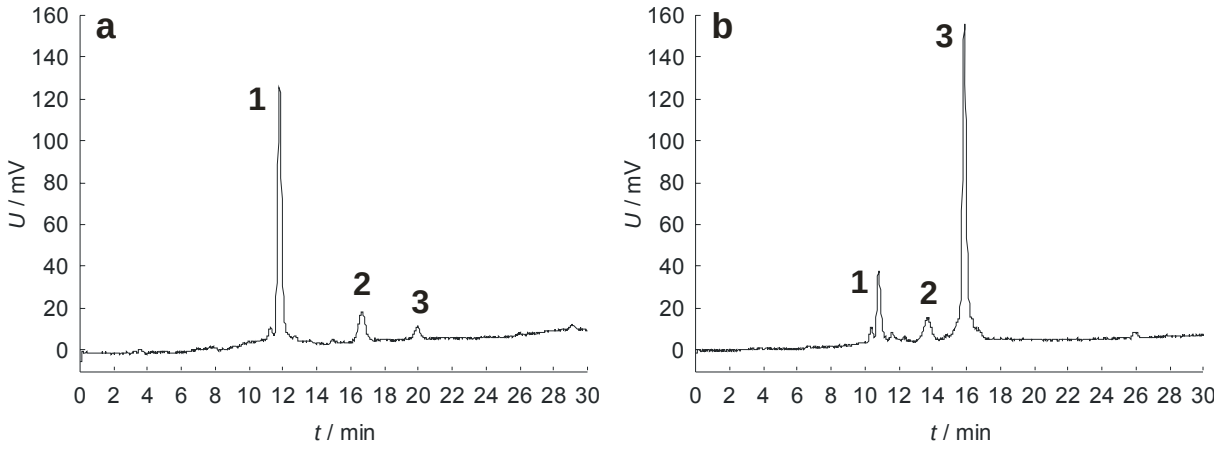


Figure 5.8: *HPLC chromatograms of the ligation reaction of the unsubstituted electrophilic fragment (E_o) and the nucleophilic fragment (N) after starting the reaction* ***(a)*** *and after three hours of reaction* ***(b)****, recorded via fluorescence detection. Peaks are marked with numbers: nucleophilic fragment (1), intermediate (2), and final product (3).*

Since the ligation reaction has to be performed on a preparative scale in order to obtain sufficient amounts of peptide dimer, and the purification by analytical HPLC allows only small volumes, the concentration of both peptide fragments had to be as high as possible. In doing so, problems occurred due to the poor solubility of the electrophilic fragment in aqueous solution at neutral pH values, which are required for the native ligation reaction. Furthermore, increasing peptide concentrations provoked side reactions, which lowered the yield and complicated the purification by HPLC. Thus, a compromise had to be made in terms of the concentration of the peptides and reaction volume.

Since the nucleophilic fragment can dimerize in solution via disulfide formation of two cysteine residues, a disulfide reducing reagent had to be added to the peptide before starting the reaction. Several reducing agents had been tested in respect of water solubility, compatibility with the chemical reaction, and stability in buffered aqueous solution. Although the addition of cystine reducing thiols to the ligation reaction can accelerate the transesterification,[196] tris(2-carboxyethyl)phosphine (TCEP) was choosen as the additive for the ligation reactions. Due to a rapid and highly efficient reduction of disulfides, TCEP can be

used in concentrations that disturb neither the native chemical ligation nor the purification by HPLC. Under acidic and neutral conditions, TCEP is stable for several hours in aqueous solution.[197]

Thermal Denaturation

The purified and lyophylized ligation products exhibited very weak solubility in aqueous buffer at neutral pH 7.4. However, thermal denaturation experiments have to be performed under physiological conditions, since the interaction behavior of fluoroalkyl groups with the model polypeptide must be comparable to that in native systems. This problem could not be solved and, the CD spectra of the ligation products could consequently not be recorded. However, the addition of guanidine hydrochloride to the buffered peptide solution resulted in total solubility of the *coiled coil*.

As described in section 4.2, the model polypeptide was designed to exhibit a very high structural stability in order to accept any single-amino acid substitutions without alterations of the overall folding. However, denaturation of the protein must be possible to detect the influence of the amino acid fluorination on side chain interactions. Attempts of thermal denaturation of the unsubstituted *coiled coil* dimer, which were monitored by CD-spectroscopy, revealed the problem that the protein showed no unfolding until 90°C. Therefore, reagents that help denaturating protein structure had to be used. The addition of urea did not result in sufficient denaturation of the protein until 90°C, even at the highest possible concentration (10 M urea). Therefore, guanidine hydrochloride (GdnHCl), a compound that is widely used as denaturating agent in probing folding and structural stability of proteins[198] was used in melting experiments.

When optimizing the concentration of GdnHCL, special attention had to be payed to the fact that most amino acid substitutions were expected to destabilize the *coiled coil* folding of the dimer. That means that a GdnHCl concentration, which effects a complete sigmoidal denaturation profile of the unsubstituted model polypeptide until 90°C, may denaturate substituted variants before melting. Several concentrations were tested, and fine-tuning resulted in a thermal denaturation protocol including 5M guanidinium chloride. Although the addition of GdnHCl led to complete solubility of the *coiled coil* dimers, CD spectra in the far UV range (190-250nm) could not be recorded, since the absorbance of the denaturating agent below 215nm disturbed the measurements. However, the decrease of the CD-minimum at 222nm, which is characteristic for α-helical peptide folding,[199] could be monitored by CD-spectroscopy during denaturation experiments.

The major question behind the use of GdnHCl in thermal denaturation experiments was whether 8M stock solutions of this compound can be produced with sufficient reproducibility,

since differences in GdnHCl concentration between the measurements of different substituted model polypeptides will strongly affect their melting profiles. Therefore, test experiments were performed with the unsubstituted model polypeptide, including completely separated ligation reactions, product purifications, protein concentration determinations, denaturation measurements, and independently made GdnHCl stock solutions (Figure 5.9).

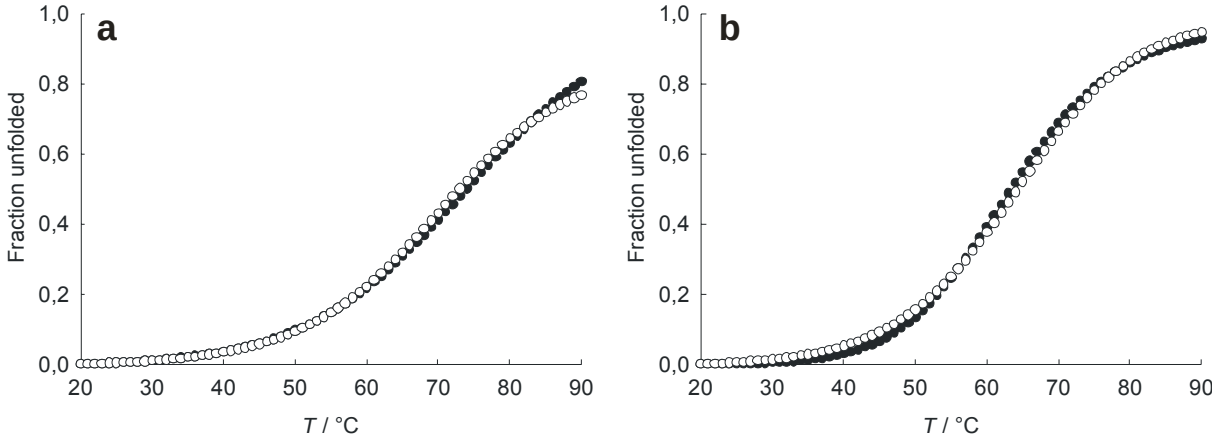


Figure 5.9: *Thermal denaturation profiles of unsubstituted coiled coil peptides of full length (P_o) from separate ligation experiments (open and closed circles, respectively) with* ***a)*** *TCEP and* ***b)*** *BzlSH as reducing agent added to the peptide solution before thermal denaturation. Melting was performed in 5M GdnHCl.*

The results prove a sufficient reproducibility of the thermal stability screen but show that the disulfide reducing reagent that has to be added to the *coiled coil* dimer influences the stability of the model polypeptide. Since two cysteine residues are present in the peptide dimer, the reduction of possible disulfides is indispensible for obtaining correct stability data. Since benzylmercaptan leads to inhomogenity of aqueous solutions and is shown to destabilize the folding motif to a greater extent than TCEP, the phosphin that was chosen as reducing agent in thermal denaturation experiments.

5.2.1.2 The Optimization of Self-Replication Experiments

The replication reaction with the nucleophilic peptide fragment N and the electrophilic peptide fragment E_o without amino acid substitutions appeared to be extremely sensitive to molecular, physicochemical, and environmental conditions. An experimental protocol was thouroughly established that reveals a sufficient reproducibility of the data obtained from the reaction rate measurements.

Since the nucleophilic fragment contains a cysteine residue that is required for native ligation, peptide dimerization via disulfide linkage can occur in the reaction solution. Such side reactions would interfere with the *coiled coil* folding and unfolding processes as well as the chemistry of the self-replication cycle and, thus, would strongly affect the replication rate. Therefore, a disulfide reducing agent was added to the reactions. Since benzylmercaptan is a side product of the native peptide ligation, when electrophilic peptide fragments are used that possess a carboxy-terminal benzylthioester, the first idea was to apply this thiol as reducing agent. Thus, no additional compound was present in the reaction solution and an increase in complexity of the reaction system was avoided. This attempt failed to provide reproducible results. One reason might have been that benzylmercaptan is not soluble in water, as it forms a separate phase in the reaction solution. A two-phase system may lead to partitioning of the peptide fragments and the ligation product, therefore, affecting the reaction rate. Alternatively, the applicability of water-soluble thiols as reducing agents for the replication experiments, such as thioglycolic acid (TS) and dithiothreitol (DTT), have been tested. DTT is a very common reagent for the reduction of cystine bonds in peptides and proteins.[200] However, reaction monitoring via analytical HPLC identified unexpected side-reactions in replication experiments with TS and DTT. However, no negative influence on the autocatalyzed ligation reactions could be found for TCEP. Therefore, this disulfide reducing reagent was used for further replication experiments.

One major problem was the monitoring of the replication reactions. Analytical HPLC was used to analyze the reaction progress at definite times. Therefore, aliquotes of the reaction mixture were taken and reactions were stopped by addition of stop solution. The resulting peptide solutions had to be stable until the start of the respective HPLC run, which could take up to six hours for later aliquotes. Therefore, the composition of the stop solution had to be optimized in respect to solubility as well as stability of the reaction components. Thereby, stop solution compositions that supported peptide solubility often decreased the stability of the reaction components. A big step forward, regarding this problem, could be made by changing the HPLC method from UV absorption to specific fluorescence detection of the three tyrosine residues in the nucleophilic fragment as well as in the ligation product (Figure 5.10).

Since fluorescence detection is much more sensitive compared to simple UV absorption, higher dilutions of the reaction aliquotes in the stop solution could be made without affecting sensitivity of the detection method. In addition, only nucleophilic peptide fragment, intermediate and end products are shown in the chromatogram, which eases the assignment of the peaks. A further advantage of the fluorescence detection strategy is the easier analysis of the results. Since UV absorption of peptides is dependent on the number and the composition of amino acids, peak areas of the peptide fragments cannot be compared to

those of the ligation product. Thus, only the absolute peak areas of the ligation product can be used for the determination of the ligation rate.

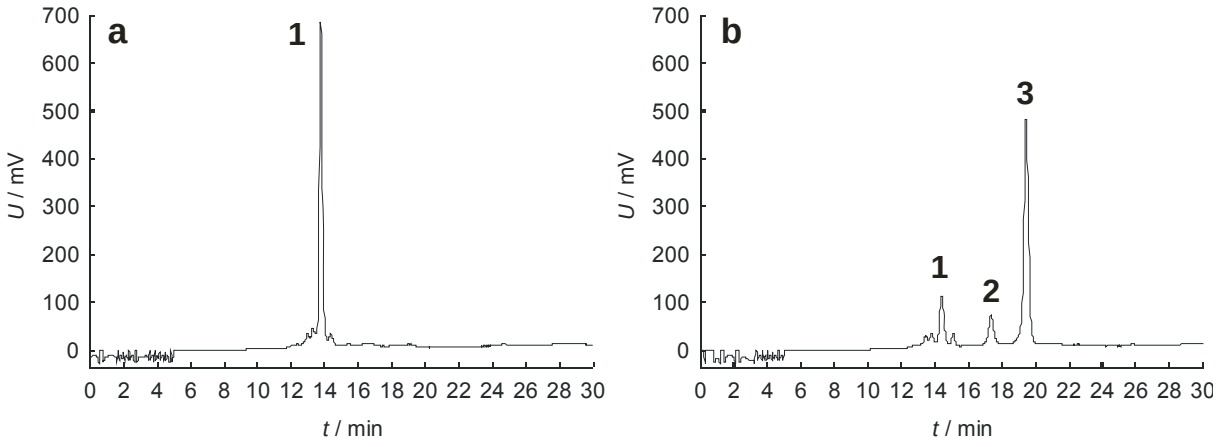


Figure 5.10: *Monitoring of the reaction progress of a self-replication with the nucleophilic fragment (N) and the unsubstituted electrophilic fragment (E_o) via analytical HPLC at the beginning **(a)** and after 120 min **(b)**. Peptide peaks are marked with numbers: 1: nucleophilic fragment, 2: intermediate and 3: end product.*

In contrast, the sum of the peak areas of the three peptides shown in the fluorescence chromatogram is supposed to be constant in all analytical runs of a single monitored reaction due to the identical content of tyrosine residues. Thus, the turnover of the replication reaction can be determined from the relations of the peak areas. This excludes error sources from the turnover calculations, such as inconstant volumes that are injected by the HPLC autosampler (Figure 5.11).

Smaller values for the sums of the peak areas could be calculated for the first aliquotes of the reactions (until 1.5 min) compared to all subsequent samples. No explanation for this phenomenon could be found. However, no consistent tendency regarding falling or rising values during the monitored reaction time could be observed, which validates the applied fluorescence detection strategy for the determination of self-replication rates.

Although the stepwise optimization of the reaction parameters like those described above as well as temperature, peptide concentrations, and pH resulted in reproducible turnover plots for sample reactions, this screening method remains extremely sensitive to several environmental factors. Thus, conditions such as reaction temperature have to be kept and the complete operating schedule including HPLC measurements must be strictly followed to assure comparable results.

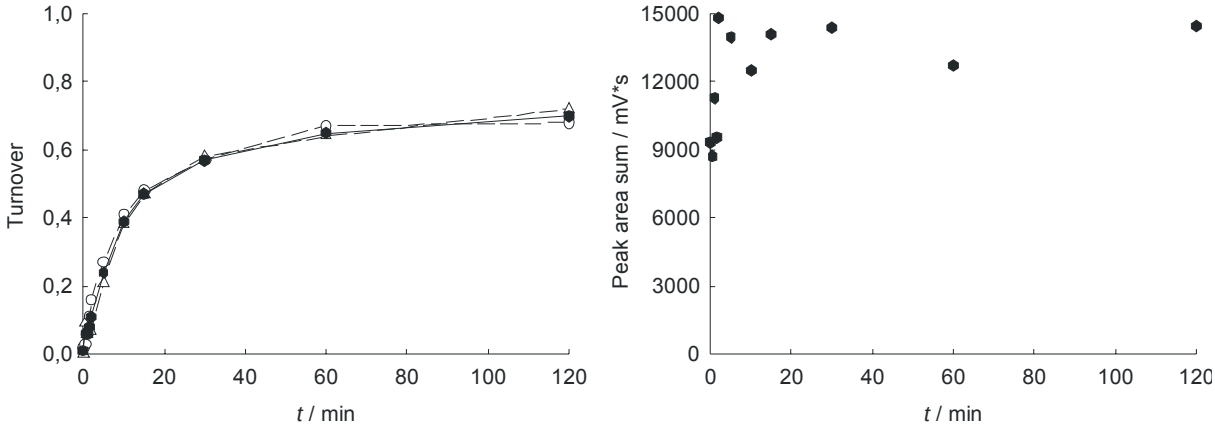


Figure 5.11: *Replicase turnover rates **(left diagram)** and peak area sums **(right diagram)** determined by HPLC with fluorescence detection. Two reactions of the nucleophilic fragment (N) with the unsubstituted electrophilic fragment (E_o) were performed independently (open symbols, dashed lines), and the average values calculated (closed hexagons, solid line). Peak area sums include peak areas of the nucleophilic fragment, the intermediate, and the end products.*

5.2.2 Proof of Concept

Before the designed *coiled coil*-based model polypeptide and the optimized thermostability and self-replication screens, respectively, could be used for the evaluation of the interaction properties of fluoroalkyl-substituted amino acids in native protein environments, the design concept of the novel screening system had to be proven. One aspect to be tested was the sensitivity of both screening methods towards amino acid side chain alterations in both substitution positions, Lys8 in the charged interaction domain and Leu9 in the hydrophobic core, respectively. Even minor changes in side chain structure, e.g., in the content of fluorine atoms, must result in discriminable melting profiles and self-replication rates, respectively. A further important requirement for the applicability of the strategy is the presence of the amino acid side chain interactions between the substitution positions and their interaction partners in the model polypeptide, which are the salt bridge Lys8-Glu29 in the charged interaction domain and the zipper-like cluster Leu9-Leu30/Leu33/Leu37 in the hydrophobic core.

5.2.2.1 The Lys8-Glu29 Salt Bridge Formation

When folded into a *coiled coil* structure, the Leu9 side chain in **a** position is buried in the hydrophobic interface of both α-helical peptide strands and is forced to interact with the

adjacent leucine residues of the interaction core. Since the *coiled coil* formation of the designed model polypeptide could be proven, the Leu9-Leu30 / Leu33 / Leu37 interactions can be assumed for the investigations on fluoralkyl-substituted amino acids in hydrophobic protein environments and do not need to be evaluated in detail. In contrast, the amino acid side chain interaction that was to be used for investigating the interactions of fluoroalkyl groups in polar protein environments, the Lys8-Glu29 salt bridge, had to be investigated before substituting Lys8 with fluorinated amino acids. Regarding the classical *coiled coil* design, Glu29 is the interaction partner of Lys8. However, since the ends of α-helices are often somewhat flexible in structure and both residues have other charged side chains in proximity, the formation of alternative salt bridges had to be considered.

In the native antiparallel *coiled coil* structure of hepatitis δ antigen, the basis for the design of the model polypeptide described here, the arginine residue comparative to Arg1 in the **g** position of the model polypeptide does not form a salt bidge with Glu36 in **g'** position on the opposite helix strand, as it was expected regarding the *coiled coil* pattern. Instead, the side chain interacts with Glu4 on the same peptide strand. Glu36 does not interact with any positively charged amino acid side chain. If this is the case for the model polypeptide, Glu36 can compete with Glu29 for salt bridge formation with Lys8. Furthermore, Lys8 can form an intrahelical salt bridge with Glu4 in an *i*, *i*-4 manner and Glu29 can form an intrahelical salt bridge with Lys32 in an *i*, *i*-3 manner. Both types of intrahelical charged interactions are very well known to stabilize helical structures in *coiled coils*[201] but may destabilize the dimer when competing with the formation of interhelical salt bridges.[202] In both of the described cases, the interhelical ionic interaction between Lys8 and Glu29 would not be present in the *coiled coil* folding motif. Consequently, Lys8 could not be applied as a substitution position for fluoroalkyl-substituted amino acids, since amino acid substitutions would not affect dimer stability. Thus, specific interactions could not be determined by the thermostability and self-replication screens.

In order to investigate this issue, molecular modeling studies were performed with the *coiled coil* screening polypeptide. Since no exact structure from x-ray crystallographic or NMR investigations were present, a peptide model had to be created. Starting from the x-ray structure of hepatitis δ antigen, an antiparallel *coiled coil* dimer with the amino acid sequence of the model polypeptide had been generated and the structure was energy-minimized and geometry-optimized. The final structure constituted the basis for molecular mechanic and molecular dynamic calculations on the ionic interactions of amino acid side chains Lys8 and Glu29.

The geometries of all side chains that may play a role in Lys8-Glu29 salt bridge formation (Lys8, Glu29, Arg1, Glu4, Lys32, and Glu36) have been varied systematically following a rational concept in order to evaluate all the ionic interactions these residues can participate in. Therefore, the side chain conformations were geometry-optimized by energy minimization and the most stable ones were compared (Figure 5.12). All the possible combinations of side chain conformations of the investigated salt bridges had to be considered.

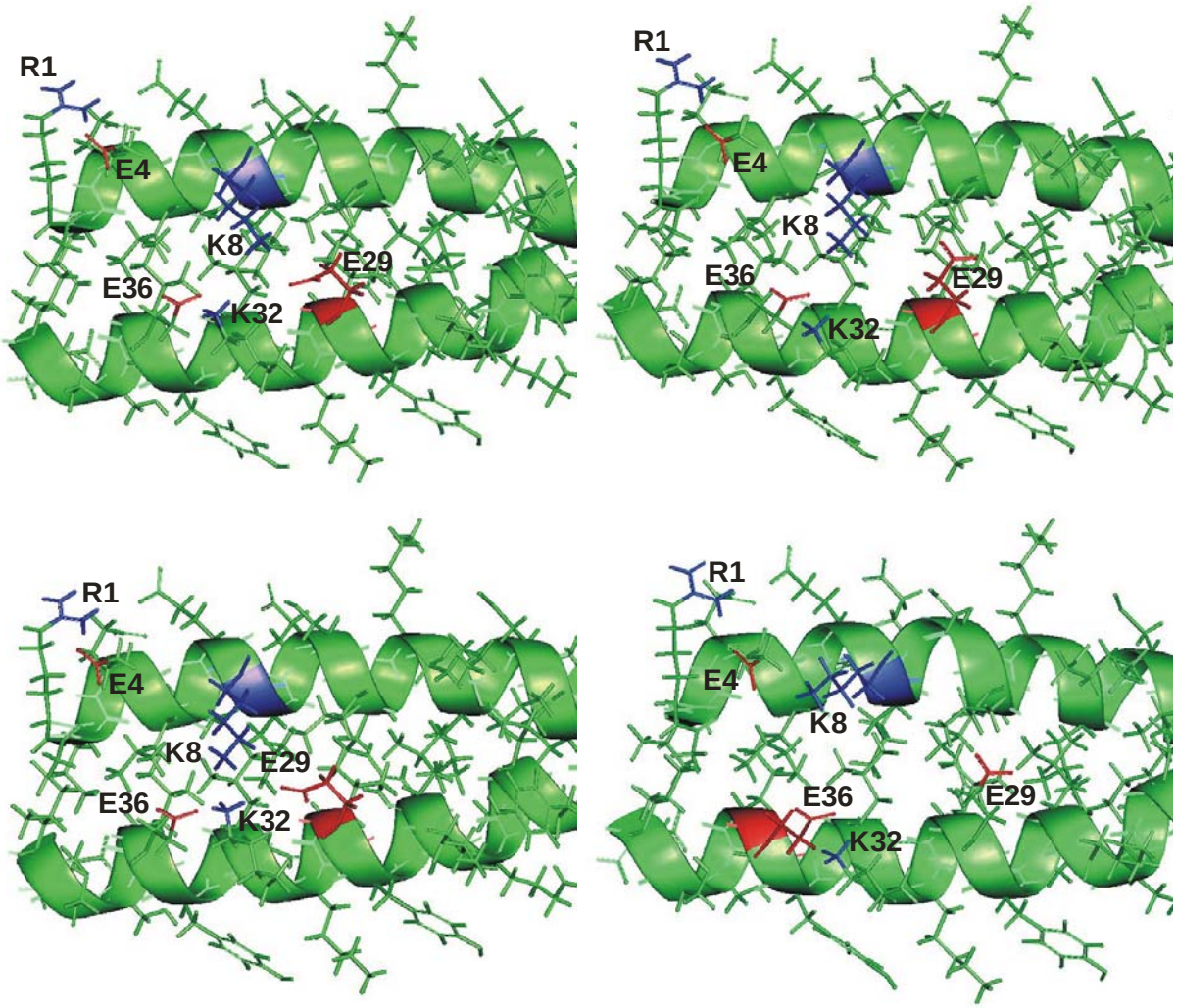


Figure 5.12: *Calculated ionic interactions of amino acids Lys8 and Glu29 (blue and red side chains and backbone ribbons, respectively) as well as charged amino acid side chains in proximity Arg1, Glu4, Lys32, and Glu36 (blue and red functional groups) within the model polypeptide. The helix backbones of the coiled coil protein are represented by green ribbons.*

In the conformation of lowest energy identified by these calculations, the salt bridge Lys8-Glu29 is formed (Figure 5.12, upper left image). In contrast, the interhelical salt bridge Arg1-Glu36 of the charged *coiled coil* interaction domain is broken by formation of intrahelical salt bridges Arg1-Glu4 and Lys32-Glu36. While the *coiled coil*-like interhelical Arg1-Glu36

interaction is destroyed by the increased flexibility of the ends of the α-helical peptide-strands, the g-g'-salt bridge of the next heptad repeat is formed. Obviously, the peptide conformation of the helical regions around Lys8 and Glu29, respectively, are much more rigid.

Two more peptide conformations were found that exhibited an energy that was only about 5kJ/mol higher than that of the most stable conformation of the protein model. In one of these structures, the salt bridge Lys8-Glu36 was found (Figure 5.12, upper right image) and within the other conformation, the functional groups of the four side chains Lys8, Glu29, Lys32, and Glu36 are oriented towards each other (Figure 5.12, lower left image). In both cases, Arg1 interacts with Glu4.

In all of the three most stable conformations that were found, Lys8 participates in charged interactions with at least one glutamic acid of the opposite helix strand, following the **g**-**g'**-*coiled coil* pattern. The calculations suggest that the salt bridge Lys8-Glu29 is formed within the model polypeptide. The peptide conformation of lowest energy with Lys8 interacting with Glu4 of the same helix strand possesses an about 37 kJ/mol higher energy compared to the most stable structure (Figure 5.12, lower right image).

These results approve the amino acid position Lys8 as substitution site for the investigations on fluoroalkyl-substituted residues in polar protein environments. Further evaluations on the interhelical salt bridge Lys8-Glu29 have been made regarding its influence on *coiled coil* stability as well as on self-replication rate of Lys8-substituted electrophilic peptide fragments. These experiments are described in the next two sections, dealing with the sensitivity of the thermostability and the self-replication screens.

5.2.2.2 The Sensitivity of the Screening Methods

The next step in proving the design concept of the screening system for the investigations of the interaction properties of fluoroalkyl side chains in native protein environments was the evaluation of the sensitivity of both the thermostability and the self-replication screens. Both screening methods detect changes in *coiled coil* dimer stability that are caused by amino acid side chain alterations in the substitution positions Lys8 and Leu9. Since single salt bridges in the polar interaction domain have a much smaller impact on *coiled coil* stability, compared to the size[203] and polarity[204] of amino acid side chains within the hydrophobic interaction core, the sensitivity of both screening methods regarding amino acid side chain alterations were tested by substituting lysine in position 8. This strategy was used in order to find out how the designed screening system responds to amino acid substitutions that cause only minor alterations in protein stability.

The Response to Minor Alterations in Protein Stability

A peptide variant was synthesized, that had Lys8 in the **g** position substituted by alanine (Lys8Ala). Therefore, an electrophilic peptide fragment was synthesized bearing the mentioned amino acid substitution (E_{K8A}). Self-replication experiments with this variant as well as thermostability studies with the ligation product (P_{K8A}) were performed, and the obtained results were compared with the results of the unsubstituted peptide P_o (Figure 5.13).

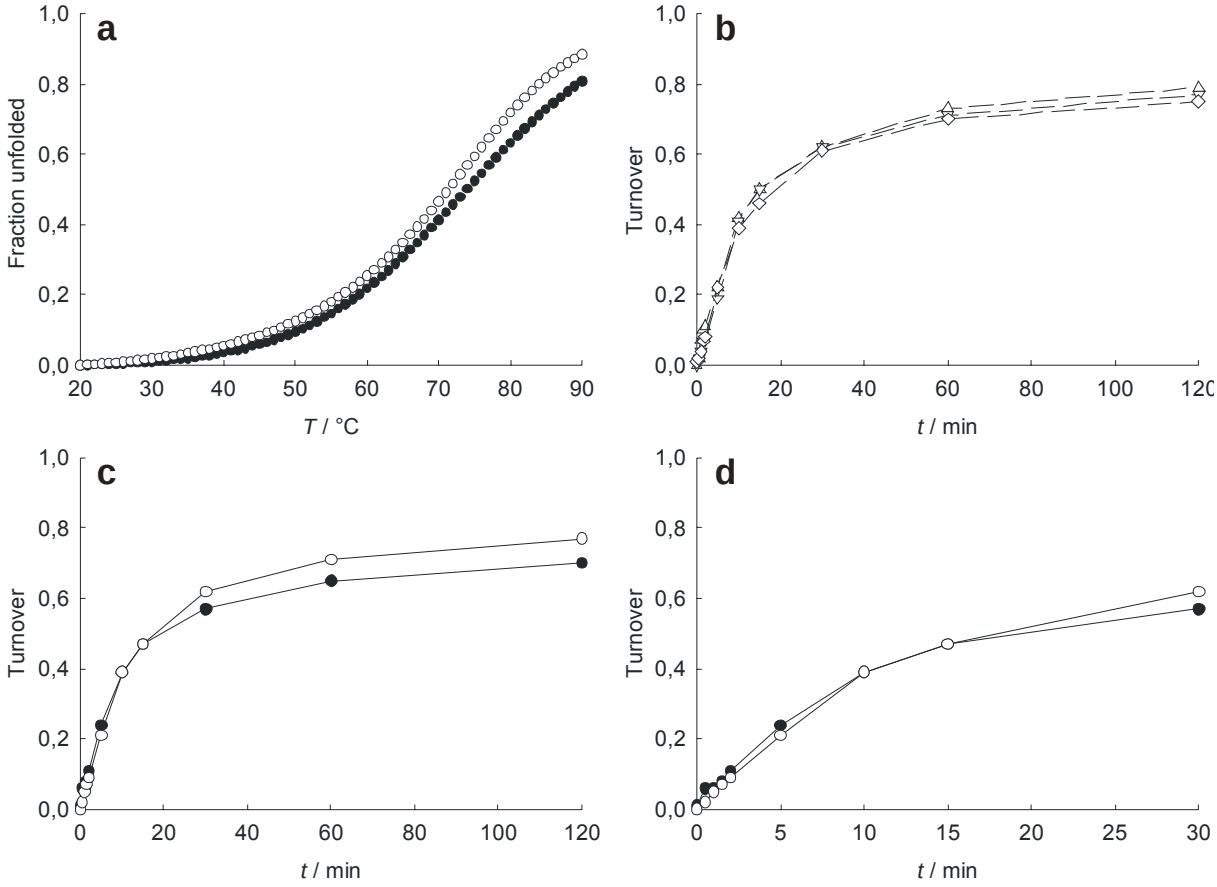


Figure 5.13: *Thermostability profiles **(a)** and average self-replication rates (c,d) of P_{K8A} (open circles) and P_o (closed circles). Average self-replication data of P_{K8A} are calculated from results of three independently performed experiments **(b)**.*

The melting profiles of P_o and its variant P_{K8A} (Figure 5.13a) show that the substitution of lysine by alanine in the **g** position of the *coiled coil* model polypeptide results in a slight destabilization of the folding motif. The calculated melting point of P_{K8A} (71.3°C) is 2.6 K lower than that of the unsubstituted dimer (73.9°C). Since the Lys8Ala amino acid substitution breaks the salt bridge Lys8-Glu29, it can be concluded that this interhelical ionic

side chain interaction stabilizes the *coiled coil* folding motif of the model polypeptide in the Lys8-unsubstituted variants.

This result does not appear to be surprising since interhelical attractive ionic interactions in the charged *coiled coil* interfaces are used in the *de novo* design of *coiled coils* to control the relative orientation of the helix strands by stabilizing the desired conformation.[,,] However, there are also examples of interhelical salt bridges in *coiled coil* peptides that destabilized the folding motif.[205] The reason for the inconsistent contribution of interhelical ionic interactions to *coiled coil* stability results from a fine balance between energetically favorable coulombic attractions and unfavorable desolvation of charges and conformational constraints of the residues involved in ion pairing. Therefore, proving the stabilizing effect of the Lys8-Glu29 salt bridge was a mandatory requirement for the application of Lys8 as substitution position for the investigations on the interactions of fluoroalkyl side chains in polar protein environments.

The results of the self-replication experiments show that this screening method is able to detect the Lys8Ala amino acid substitution (Figure 5.13c). However, the data reveal a more complex situation about the influence of the Lys8Ala substitution on protein folding, compared to its impact on stability of the *coiled coil* dimer. While the removal of the interhelical salt bridge led to an acceleration of the self-replication reaction after 15 minutes, in the first 10 minutes the replication rate of the alanine variant was slightly slower than that of the unsubstituted peptide. Obviously, a minor destabilization of the *coiled coil* dimer by alteration of the charged interface causes a diminished reaction rate at the beginning of the replication process. The reversed effect can be observed at later reaction times. In order to verify these interesting results, self-replication experiments with the standard electrophilic peptide fragment (E_o) were performed under *coiled coil*-destabilizing conditions (Figure 5.14).

The results of the self-replication experiments with the *coiled coil*-destabilizing guanidine hydrochloride show that rising concentrations of the protein-denaturating agent, which correlate with an increasing destabilization of the folding motif, led to a stepwise lowering of the replication rate until almost complete inhibition of the reaction in 2 M GdnHCl (Figure 5.14). However, the reaction in 0.25 M GdnHCl shows that a slight destabilization of the *coiled coil*, which is comparable to the removal of one interhelical salt bridge within the polar *coiled coil* interface, diminishes the reaction rate at the beginning of the replication, while an acceleration is observed at later reaction times. This finding proved the interpretation of the self-replication data of the Lys8Ala variant.

The self-replication cycle consists of protein folding, native ligation, and dissociation of the *coiled coil* dimer. While destabilization of the folding motif does not affect the chemistry of the ligation process, it influences its catalysis: *coiled coil* folding and unfolding. On the one hand,

the annealing of the peptide fragments to the template is inhibited, which decelerates the product formation. On the other hand, the dissociation of the ligation product is promoted, which accelerates the process.

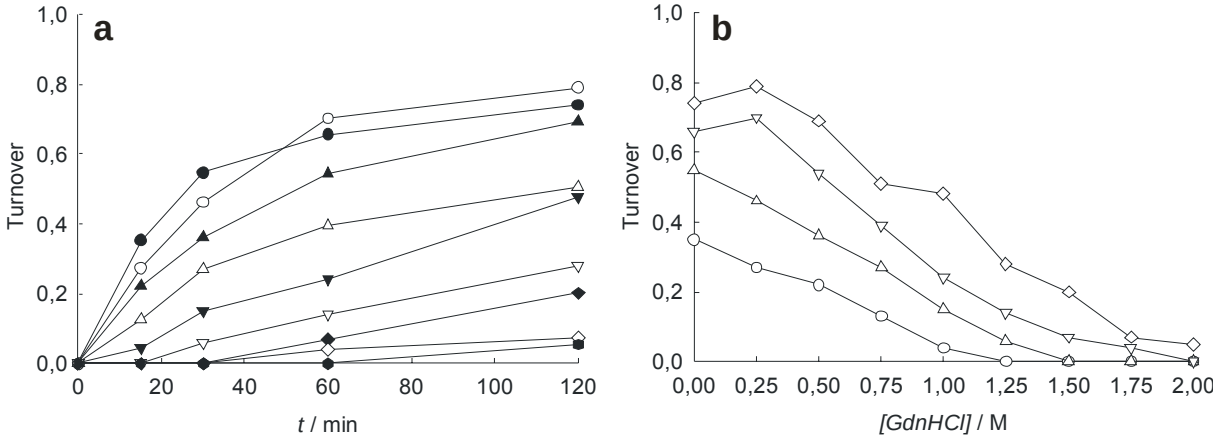


Figure 5.14: *Self-replication experiments of the standard electrophilic peptide fragment (E_o) with the nucleophilic one (N) in the presence of GdnHCl in rising concentrations;* ***a)*** *reaction plots of different GdnHCl concentrations: w/o (closed circles), 0.25M (open circles), 0.5M (closed triangles up), 0.75M (open triangles up), 1M (closed triangles down), 1.25M (open triangles down), 1.5M (closed diamonds), 1.75M (open diamonds), 2M (closed hexagons);* ***b)*** *reaction plots of different times: 15min (circles), 30min (triangles up), 60min (triangles down), 120min (diamonds).*

Since the stability of the ternary complex of both peptide fragments and the template is much weaker compared to that of the dimer of two full-length peptide strands, stonger destabilizing conditions (0.5M GdnHCl and more) have a higher impact on the annealing process than on the dimer dissociation. Consequently, the reaction rate is diminished. This effect correlates with the degree of protein destabilization. At conditions of minor structure destabilization (0.25M GdnHCl; removal of an interhelical salt bridge) the situation seems to be more complex. Which effect, inhibited annealing or promoted dissociation, has a higher impact on catalysis, is obviously dependent on the ratio of peptide fragments and ligation product, which means the reaction progress. Although the interhelical salt bridge Lys8-Glu29 stabilizes the *coiled coil* dimer, it reveals no accelerating impact on folding kinetics of the full-length protein. This finding might be surprising but could be described for comparable ionic interactions in other *coiled coil* proteins.[206]

The sensitivity of the designed *coiled coil*-based screening system towards amino acid substitutions that cause minor alterations in dimer stability was proven and the response of

both screening methods characterized. Now it appeared to be interesting to sees how the system would respond to major decreases in protein stability.

The Response to Major Alterations in Protein Stability

Since the hydrophobic core in a *coiled coil* protein represents the main driving force for the formation of the folding motif and provides the major contribution to structural stability, a variant of the model polypeptide was chosen that possesses glutamic acid, a negatively charged amino acid, in the substitution position Leu9 within the hydrophobic core. The respective electrophilic peptide fragment was synthesized (E_{L9E}) and self-replication experiments as well as thermostability measurements of the ligation product (P_{L9E}) of E_{L9E} with the nucleophilic peptide fragment (N) were performed (Figure 5.15).

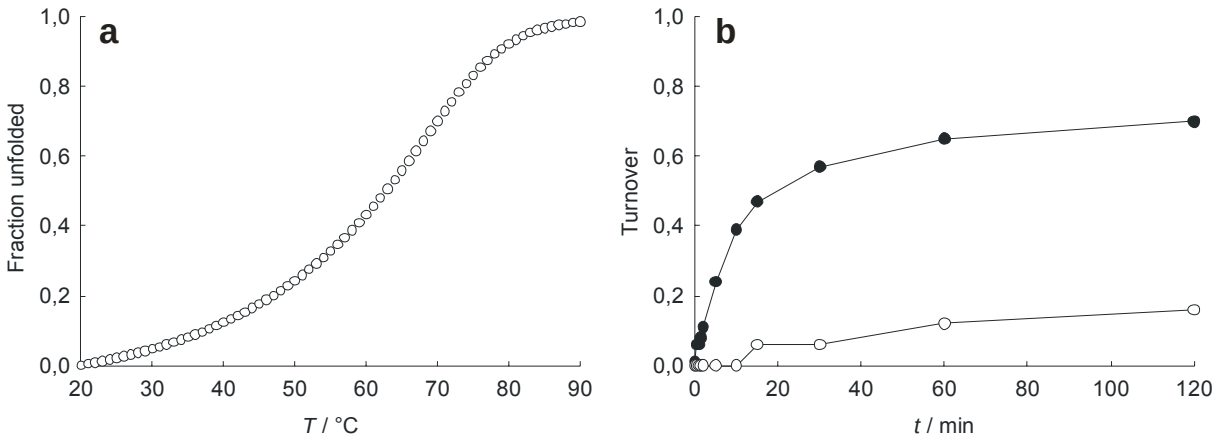


Figure 5.15: *Melting profile of P_{L9E}* ***(a)*** *and product formation of P_{L9E} (open circles) in self-replication experiments, compared to the reaction of the unsubstituted peptide P_o (closed circles).* ***b)*** *Thermal denaturation was performed in 2M instead of 5M guanidine hydrochloride.*

In 5M guanidine hydrochloride, P_{L9E} was already fully unfolded at 20°C. In contrast, the standard *coiled coil* dimer P_o could not be completely denaturated until 90 °C under the same conditions. Instead, a melting profile of P_{L9E} could be recorded in 2 M GdnHCl (Figure 5.15a). Obviously, a single amino acid replacement within the hydrophobic *coiled coil* interaction core can result in a relatively strong destabilization of the folding motif. Comparable to the results of the self-replication experiments with increasing concentrations of GdnHCl (Figure 5.14), the destabilization caused by Leu9Glu substitution led to a dramatic drop in self-replication rate (Figure 5.15b). These findings prove the *coiled coil*-based screening system

to react much more sensitively to substitutions in position Leu9 of the hydrophobic core, compared to those in position Lys8 of the polar dimer interface.

In order to evaluate the behavior of the screening system under extreme protein-denaturating conditions, self-replication experiments with the nucleophilic peptide fragment (N) and the standard electrophilic peptide fragment (E_o) as well as two Leu9-substituted analogues were performed in 6 M guanidine hydrochloride (Figure 5.16). The turnover plots show very slow reactions with less than 10% product yield after 20 hours. After 120 minutes less than 1% turnover could be interpolated for the peptide ligations. The curve linearity suggests an uncatalyzed ligation reaction. Thus, the guanidine hydrochloride in high concentration completely inhibits the reaction-catalyzing *coiled coil* formation. These results indicate that the uncatalyzed background reactions of the *coiled coil* self-replication experiments proceed very slowly and are negligible.

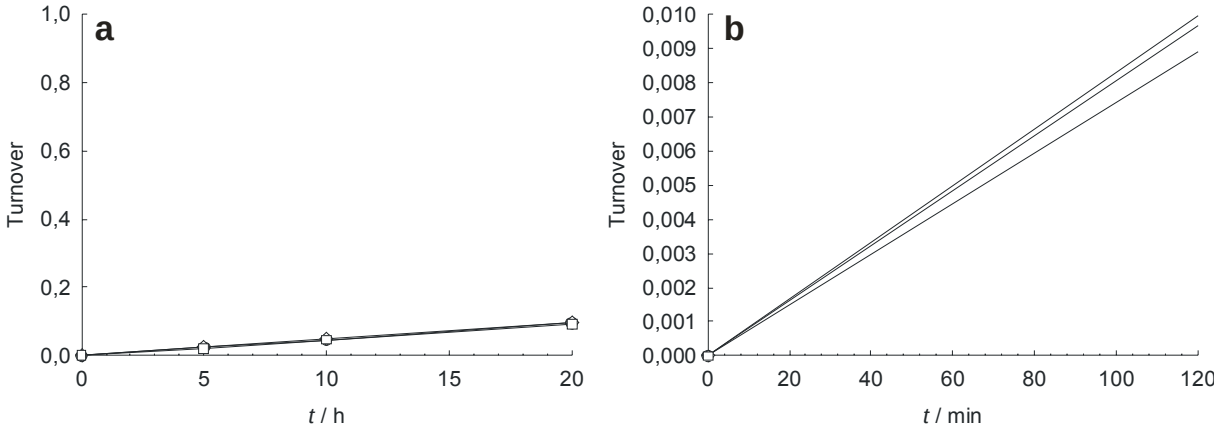


Figure 5.16: *Turnover plots of ligation-reactions of N with E_o and two L9-substituted analogues in 6M guanidine hydrochloride **(a)**. Since no product could be detected by HPLC below five hours of reaction, curves for the first 120 minutes of ligation are interpolated **(b)**.*

Consequently, differences in turnover rate of self-replication reactions between peptide variants that have amino acids substituted in positions Lys8 and Leu9 are caused by the influences of these residues on *coiled coil* folding and unfolding and can, thus, be attributed to the interaction properties of their side chains. This feature of the self-replication screen is an indispensible condition of its applicability for the evaluation of the properties of fluoroalkyl-substituted amino acids in native protein environments.

The thermostability and self-replication screening methods were optimized and the sensitivity of the *coiled coil*-based screening system towards amino acid substitutions in positions Lys8 and Leu9 was tested. Now it was possible to insert fluoroalkyl-substituted amino acids into

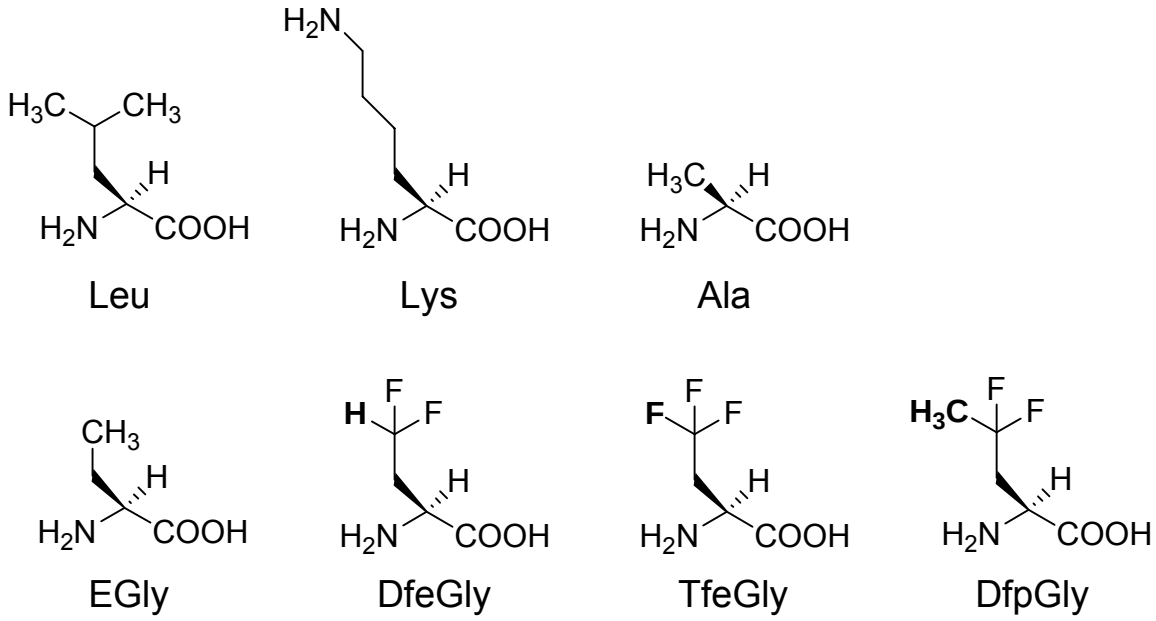
either of the substitution positions in order to investigate the impact of alkyl fluorination on the properties and molecular interactions of amino acid side chains within hydrophobic and polar native protein environments, respectively. Therefore, two classes of fluorinated peptide and protein building blocks were evaluated: side chain-modified L-amino acids, on the one hand, and fluoroalkyl-substituted $C^{\alpha,\alpha}$-dialkylated amino acids on the other.

5.2.3 Investigations on Side Chain-Fluorinated Amino Acids

Amino acids with systematically varied side chains have been screened in order to investigate properties of fluoroalkyl groups, such as space filling, hydrophobicity, polarity, as well as their interaction behavior with native and fluoroalkyl-substitiuted amino acid side chains, respectively. As a first step, an amino acid side chain had to be chosen as starting point for structural alterations, such as changes in chain length, branching, and content of fluorine atoms.

Incorporated into the substitution position 9 of the model polypeptide, the fluorinated amino acids replace the native residue leucine. Referring to the widespread argument that the spatial demand of a trifluoromethyl group is approximately as high as that of isopropyl, the side chain of *(S)*-4,4,4-trifluoroethylglycine is supposed to have space filling properties like native leucine. Therefore, the investigations started from *(S)*-ethylglycine (EGly). The first alteration of the EGly side chain made was the substitution of two hydrogens of the terminal methyl group by two fluorine atoms, resulting in *(S)*-4,4-difluoroethylglycine (DfeGly). In order to systematically enhance the steric demand of the DfeGly side chain, the remaining hydrogen of the terminal difluoromethyl group was either substituted by a third fluorine, which resulted in *(S)*-4,4,4-trifluoroethylglycine (TfeGly), or by a methyl group in the resulting *(S)*-4,4-difluoropropylglycine (DfpGly). A further canonical residue, alanine, was used as a control amino acid for the investigations (Scheme 5.1). The Fmoc-protected fluoroalkyl-substituted amino acids DfeGly, TfeGly, and DfpGly as well as alanine and EGly were inserted into the substitution positions Lys8 and Leu9, respectively, of the electrophilic peptide fragments. Self-replication experiments with these substituted electrophilic peptide fragments and the nucleophilic peptide fragment (N) were performed, and thermostability profiles of the ligation products were recorded.

As described in section 4.3.2, self-replication reactions of *coiled coil* peptides are influenced by multiple factors such as amino acid side chain interactions within the specific polar and hydrophobic interfaces as well as between the unfolded monomeric peptide fragments and full-length ligation products in the replication cycle. Consequently, side chain interactions between the non-natural building blocks in both of the substitution positions as well as their impact on the stability of the *coiled coil* interaction core will affect the turnover rate.



Scheme 5.1: *Stuctures of the native amino acids **(upper row)** as well as (S)-ethylglycine and its fluorinated derivatives **(lower row)** that are inserted into the substitution positions Leu9 and Lys8 of the model system.*

In order to obtain evaluable data from the self-replication experiments, the diversity of the substituted amino acids had to be limited. Therefore, only these peptides containing substituted amino acids of same side chain length but differing in fluorine content were investigated by the self-replication screen. The thermostability profile is mainly influenced by the impact of the physicochemical properties and interaction behavior of fluoroalkyl groups with native amino acid side chains on the stability of the specific *coiled coil* interaction domains in the model polypeptide. Therefore, thermostability screens of all of the substituted *coiled coil* variants were performed.

5.2.3.1 Interaction Properties in a Hydrophobic Protein Environment

The thermostability profiles as well as the turnover plots of the investigated Leu9-substituted *coiled coil* variants reveal that both of the screens are sensitive enough to discriminate between all of the screened peptides (Figure 5.17). Even the variants L9DfeGly and L9TfeGly yield distinguishable results from both screens, although the whole model polypeptides do not differ in more than two fluorine-by-hydrogen substitutions (one in each of both fluorinated amino acid side chains within the *coiled coil* dimer).[207] The melting profiles of the substituted *coiled coil* variants show that all substitutions in position 9 of the hydrophobic

core strongly destabilize the protein structure, compared to the original sequence (Figure 5.17a).

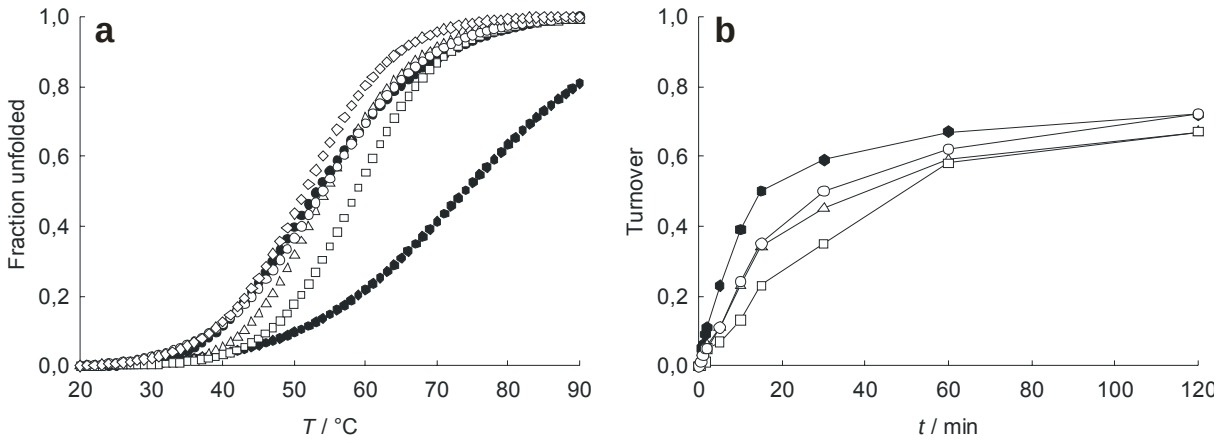


Figure 5.17: *Thermostability profiles **(a)** and self-replication turnover plots **(b)** of coiled coil protein variants with amino acid substitutions in position Leu9. Original residue (closed hexagons), Ala (closed circles), EGly (open circles), DfeGly (open triangles), TfeGly (open squares), DfpGly (open diamonds).*

The melting temperatures (T_m) of all substituted variants are lowered by 14.9-22.3K (Table 5.1). Considering the hypothesis that the steric bulk of TfeGly is approximately as large as leucine and the postulated enhancement of hydrophobicity of alkyl groups upon fluorination, an increase of structural stability should have been expected for the Leu9TfeGly substitution. The opposite result was received.

Table 5.1: *Melting temperatures of P_o and the Leu9-substituted variants.*

Protein	P_o	P_{L9A}	P_{L9EGly}	$P_{L9DfeGly}$	$P_{L9TfeGly}$	$P_{L9DfpGly}$
T_m **/ °C**[*]	73.9	53.2	54.0	54.4	59.0	51.6

[*] T_m is defined as the temperature at which 50% of the protein is unfolded.

From this finding it can be concluded that either the volume of a trifluoromethyl group is smaller than that of an isopropyl group or fluoroalkyl groups are less lipophilic than alkyl groups. The latter may be caused by the high electronegativity of the fluorine atoms that disturbs hydrophobic core interactions. Both issues are still controversially discussed in literature.

The incorporation of alanine into the hydrophobic interaction domain (P_{L9A}) causes a hole in the leucine core packing that strongly destabilizes the folding motif. The elongation of the Ala

methyl side chain by one methyl group (P_{L9EGly}) only slightly diminished this effect (ΔT_m = 0.8K). A stepwise increase in side chain volume by subsequent fluorination of the terminal methyl group in EGly further elevated the melting points of the dimer ($P_{L9DfeGly}$ and $P_{L9TfeGly}$). Very interestingly, this effect was only marginal for the difluoro analogue $P_{L9DfeGly}$ (ΔT_m = 0.4K), while the variant possesing just one more fluorine atom per fluorinated amino acid side chain ($P_{L9TfeGly}$) showed an increase in thermal stability by 5K. Unexpectedly, an additional enlargement of the TfeGly side chain by substitution of one of the fluorine atoms by methyl in DfpGly led to a dramatic loss of structural stability ($P_{L9DfpGly}$), compared to $P_{L9TfeGly}$ (Figure 5.18).

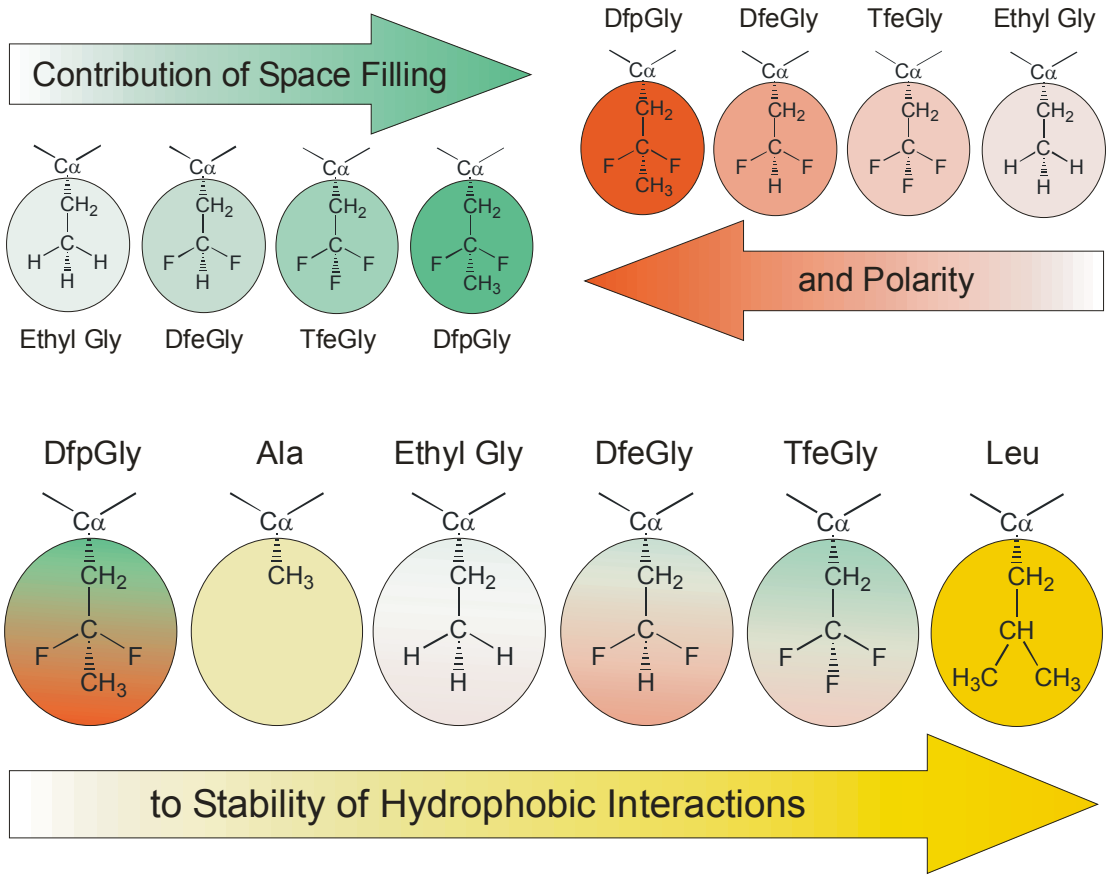


Figure 5.18: *Schematic representation of the influence of polarity and space filling of fluoroalkyl groups on stability of hydrophobic protein domains.*

These findings indicate that the space filling of the fluoroalkyl-substituted side chain cannot be the only property that affects hydrophobic core packing and the structural stability of this protein interaction domain. Consequently, the polarity of the systematically altered amino acid side chains plays an additional role.

Obviously, the strong electronegativity of the fluorine substituents causes a polarization of hydrogen atoms in all adjacent alkyl groups of the amino acid side chains that dramatically

disturbed hydrophobic core packing in the case of DfpGly. Therefore, the extent of destabilization increased along with the number of alkyl hydrogen atoms within the side chains that can be polarized by fluorine, which are two in TfeGly, three in DfeGly, and five in DfpGly. Although the latter carries the most space filling side chain of the investigated *(S)*-ethylglycine derivatives, the thermal stability of $P_{L9DfpGly}$ is lower than that of the alanine variant (P_{L9A}). This result clearly demonstrates that the electrostatic consequences of alkyl fluorination can have a much stronger influence on hydrophobic protein interactions than the increase in molecular volume upon stepwise fluorination. A recent report about polar interactions of a CF_3 group that stabilize peptide secondary structure and, thereby, overcompensate the destabilizing impact of enhanced steric demand, confirm the finding of this work.

The self-replication experiments of the electrophilic peptide fragments with original sequence (E_o) and EGly in position Leu9 (E_{L9EGly}) show that the moderate destabilization of the *coiled coil* folding motif that is caused by a decrease in space filling of the amino acid side chain in this substitution position of the hydrophobic core remarkably reduces the product formation rate (Figure 5.17b). This observation is in agreement with the results obtained from self-replication experiments with different concentrations of the secondary structure-disturbing agent guanidine hydrochloride (Section 5.2.2.2; Figure 5.14). Although the amino acid substitution disturbed the tertiary structure of the *coiled coil* rather than the secondary structure, a similar effect could be observed.

Consequently, a stepwise stabilization of the peptide dimer by subsequent fluorination of the EGly side chain in $P_{L9DfeGly}$ and $P_{L9TfeGly}$, that was concluded from the thermostability screens, should result in an acceleration of the self-replication reactions compared to the L9EGly variant. Interestingly, the opposite effect is found when analyzing the turnover plots of both reactions. Fluorination of the ethyl side chain led to an inhibition of the replication cycle, which cannot be caused by altered amino acid side chain interactions within the hydrophobic core. Since the observed effect correlates with the content of fluorine atoms in the amino acid side chain, it can be attributed to specific fluorine-fluorine interactions between substituted amino acids of the electrophilic peptide fragments as well as the ligation products in the unfolded monomeric state (Figure 5.19).

Such interactions between fluoroalkyl groups impede the α-helical *coiled coil* folding and the annealing of both peptide fragments with the full-length template and, therefore, inhibit the reaction cycle. Consequently, the “fluorous effect” causes a slower product formation which can be shown by turnover plots. Since these fluorous interactions do not directly interfere with the hydrophobic core of the *coiled coil*, they should also be detectable in fluorinated *(S)*-ethylglycine variants of Lys8 in the polar interface.

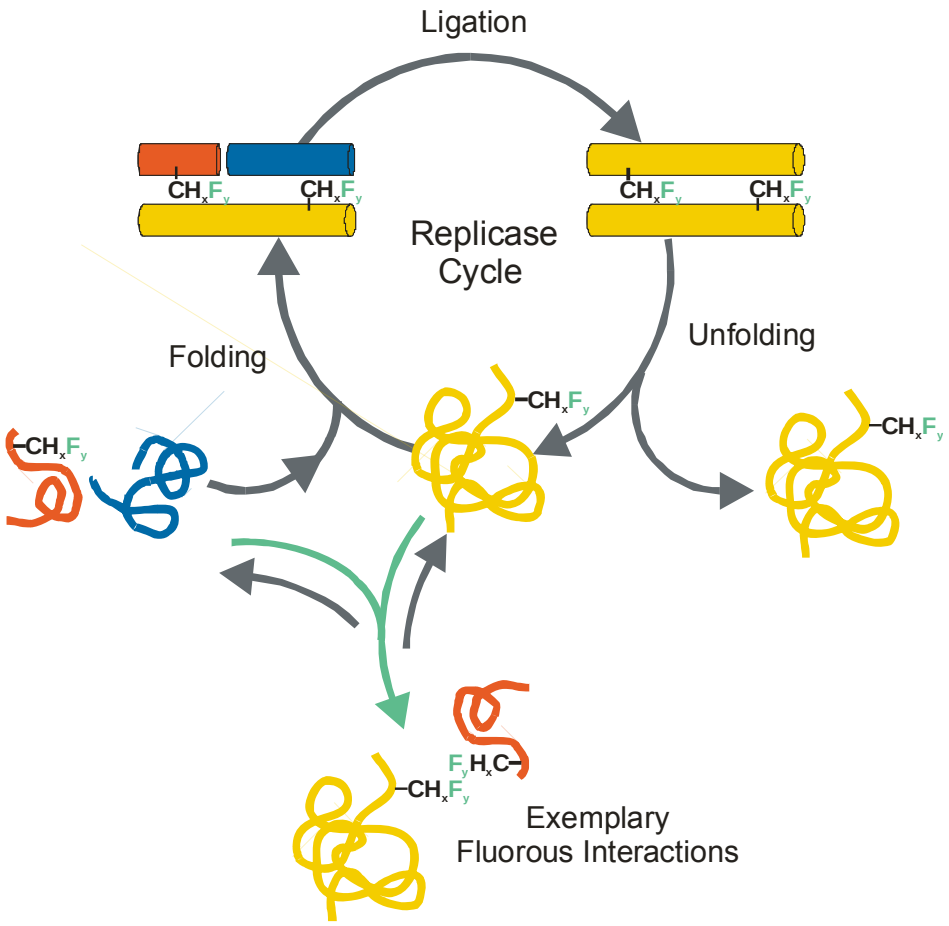


Figure 5.19: *Schematic representation of the self-replication reaction of the designed screening protein and the influence of fluorous interactions on the reaction cycle. Substituted electrophilic peptide fragments are shown in red, the nucleophilic peptide fragment in blue, and the ligation product in yellow. Amino acid side chain substitutions are marked with fluoroalkyl groups.*

5.2.3.2 Interaction Properties in a Polar Protein Environment

As expected, all substitutions of Lys8 in the polar *coiled coil* interaction domain by the described amino acids have a much lower impact on the structural stability of the α-helical *coiled coil* motif. Both the thermostability profiles and the turnover plots of the self-replication reactions of these variants show minor influences, compared to substitutions in position Leu9 of the hydrophobic core (Figure 5.20). These results confirm the investigations made on the sensitivity of the screening system by Lys8Ala substitution (Section 5.2.2.2; Figure 5.13). Both non-fluorinated variants P_{K8A} and P_{K8EGly} revealed a similar stability, which is represented by the melting points that are lowered by 2.0 and 2.6K, respectively, compared to the unsubstituted *coiled coil* dimer (Table 5.2).

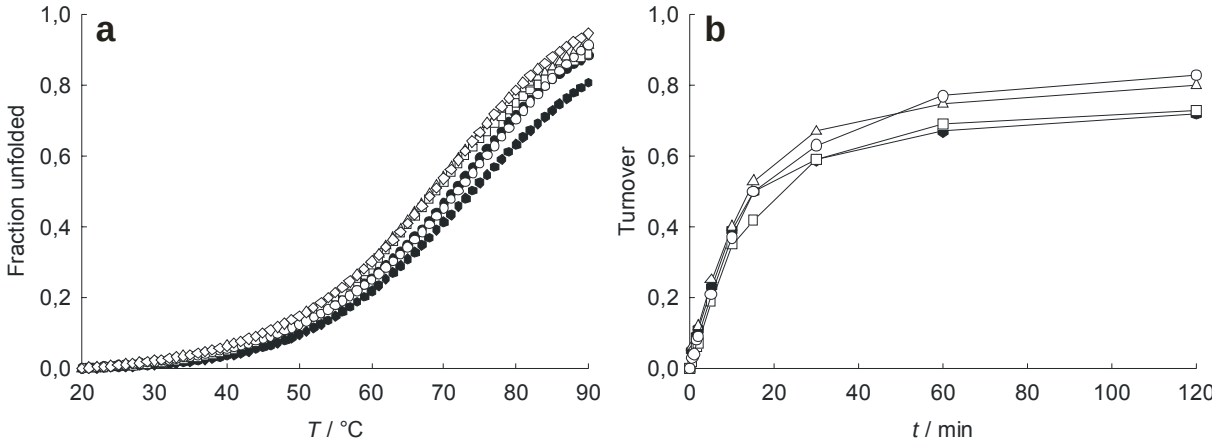


Figure 5.20: *Thermostability profiles **(a)** and self-replication turnover plots **(b)** of coiled coil protein variants with amino acid substitutions in position Lys8. Original residue (closed hexagons), Ala (closed circles), EGly (open circles), DfeGly (open triangles), TfeGly (open squares), DfpGly (open diamonds).*

Table 5.2: *Melting temperatures of P_o and the Lys8-substituted variants.*

Protein	P_o	P_{K8A}	P_{K8EGly}	$P_{K8DfeGly}$	$P_{K8TfeGly}$	$P_{K8DfpGly}$
T_m **/ °C**[*]	73.9	71.3	71.9	68.3	68.9	68.6

[*] T_m is defined as the temperature at which 50% of the protein is unfolded.

This weak destabilization of the *coiled coil* folding motif is caused by the loss of the stabilizing salt bridge Lys8-Glu29. The melting point difference of 0.6K between both dimers may be caused by the longer side chain of EGly compared to Ala, which can stabilize the folding motif by attaching to the hydrophobic core. Since the difference was marginal, this interpretation is highly speculative.

Remarkably, fluorination of the solvent-exposed EGly side chain in position Lys8 caused an additional loss of structural stability of the screening protein by 3.0-3.6K. Thereby, all three variants with fluorinated amino acids showed similar thermostabilities. Attractive polar interactions of the fluoroalkyl groups within the charged *coiled coil* interface should lead to a stabilization of the *coiled coil* dimer. In case of repulsive polar interactions between the carboxylate of Glu29 and the fluorine atoms of the fluoroalkyl groups, differences between the fluorinated EGly derivatives should have been observed. Since the side chain of DfpGly would orientate with the polarized terminal methyl group facing towards the negatively charged glutamic acid, a stabilization compared to P_{K8EGly} would be expected. The observed destabilization is caused rather by fluorine-fluorine interactions in the unfolded state of the peptide strands that have an impact on the monomer-dimer equilibrium. A more detailed

analysis is not possible due to the unsufficient resolution of the stability data for the Lys8 variants by this screen.

As expected, the turnover plots of the self-replication experiments show a reaction profile for the Lys8EGly variant that is comparable to the one of the Lys8Ala substitution (Figure 5.20b). In both self-replication reactions, only the loss of the salt bridge Lys8-Glu29 had an impact on product formation rate. Since both of the fluorinated ethyl side chains of DfeGly and TfeGly have a similar impact on *coiled coil* stability, their influences on the folding and unfolding processes within the polar interaction domain during self-replication and, therefore, the measured reaction rates should not differ. While a slight retardation of product formation can be concluded for the DfeGly variant compared to te reaction with EGly, TfeGly in position Lys8 led to a more significant drop of the reaction rate. Comparable to the self-replication reactions with the Leu9 substituted *coiled coil* variants, the observed effect correlates with the content of fluorine atoms and can, therefore, be attributed to the “fluorous effect”. Specific fluorine-fluorine interactions between substituted monomeric peptides in the unfolded state inhibit *coiled coil* folding and, thus, lower the self-replication rate (Figure 5.19).

Investigations of the Lys8DfpGly and Leu9DfpGly variants were the subject matter presented in a diploma thesis in our group.[208]

Recapitulating the results obtained from the investigations of the interaction properties of fluoroalkyl-substituted amino acid side chains in native protein interfaces, very interesting conclusions can be drawn.[209] The screening performed by systematic variation of the structure of fluorinated *(S)*-ethylglycine side chains provided new information about controversially discussed physicochemical properties of fluoroalkyl groups, such as space filling and polarity (Section 2.1). Furthermore, their interaction behavior in native protein environments, which comprises the controversy “increase in hydrophobicity versus fluorous effect” (Section 2.3.3):

- A trifluoromethyl group is not able to sufficiently mimic an isopropyl group in hydrophobic protein environments. Either space filling or lipophilicity of CF_3 or both properties were lower than for isopropyl.

- Fluorination of alkyl groups has opponent electrostatic and steric consequences on stability of hydrophobic protein domains. While the increase in amino acid side chain volume stabilizes hydrophobic cores, the polarity of the C-F bond and, most of all, the polarization of alkyl hydrogen atoms in proximity to C-F strongly disturb hydrophobic

interactions. Fluoroalkyl groups should prefer weak polar contacts over hydrophobic interactions.

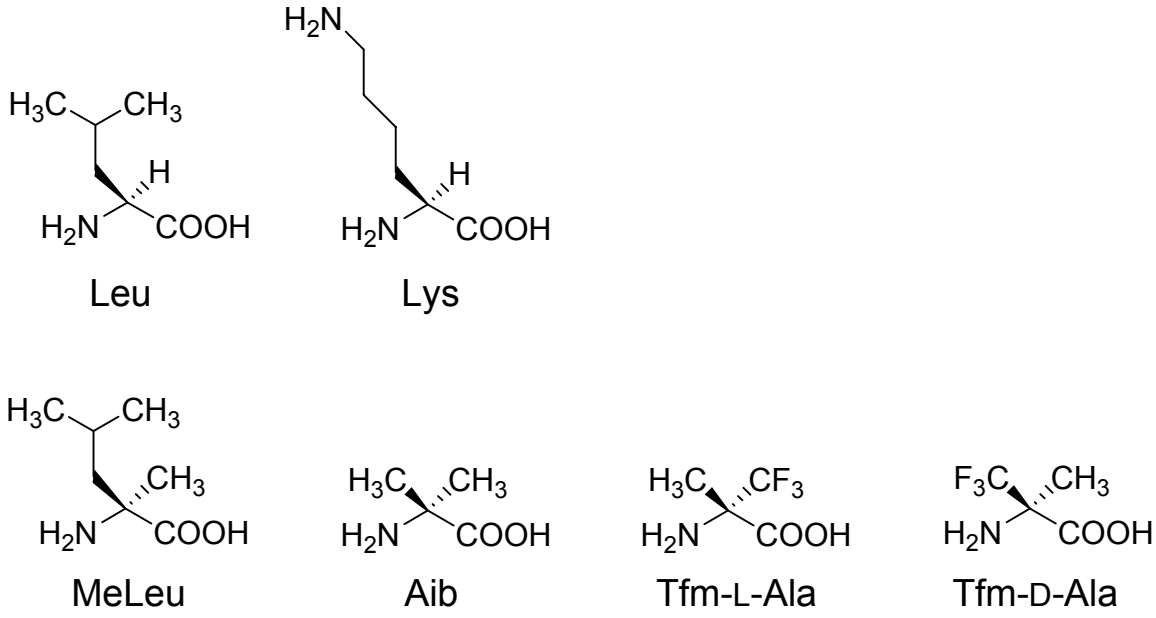
- Specific fluorine-fluorine interactions between fluoroalkyl-substituted amino acid side chains can affect protein folding which supports the theory of the “fluorous effect”.

The information obtained from these studies are supposed to be helpful for the application of fluoroalkyl-substituted amino acids in the rational design of peptide-base drugs with improved biological and pharmacokinetical features, such as increased *in vivo* stability and better binding properties. Enhanced metabolic stability of peptides by resistance against proteolysis can be achieved by the incorporation of $C^{\alpha,\alpha}$-dialkylated amino acids.[210] To combine both the resistance towards proteolysis and the unique interaction properties of fluoralkyl groups, fluorinated $C^{\alpha,\alpha}$-dialkylated amino acids can be used for peptide design. Therefore, properties and interaction behavior of this class of non-natural peptide and protein building blocks have been investigated by applying the developed *coiled coil*-based screening system.

5.2.4 Investigations on Fluoroalkyl-substituted, $C^{\alpha,\alpha}$-Dialkylated Amino Acids

To evaluate the interaction properties of fluorine-containing $C^{\alpha,\alpha}$-dialkylated amino acids, both enantiomers of α-trifluoromethylalanine (TfmAla) have been incorporated into both of the substitution positions Lys8 within the polar interface and Leu9 within the hydrophobic core of the *coiled coil*-based screening system (Scheme 5.2). In addition, the non-fluorinated analogue aminoisobutyric acid (Aib) was used as control in order to distinguish between the contributions of fluorine substitution and dialkylation, respectively, to the effect TfmAla exerts on stability and folding of the specific protein interaction cores. Furthermore, α-methyl-L-leucine (MeLeu) was investigated in position Leu9 to separate the dialkylation effect on hydrophobic core stability.

Electrophilic peptide fragments (E) containing the $C^{\alpha,\alpha}$-dialkylated building blocks in one of the substitution positions Lys8 and Leu9 were synthesized. Since TfmAla, like other fluorinated, $C^{\alpha,\alpha}$-dialkylated amino acids, cannot be incorporated into peptides via standard SPPS strategies, tripeptides with TfmAla in the middle position had to be synthesized via solution phase peptide synthesis (Section 4.1.2.2; Scheme 4.5).[211] In a following step, these tripeptides were incorporated into the appropriate position of the electrophilic peptide fragments on solid support via fragment condensation. For the syntheses of the tripeptides, racemic mixtures of TfmAla were used, and the resulting diastereomers were separated from each other at the dipeptide level using *reversed phase* HPLC.



Scheme 5.2: *Stuctures of the native amino acids **(upper row)** as well as the $C^{\alpha,\alpha}$-dialkylated building blocks **(lower row)**, that are inserted into the substitution positions Leu9 and Lys8 of the model system.*

Since the absolute configuration of the final electrophilic peptide fragments were not known, diastereomers were marked with Roman numbers ($E_{K8TfmAla(I)}$, $E_{K8TfmAla(II)}$, $E_{L9TfmAla(I)}$, $E_{L9TfmAla(I)}$) according to the order of elution during HPLC-based separation.Thermostability measurements of the full-length ligation products (P) as well as self-replication experiments of the substituted electrophilic peptide fragments (E) with the nucleophilic peptide fragment (N) were performed.Comparable to the investigations of fluorinated, side chain-modified amino acids (Section 5.2.3), only E-variants that contain amino acids of same structure but differing in fluorine content (TfmAla diastereomers and Aib derivatives) have been evaluated by the self-replication screen. This restriction was made in order to reduce the diversity of the substituted amino acids. The limitation is demanded by the high complexity of the factors that affect the self-replication cycle and, therefore, are reflected by the product formation rate.

5.2.4.1 Interaction Properties in a Hydrophobic Protein Environment

Following the amino acid sequence of the screening model, tripeptides Lys-TfmAla(I/II)-Glu were used for the production of $E_{L9TfmAla(I)}$ and $E_{L9TfmAla(II)}$. After syntheses of both diastereomers of electrophilic peptide fragments that contain TfmAla in position Leu9 in the hydrophobic core of the *coiled coil*-based model polypeptide, two main products of correct

molecular mass could be detected for $E_{L9TfmAla(I)}$ and $E_{L9TfmAla(II)}$ by HPLC and MALDI-MS. Obviously, racemization occurred during solid phase peptide synthesis.

Loss of optical purity at the α-carbon of the activated C-terminal amino acid via oxazolon formation during peptide segment condensation is a well known problem. It was, therefore, concluded that epimerization occurred at residue Glu10. Consequently, D-Glu was present in position 10 within the "wrong" diastereomers of the electrophilic peptide fragments. The well known fact that D-amino acids strongly destabilize right-handed α-helices was used to determine the configurations of Glu10 in all four $E_{L9TfmAla}$ variants.[212] Therefore, thermostability profiles have been recorded in 3M guanidine hydrochloride for their full-length ligation products (Figure 5.21).

In both cases, $P_{TfmAla(I)}$ and $P_{TfmAla(II)}$, the melting profiles clearly allow differentiation between the diastereomers obtained from peptide synthesis. Differences in melting points of 14.5K and 9.3K could be determined (Table 5.3).

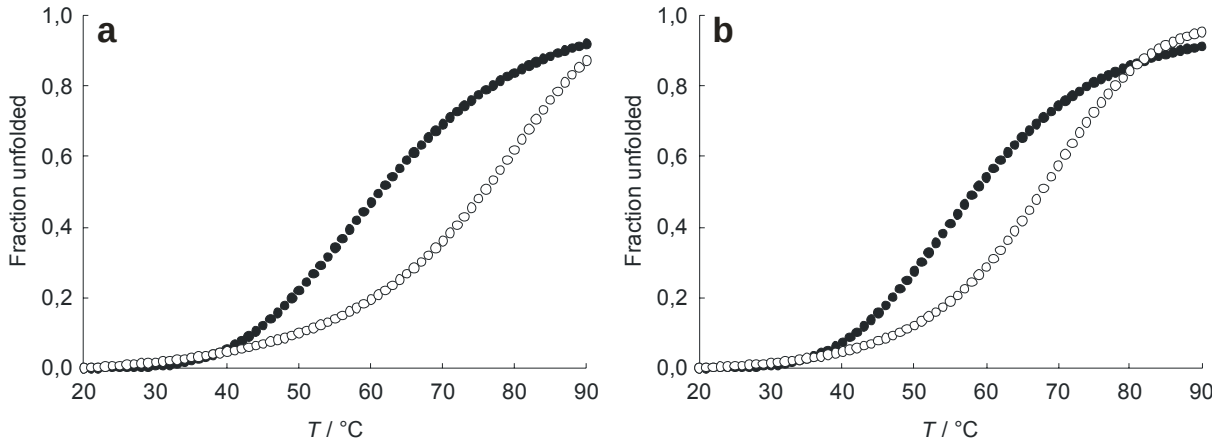


Figure 5.21: *Thermostability profiles in 3M GdnHCl of **a)** both Leu9TfmAla(I) and **b)** both Leu9TfmAla(II) diastereomers. Coiled coil variants, containing the L-amino acid in position Glu10, are represented by open circles the others by closed circles.*

Table 5.3: *Melting temperatures of all four $P_{L9TfmAla}$ diastereomers.*

Protein	$P_{L9TfmAla(I)L}$	$P_{L9TfmAla(I)D}$	$P_{L9TfmAla(II)L}$	$P_{L9TfmAla(II)D}$
T_m / °C*	75.7	61.2	67.7	58.4

* T_m is defined as the temperature at which 50% of the protein is unfolded.

It was concluded that the more stable diastereomers represent the "right" peptides that contain the L-glutamic acid in position 10. Consequently, the peptides $E_{L9TfmAla(I)}$ and $E_{L9TfmAla(II)}$ were used for further investigations.

The results of both the thermostability measurements and the self-replication experiments of the Leu9-substituted *coiled coil* variants were somewhat surprising (Figure 5.22).

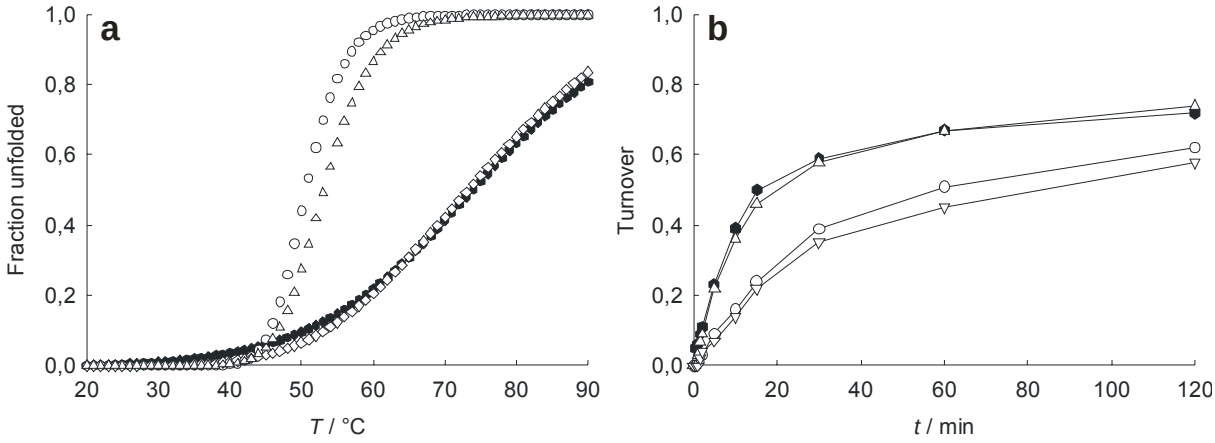


Figure 5.22: *Thermostability profiles **(a)** and self-replication turnover plots **(b)** of coiled coil protein variants with amino acid substitutions in position Leu9. Original residue (closed hexagons), Aib (open circles), TfmAla(I) (open triangles up), MeLeu (open diamonds). No melting profile is shown for the TfmAla(II) substitution.*

The melting profiles show that the substitution of leucine in the hydrophobic core by Aib and both TfmAla enantiomers strongly destabilize the *coiled coil* structure (Table 5.4), while $P_{L9TfmAla(II)}$ was almost completely unfolded in 5M GdnHCl at 20°C. Therefore, no melting curve could be recorded for this *coiled coil* variant. The denaturation profiles of P_{L9Aib} and $P_{L9TfmAla(II)}$ showed sigmoidal melting curves but also revealed a very weak CD_{222} signal already at the beginning of the denaturation process.

Table 5.4: *Melting temperatures of P_o and the Leu9-substituted variants.*

Protein	P_o	P_{L9Aib}	$P_{L9TfmAla(I)}$	$P_{L9TfmAla(II)}$	$P_{L9MeLeu}$
T_m **/ °C**[*]	73.9	50.6	53.1	---	73.3

[*] T_m is defined as the temperature at which 50% of the protein is unfolded.

A destabilization compared to the unsubstituted P_o was expected since all three building blocks miss the voluminous side chain for effective hydrophobic core packing. However, the stability of P_{L9Aib} (T_m = 50.6°C) is even lower than for P_{L9Ala} (T_m = 53.2°C). It is well known that $C^{\alpha,\alpha}$-dialkylated amino acids like Aib stabilize native-like helical peptide structures in general and induce the formation of α- and 3_{10}-helices.[,213,214] Thus, $C^{\alpha,\alpha}$-dialkylated amino acids are supposed to have a *coiled coil*-stabilizing effect on the backbone of the helix strand and P_{L9Aib}

should be at least as stable as P_{L9Ala}. However, the opposite result was observed. Obviously, the impact of dialkylation on backbone folding results in an amino acid side chain conformation that is unfavorable for an efficient hydrophobic core packing.

Since D-amino acids destabilize right-handed helices due to the side chains that point into the helix core, it was concluded that the least stable of both $P_{L9TfmAla}$ diastereomers ($P_{L9TfmAla(II)}$) possesses the TfmAla enantiomer that has the larger group (CF_3) in position of the side chain in a D-amino acid (Tfm-L-Ala). Consequently, $P_{L9TfmAla(I)}$ has to contain Tfm-D-Ala in position 9 of the hydrophobic core. Due to the enhanced volume of the trifluoromethyl side chain in TfmAla(I), compared to methyl in Aib, which participates in hydrophobic core packing, the fluorinated variant caused less destbilization (ΔT_m = 2.5°C).

Interestingly, no destabilization occurred on substituting the native leucine by its $C^{\alpha,\alpha}$-dialkylated analogue MeLeu. Obviously, the *coiled coil*-destabilizing impact of the dialkylation is compensated by the hydrophobic core packing that forces the isobutyl side chain of MeLeu into an appropriate conformation and, therefore, has more impact on folding of the helix backbone. Although this substitution did not cause a structural stabilization of the *coiled coil* protein, the dialkylation provides resistance towards proteolytic degradation and, thus, this non-natural building block may be a powerful tool in the design of *coiled coil*-based pharmaceuticals such as inhibitors of HIV-entry.[,,]

As expected, strong destabilization of the *coiled coil* structure by Leu9Aib substitution led to a significant deceleration of the self-replication reaction (Figure 5.22b). Considering the additional loss of stability for the L9TfmAla(II) variant and the “fluorous effect” that could be proven by earlier investigations on fluorinated, side chain-modified amino acids (Section 5.2.3), a further inhibition of product formation for both L9TfmAla substitutions was expected. However, while this effect was only minor in case of L9TfmAla(II), the more stable of the two diastereomeric peptides even showed a significantly accelerated reaction profile, compared to the L9Aib variant. The impact of dialkylation on helix backbone conformation, which was detected by the thermostability screen of the ligation products, combined with fluoroalkyl-substitution within the hydrophobic *coiled coil* domain results in an acceleration of self-replication. Since a complex combination of factors influences the reaction cycle (hydrophobic core destabilization, fluorous interactions, altered backbone conformation by $C^{\alpha,\alpha}$-dialkylation), an explanation of this finding would be highly speculative and should not be issued without further investigations, e.g., with more $C^{\alpha,\alpha}$-dialkylated amino acid analogues that help to separate the effects of the mentioned factors.

5.2.4.2 Interaction Properties in a Polar Protein Environment

Considering the amino acid sequence of the *coiled coil* model polypeptide, the tripeptides Glu-TfmAla-Leu had to be synthesized and incorporated into the appropriate position of electrophilic peptide fragments that contain either of both enantiomers of TfmAla in amino acid position Lys8 of the polar *coiled coil* interface. Since amino acid coupling to the N-terminus of TfmAla is highly challenging due to steric hindrance and the weakened basicity of the amino function, it demands special synthetic strategies (Section 4.1.2.2). Unfortunately, the solution phase synthesis methods that were applied for the synthesis of Glu-TfmAla-Leu were incompatible with appropriate protecting group strategies for the side chain carboxy group of the glutamic acid residue. Therefore, Glu was replaced by Ala, the canonical amino acid with the highest helix propensity. Since this amino acid substitution took place in an **f** position of the heptad repeat pattern in the *coiled coil*, it made no impact on the stability of secondary or tertiary structure had to be feared.

The thermostability profiles of the model polypeptide variants, possessing $C^{\alpha,\alpha}$-dialkylated amino acids in position Lys8, showed that the substitution of lysine by aminoisobutyric acid resulted in a slight stabilization of the tertiary structure (Figure 5.23).

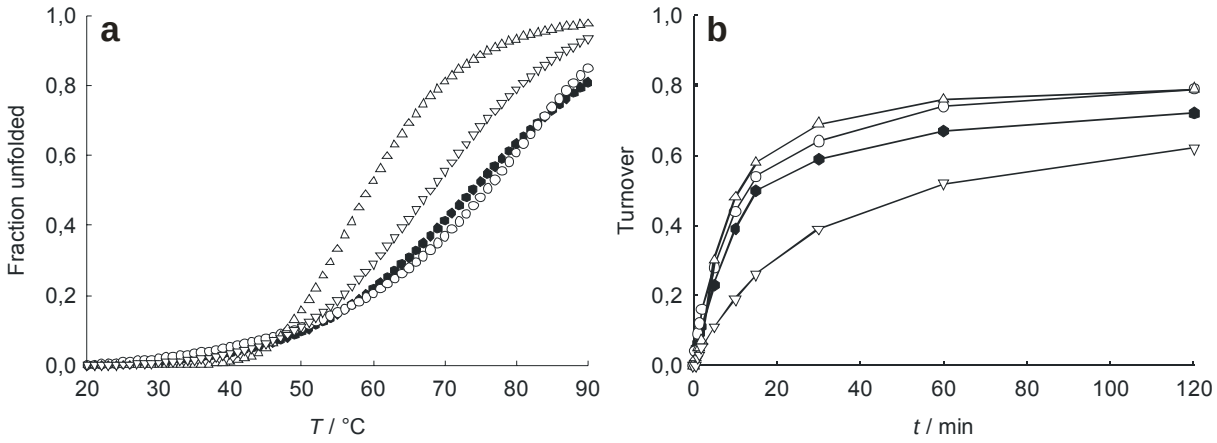


Figure 5.23: *Thermostability profiles **(a)** and self-replication turnover plots **(b)** of coiled coil protein variants with amino acid substitutions in position Lys8. Original residue (closed hexagons), Aib (open circles), TfmAla(I) (open triangles up), TfmAla(II) (open triangles down).*

In contrast to the destabilizing impact of Aib within the hydrophobic core, this kind of modification in the polar interface even overcompensates the destabilizing consequence of removing the salt bridge Lys8-Glu29 and elevated the melting point by 4.5K, compared to the K8Ala variant (Table 5.5).

Table 5.5: *Melting temperatures of P_o and the Lys8-substituted variants.*

Protein	P_o	P_{K8Aib}	$P_{K8TfmAla(I)}$	$P_{K8TfmAla(II)}$
T_m / °C*	73.9	75.8	59.3	68.0

* T_m is defined as the temperature at which 50% of the protein is unfolded.

It can, therefore, be concluded that $C^{\alpha,\alpha}$-dialkylation stabilizes *coiled coil* structure in general but has destabilizing consequences on core packing in the hydrophobic interaction domain. Both diastereomeric variants of $P_{K8TfmAla}$ are destabilized, compared to its unfluorinated analogue (P_{K8Aib}). Thereby, $P_{K8TfmAla(II)}$ is the more stable protein. According to the assignment of the absolute configurations of the TfmAla enantiomers in position Leu9 (Section 5.2.4.1), $P_{K8TfmAla(I)}$ contains Tfm-L-Ala and $P_{K8TfmAla(II)}$ contains Tfm-D-Ala. The melting point of $P_{K8TfmAla(II)}$ is comparable to the *coiled coil* variants that contain the fluorinated ethylglycine derivatives in substitution position Lys8. Therefore, the destabilization compared to the unfluorinated variant may as well be caused by fluorous interactions that shift the monomer-dimer equilibrium. Most probably, the same effect occurs during denaturation of $P_{K8TfmAla(I)}$, while an additional destabilization of the folding motif might be due to the α-helix-destabilizing impact of the bulky Tfm group in the position of the side chain in a D-amino acid.

The turnover plots of the self-replication reactions show an acceleration of product formation caused by Lys8Aib substitution, which is obviously due to helix stabilization during fragment annealing in the replication cycle. In contrast to the self-replication profiles of the peptides that were substituted by TfmAla in position Leu9 of the hydrophobic core, in the case of TfmAla incorporation in position Lys8, the less stable protein of both diastereomers revealed the higher reaction rate, while self-replication of the other is strongly inhibited. Comparable to both TfmAla diastereomers in the hydrophobic core, the results of the self-replication experiments are hard to interprete, since the contributions of *coiled coil* destabilization, fluorous interactions, and altered helix-backbone conformation by $C^{\alpha,\alpha}$-dialkylation are not distinguishable by the methods used here.

The results obtained from the investigations on $C^{\alpha,\alpha}$-dialkylated amino acids clearly demonstrate that this kind of peptide and protein building blocks strongly affects backbone conformation and, therefore, the folding of the molecule. It could be found that Aib in a position that is exposed to the solvent and MeLeu in the hydrophobic core contribute to structural stability of the *coiled coil* motif. These findings combined with the well known feature of $C^{\alpha,\alpha}$-dialkylated amino acids to inhibit proteolysis, prove these building blocks to be powerful tools for the design of biologically active *coiled coil* peptides, e.g., as inhibitors of

viral infection. The fluorinated Aib analogues, however, destabilize the α-helical *coiled coil* motif.

Interesting results have been obtained from the systematic investigations on the interaction properties of fluoroalkyl-substituted amino acids in native protein environments, applying the thermostability and self-replication measurements of the newly developed screening system. The new information that has been won may be helpful for the rational design of stable protein domains and improved biocative peptides. Fluoroalkyl-substituted $C^{\alpha,\alpha}$-dialkylated building blocks can be used for the formation of fluorous protein cores that reveal enhanced structural and metabolic stability due to perfluorination and dialkylation, respectively. The finding that fluorination polarizes alkyl groups is already very useful for the design of fluorinated peptide-based pharmaceuticals. However, more detailed information about the preferred native partners of specific fluoroalkyl groups for amino acid side chain interactions would enable the rational design of peptide-protein interaction sites with highly specific interactions between fluorine-containing, non-natural and native residues. This has been the aim of the second part of this thesis, a library screening applying the phage display methodology.

5.3 Selection of Preferred Interaction Partners for Fluorinated Amino Acids by Phage Display

The designed *stem loop* peptide with the randomized amino acid positions had to be displayed on the surface of the bacteriaphage M13 as pIII fusion (Section 4.4). Therefore, the DNA fragment that encodes for the *stem loop* sequence had to be inserted into the phagemid vector pComb3H to the 5'-end of the gene that encodes for the C-terminal part of the minor coat protein pIII (amino acids 230-406). After cloning of the *stem loop*, the DNA region that encodes for the peptide sequence with the amino acid positions to be randomized was cut out by specific restriction enzymes and a DNA cassette containing the library was inserted. Amplified phage that present the peptide library can then be used for the selection procedure.

5.3.1 Cloning of the *Stem Loop* Peptide

For the insertion of the DNA that encodes for the *stem loop* peptide into pComb3H, a cloning site of the restriction endonuclease SfiI is used. The main advantage of the Sfi I strategy is

the variable region within the palindromic recognition sequence of this enzyme (GGCCNNNN˅NGGCC) that can be used to design two different Sfi I restriction sites within the vector. A DNA-stuffer segment between both SfiI-sites is replaced by the peptide-encoding DNA. A self-ligation of the vector without the new DNA fragment is prohibited, since the ends that were produced by Sfi I digestion are not complementary. (Figure 5.24).

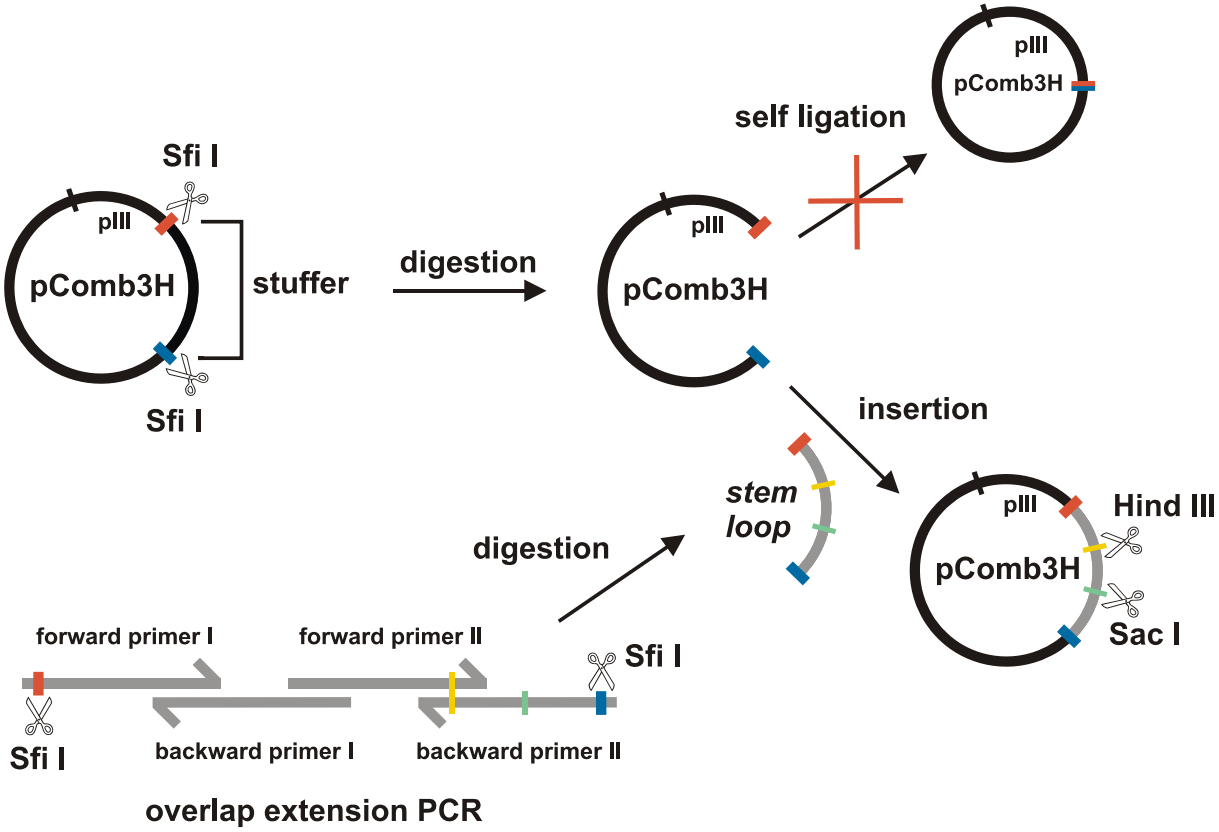


Figure 5.24: *Schematic representation of PCR and cloning strategy for the insertion of the stem loop-encoding DNA to the 5'-end of the truncated gene III into the phagemid vector pComb3H.*

The *stem loop* DNA contains restriction sites for the endonucleases Hind III (A˅AGCTT) and Sac I (GAGCT˅C) that surround the region of the six amino acid positions that have to be randomized in the library and two Sfi I sites at both ends that are complementary to the pComb3H restriction sites. The amino acid positions Leu33 and Glu34, which are placed within the library region that is replaced by a randomized fragment later, are each substituted by a stop codon to avoid translation of clones that possess the non-randomized library region. Since the designed *stem loop* DNA fragment is too long for complete synthesis by chemical means (262 nucleotides per strand), overlap extention PCR was used to produce the full-length DNA fragment from four primers, 80-90 nucleotides in length (Figure 5.24).

Since overlap extension PCR can be critical due to possible secundary structure formation of the long primer fragments, two different strategies were tested. In a one-step-procedure, one single PCR was performed with all four primers in same amounts. The second strategy, a two-step-procedure started with two PCRs: one extension reaction with forward primer I and backward primer I and another one with forward primer II and backward primer II. The purified products of both PCRs were then extended in a further PCR (second step) to yield the full-length *stem loop*-DNA. Very low yield was obtained from both procedures. Since less PCR cycles were performed during one-step strategy and the risk of mutation increases along with the rounds of amplification, the product of this reaction was used for cloning.
After digest of pComb3H and the *stem loop* DNA with Sfi I and purification, the insert was ligated into the phagemid vector. Several bacteria clones that were transfected with the ligated vector were selected and proved for correct cloning. The presence of the insert was tested by colony scans, and the right DNA sequence was proven by DNA sequencing. A clone with the correct *stem loop* DNA insert was chosen and its phagemid DNA was amplified for library construction, applying a maxiprep procedure.

5.3.2 Construction of the *Stem Loop* Library

For library construction, two complementary oligonucleotides were purchased that have all six library codons randomized by applying the NNK-strategy. Positions one and two within the codon can contain any (N) of the four nucleotides (A, C, G, T) and position three can contain G or T (K). The background of this strategy is the exclusion of two of three stop codons (TAA and TGA), while only TAG can occur in the library positions. The oligonucleotides were designed to form sticky ends upon annealing, complementing the ends of the phagemid vector that was digested with Hind III and Sac I in order to cut out the non-randomized library-fragment (Figure 5.25). After annealing of both randomized DNA strands, the library could be inserted into the *stem loop* DNA fragment of pComb3H.
Before the library construction procedure was optimized in order to maximize its size, background and diversity had to be tested. Therefore, small amounts of the library had been produced and transfected into bacteria. A ligation reaction without library DNA fragment was performed as a negative control to test the background of the library construction. The background consists of ligation reactions of the phagemid vector with traces of non-randomized DNA fragments that were not removed after digestion and vector self-ligation. Titering of the bacteria clones, which was performed by plating of the transfection reactions, revealed that the background reaction was very low (3,4 % of ligation with randomized DNA fragment).

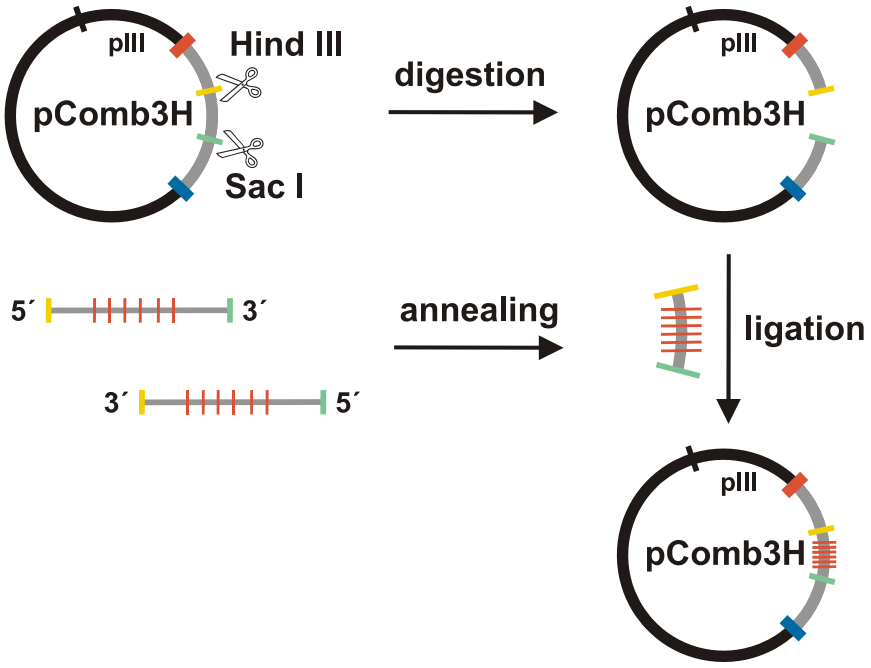


Figure 5.25: *Schematic representation of library construction and its insertion into the stem loop DNA fragment within the pComb3H phagemid vector. Randomized codons within the synthesized oligonucleotides are shown by thin red bars.*

20 bacteria clones of the ligation reaction were amplified and the insert of their phagemid DNA sequenced. 75 % of the tested clones had the right DNA sequence and an absolute diversity could be found within the six randomized amino acid positions (Table 5.6). Single mutations were found within the sequences of the remaining 25% bacteria clones. All of them resulted in frame shifts.

After a low background and good diversity could be proven, the maximum library size had to be determined. Since six randomized amino acid positions result in an overall variability of 6.4×10^7 different *coiled coil* protein fragments, an absolute library size of 6.4×10^8 had to be reached to obtain a sufficient chance for every member to be displayed on phage surface. Since a limited volume of vector DNA solution can be transfected by electroporation, the amount of correctly produced phagemid-DNA per sample has to be as high as possible. The efficiency of library construction depends on several factors that had to be optimized to reach maximum library size. Therefore, annealing and ligation reactions were performed with maximum electroporation volume and varied reaction parameters. The resulting library size was determined by titering transfected bacteria on agar plates.

The efficiency of the annealing and the high purity of the annealed products have a significant impact on the yield of the ligation reaction. One important factor, the concentration of the oligonucleotides, was optimized. Samples with DNA concentrations of each oligonucleotide between 75ng and 7.5µg per 100µL reaction volume were tested. Further

experiments revealed that the purity of the annealing product can be increased when the library cassette is incubated with the restriction enzymes Hind III and Sac I to achieve cleaner sticky ends. Precipitation and washing of the annealing product were also able to improve the procedure. Purification by agarose gelelectrophoresis failed since DNA fragments were too small.

Table 5.6: *Amino acids found in the library positions for the 15 non-mutated clones.*

clone	Glu29	Leu30	Leu33	Glu34	Glu36	Leu37
1	stop	Met	Ser	Val	Met	His
2	Glu	Ser	Gly	Thr	Arg	Ser
3	Trp	Glu	Trp	Val	His	Leu
4	Arg	Val	Ser	Lys	Arg	Pro
5	Leu	Leu	Tyr	Cys	Val	Ile
6	Asp	Ala	Asp	Leu	Val	Leu
7	Ala	Met	Met	Arg	Cys	Leu
8	Ala	Pro	Ser	Ala	Val	Trp
9	Leu	Lys	Thr	Lys	Tyr	Asp
10	Val	Val	Pro	His	His	His
11	Pro	Ile	Leu	Arg	Ala	Val
12	Asn	Glu	Leu	Phe	Glu	Arg
13	Ser	Leu	Pro	Ser	Ile	Ala
14	His	His	Val	Asp	Ser	Arg
15	Ser	Ser	Arg	Ala	Thr	Thr

The purity of the digested vector DNA as well as the ratio of phagemid and insert are most important for ligation efficiency. When optimizing these factors, it was found that a further purification of the phagemid vector by electroelution from the agarose gel and filtering through selective membranes improved the ligation reaction. The ligation of 1µg phagemid vector was tested with different amounts of annealing product between 300ng and 30µg in

order to find the optimal ratio for maximum library size. All test ligations for library optimization were infected with helper phage after transfection of bacteria, and preparations of produced library phage were made. The resulting phage libraries were combined and stored for test selection experiments.

Although all important factors that influence library size were optimized, the highest number of transfected clones that could be achieved for a single electroporation sample was 3.2 x 10^6. Consequently, 200 reactions had to be performed to obtain the required size of 6.4 x 10^8 library members, which is technically impossible. Since only one of both amino acid positions Lys8 and Leu9 in each of the electrophilic peptide fragments that are used for library screening is substituted by a fluorinated amino acid, either the library positions within the charged domain or the one within the hydrophobic core are screened. Consequently, two libraries can be constructed: one with the amino acid positions Leu30, Leu33, Glu34, and Leu37 randomized for screenings with Leu9-substituted electrophilic peptide fragments (variability of 1.6 x 10^5) and another with residues Glu29 and Glu36 randomized for screenings with Lys8-substituted peptides (variability of 400). This means that libraries with 1.6 x 10^6 and 4.0 x 10^3 members, respectively, would be sufficient. Both can be easily obtained by using the optimized procedure for library construction.

In the next step, the selection procedure, which is called panning, had to be optimized. Since these were test experiments with the original electrophilic peptide fragment, the combined phage library that was obtained from the library optimization experiments could be applied.

5.3.3 Verification of the Library Screening Strategy

Since phage presented library protein fragments may be proteolyzed during storage, phage libraries have to be reamplified before being used for panning. In order to evaluate the efficiency of this procedure, a reamplification test of the combined phage library was performed. Therefore, bacteria were infected with the library phage and new phage were produced after superinfection with helper phage. Reinfection of bacteria with phage preparation was titered on agar plates. 2 x 10^{10} phages were produced, which was three orders of magnitude higher than the size of the library to be reamplified.

However, DNA sequencing of 10 clones from the agar plates reavealed that 80% of the reamplified phage possessed the non-randomized *stem loop* gene with two stop codons in amino acid positions 29 and 30. Consequently, the library background was much better amplified compared to the phage with the randomized *stem loop*. Due to the stop codons, bacteria that are infected by background phage produce no *stem loop*-pIII fusion protein,

which seems to be beneficial for cell growth during phage amplification. Consequently, the phage library could not be reamplified but had to be produced right before panning.

Test panning with a new library (size: 1.2 x 10^7) was performed, using the non-substituted electrophilic peptide fragment (E_o) as target peptide. To monitor the background binding during selection, which comprises unspecifically bound phage that could not be removed by washing, a control panning was performed without target peptide. There are two main panning strategies: solution phase panning and solid phase panning. During solution phase panning, binding reaction between phage-displayed library and target peptide occurs in solution and bound phage is captured via an anchor molecule that is attached to the target peptide. No steric hindrance may disturb the binding, but unbound target molecules cannot be removed before capture and, therefore, compete with phage-bound target molecules for binding to the solid phase (Figure 5.26).

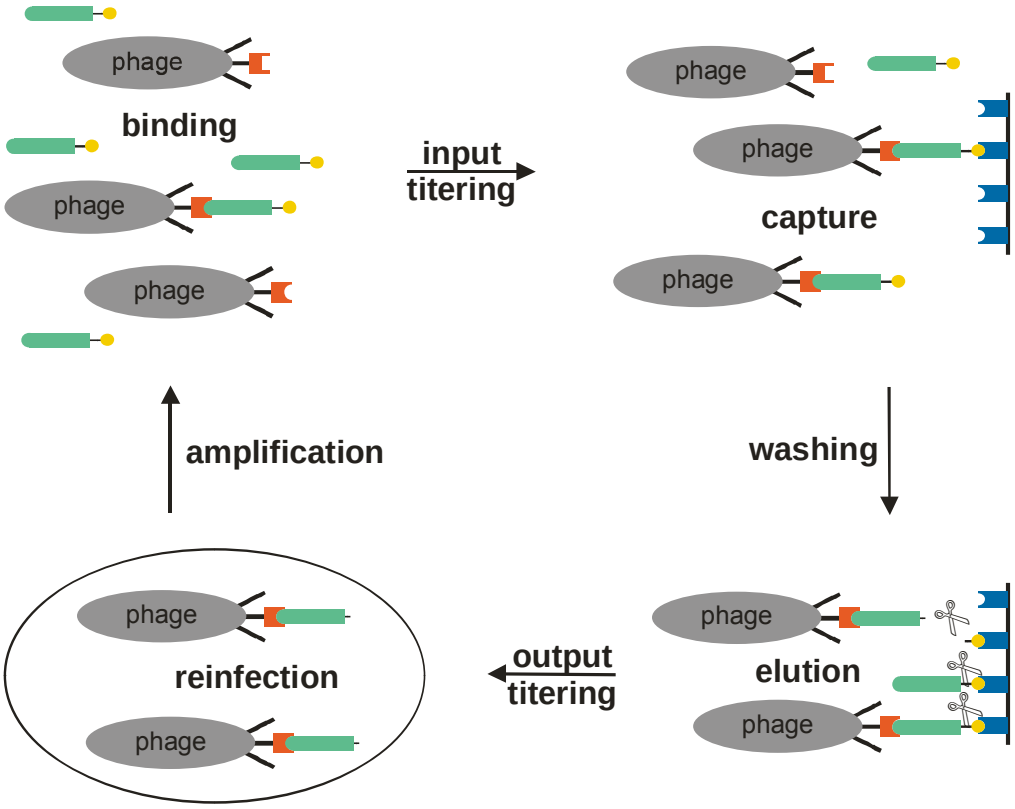


Figure 5.26: *Schematic representation of a solution phase panning procedure. Phage-displayed library is represented in red, screening peptide in green, biotin in yellow, and streptavidin in blue.*

In contrast, during solid phase panning, target molecules are immobilized first. After removing unbound target molecules, binding to the peptide library occurs on solid phase.

Screening of the *stem loop* library with the electrophilic peptide fragment was performed applying solution phase panning since hindrance of *coiled coil* formation during binding reaction on solid phase should be precluded.

Capture of the binding reaction occurred via binding of biotin (attached to the electrophilic peptide fragment) to streptavidine that was immobilized on magnetic particles. After washing and release from the magnetic particles by proteolysis, captured phage were used to reinfect bacteria. A preparation of the amplified phage selection was then used for further panning rounds. In each panning round, the number of phage that were given into binding reaction (input) and the number of phage that could be recovered from solid phase (output) were determined by bacteria infection and titering. After six rounds of selection and reamplification, the results of the test panning were analyzed. The input/output statistics show almost identical results for both the selection procedure with E_o and the control panning without target peptide (Figure 5.27).

The output values were relatively low and almost identical for both selections in all panning rounds. Although the output/input ratio of the E_o-panning continously increased, which is typical for the enrichment of specific phage clones during panning procedure, the same was found for the control-panning without target peptide. The same amounts of *stem loop*-displaying phage are captured with and without electrophilic peptide fragment. This result suggests that native ligation between E_o and the *stem loop* on phage surface did not proceed. This conclusion could be proven by analyzing single phage clones from the E_o-output titering of each panning round.

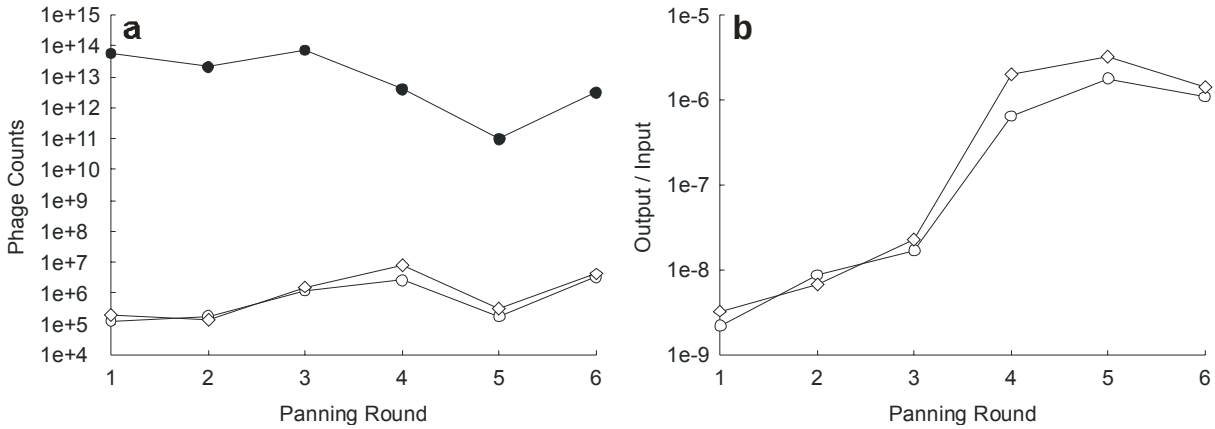


Figure 5.27: *Statistics of the test panning showing **a)** phage titers and **b)** output/input ratios of selection procedures with E_o (open circles) and without target peptide (open diamonds). Since the same phage preparation was used for both selections, input was determined only once each round (closed circles).*

A test PCR with DNA primers that flank the *stem loop* region within pComb3H was performed for five phage clones of each library production and the output of all panning rounds (Table 5.7).

Table 5.7: *Occurrence of DNA deletion during panning with E_o in %.*

library	*1*	*2*	*3*	*4*	*5*	*6*
0	0	0	80	80	100	100

The results show that after 3 and 4 rounds of panning 80 % of the screened clones had the DNA encoding for the *stem loop* protein deleted from the phagemid vector, while clones taken from the library production and after panning rounds one and two show no DNA deletions. After rounds five and six no intact vector DNA could be detected. Deletion of library DNA of the phagemid often occurs during panning procedures if no selection via binding of the displayed library peptide or protein to the target molecule takes place. Consequently, only background phage are captured and amplified. The deletion of the unprofitable foreign DNA provides an advantage to these phage, since particles with smaller genome inside are amplified more efficiently. In addition, the foreign DNA could be toxic to bacteria cells.

The sequencing results for the four clones of each library production and the output of panning rounds one and two revealed that at the beginning of the test panning before deletions occurred, preferably phage with stop codons in the randomized positions or mutations that disturb *coiled coil* structure were amplified. This result proves the finding from panning statistics and test PCR that no capture of phage by *coiled coil* formation and native ligation of the phage-bound *stem loop* fragment with E_o could be achieved during test panning.

The failing of the panning procedure was somewhat surprising since the successful presentation of both *coiled coil*-based *stem loop* proteins[215,216] as well as protein fragments with N-terminal cysteine residues for subsequent native ligation[217] as fusion of the minor coat protein pIII on the surface of filamentous phage M13 have been described in literature. It was concluded that the binding of the *stem loop* fragment to the electrophilic peptide fragment or the subsequent native ligation could not be accomplished under the conditions used for test panning. Consequently, these conditions had to be optimized before libraries could be screened further with electrophilic peptide fragments that are substituted with fluoroalkyl-substituted amino acids.

5.3.4 Optimization of the Phage Display System

In order to address the issue of the *coiled coil* binding and native ligation on phage surface and to optimize the panning procedure, binding tests of the biotinylated non-substituted electrophilic peptide fragment with a phage-displayed *stem loop* of the original amino acid sequence without randomizations had to be performed. Since the phagemid clone of the original *stem loop* sequence possessed two stop codons that encode for amino acid positions Glu29 and Leu30, a new *stem loop* gene was produced for binding tests and inserted into pComb3H that encodes for the complete protein fragment.

5.3.4.1 Optimization of the *Coiled Coil* Binding on Phage Surface

For the optimization of the binding of the electrophilic peptide fragment to the phage-displayed *stem loop* protein fragment, binding tests had to be performed with stepwise variation of binding parameters and conditions. One strategy for the quantification of binding in such binding tests is the elution of the captured phage with subsequent reinfection of bacteria and titering of infected bacteria clones. Consequently, each sample of an experiment has to be plated on at least one agar plate. Much faster and more efficient is the application of *enzyme linked immunosorbent assays* (*ELISA*). Thereby, an antibody that is conjugated with an enzyme binds the last molecule of a specific binding cascade. A color reaction of the enzyme is used for the quantification of the binding to be evaluated. The outstanding advantage of this technique is the possibility to perform many different binding tests in one experiment. The use of reaction plates with 96 cavities for separate binding tests (96-well plates) is most common.

The phagemid clone without the stop codons (denoted as “t” for template clone) was used to test the binding of the electrophilic peptide fragment (E). The original phagemid clone with the two stop codons (denoted as “o” for original clone), which produces phage particles that exclusively display wild type pIII on the surface, and a biotinylated target peptide of non-*coiled coil* amino acid sequence without a C-terminal benzylthioester (denoted as “P” for peptide) have been used as negative controls for the binding tests. For each varied parameter, the following phage/target binding combinations were tested:

(t/E) = positive sample; (t/P), (o/E), and (o/P) = negative controls

This strategy assured that unspecific binding of the phage particle as well as the electrophilic peptide fragments could be detected and separated from specific *coiled coil* binding and native ligation. A so-called *reversed phage ELISA* was used at the beginning of the tests. Thereby, the wells of the *ELISA* plate that specifically bind protein are coated with streptavidin. In the following binding cascade the target peptide binds to streptavidin via the conjugated biotin, phage binds to target peptide, and finally an anti-M13-antibody that is conjugated with horsereddishperoxidase (HRP) is specifically bound to phage. A color reaction of HRP quantifies the amount of bound phage. A *solution phase binding* strategy, which comprises binding of free target molecule and phage in solution and capture of the bound phage, was applied first (Figure 5.28). Since the target peptide is not anchored during binding reaction, *coiled coil* formation can proceed without any steric hindrance.

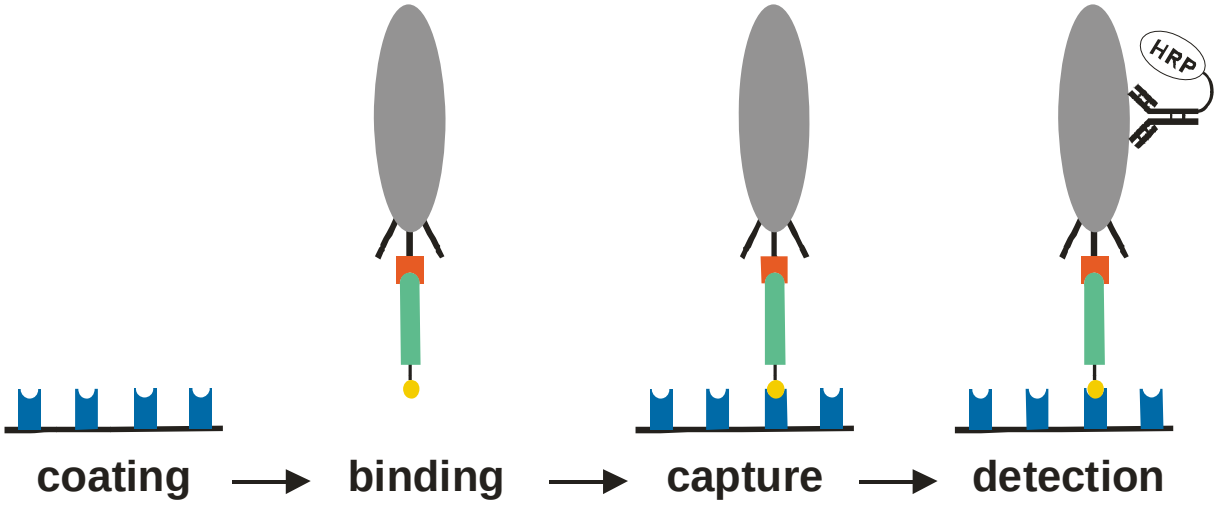


Figure 5.28: *Reversed phage ELISA with solution phase binding of phage and target molecule. Phage is represented in grey, stem loop part of the pIII-fusion protein in red, screening peptide in green, biotin in yellow, streptavidin in blue, and the antibody in black.*

Since the phage-displayed *stem loop* protein fragment can covalently dimerize via cystine formation at the N-terminal cysteine residue, disulfide reducing dithiothreitol (DTT) was added to each phage preparation. Several parameters have been varied in combinatorial binding tests, such as the concentration of the target peptide during *coiled coil* formation in the binding reaction (100nM – 1µM) and the duration of this reaction (1h – 24h), the use of protease inhibitors to avoid proteolysis of the *stem loop* during longer incubation times, and the concentration of the detergent Tween 20 in all washing steps (0.1% - 2%). Furthermore, the phage/target binding reactions were either directly used for capture on streptavidin-coated plates or phage particles were precipitated before capture and resolved to remove unbound biotinylated target peptides that would compete for the biotin binding sites of

streptavidin. The consistent results of these tests exhibited no specific binding of E_o with the phage-displayed *stem loop* fragment. Signals for this phage/target binding couple did not differ from the signals for the negative control under any of the tested conditions.

Since none of the parameter combinations resulted in the detection of specific *coiled coil* binding, the general protocol of the *reversed phage ELISA* was changed. A *solid phase binding* strategy was used, which comprises the immobilization of the biotinylated target peptide on streptavidin and a subsequent capture of the phage via phage/target binding reaction (Figure 5.29).

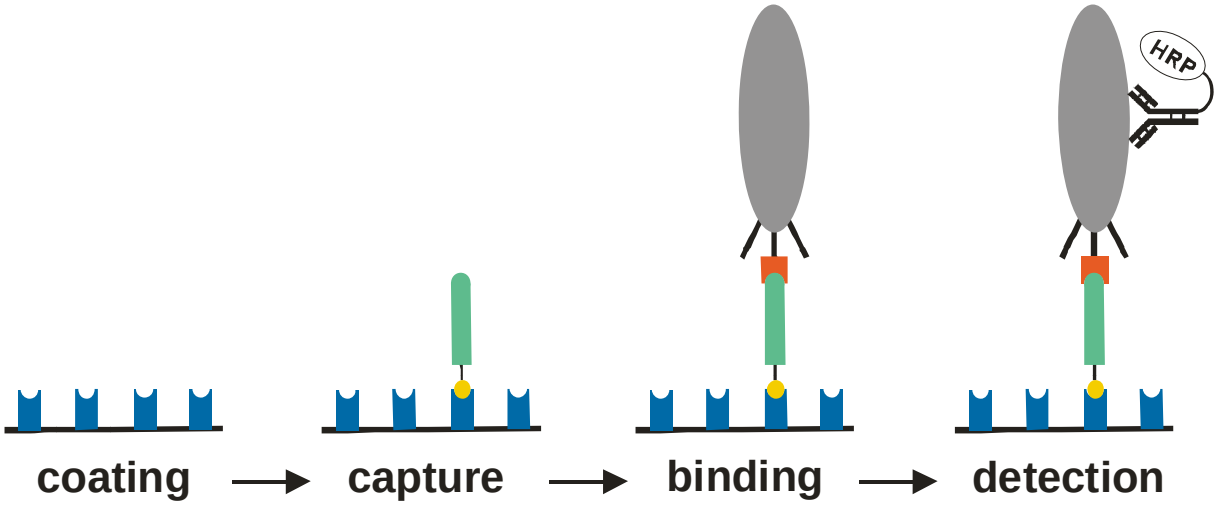


Figure 5.29: *Reversed phage ELISA with solid phase binding of phage and target molecule. Phage is represented in grey, stem loop part of the pIII-fusion protein in red, screening peptide in green, biotin in yellow, streptavidin in blue, and the antibody in black.*

In the binding assays with this new strategy, parameters were varied in a combinatorial manner as described for the *reversed phage ELISA* experiments with *solution phase binding* strategy. It could be found for all parameter combinations that samples with the electrophilic peptide fragment, (t/E) and (o/E), exhibit signals that are about three times the signal of the corresponding binding reactions with the non-*coiled coil* target peptide (t/P) and (o/P). Since E_o binds both the phage that presents the *stem loop* and the wild type phage, the binding occurs unspecifically to the phage coat. This unspecific binding may interfere with the *coiled coil* formation and, thus, had to be avoided.

Therefore, the experiments were repeated with detergent added to the binding reactions. However, neither Tween 20 (0.1% - 0.5% in washing solutions) nor guanidine hydrochloride (0.5M – 2M in binding reactions and washing solutions) could prevent this unspecific binding during *coiled coil* formation or remove unspecifically bound peptide during washing steps. Changes in temperature (RT or 37°C) did not help either. Since *coiled coil* formation should

be preferred over unspecific peptide/phage binding under conditions that destabilize protein structures (increasing GdnHCl concentrations), it was taken into account that *stem loop* protein fragments dimerize via *coiled coil* formation on phage surface right after release from the bacteria cell. As a result, the electrophilic fragment cannot compete with this strong binding.

In further binding experiments, phage were used which have either been incubated in 6M GdnHCl for different time periods (1h or 24h) or the phage preparation was directly made in 6M guanidine hydrochloride to provide *stem loop* monomers for binding reaction. It is described in literature that both phage particles[218] and the streptavidin-biotin complex[219] are stable at high concentrations of GdnHCl and neutral pH. Both *reversed phage ELISA* strategies, *solution* and *solid phase binding*, have been tested with different concentrations of GdnHCl added to the binding reaction (0.75M – 3M) and to the washing solutions (0.1% - 2%). Also other parameters were varied, such as binding time and concentration of the target peptide. However, no native ligation between *stem loop* and E_o was detected.

To prevent *stem loop* monomers from refolding during binding reaction after denaturation in 6M GdnHCl, a new *ELISA* strategy was applied. In the so-called *phage ELISA* an anti-M13-antibody is adsorbed to the plate wells and phage is specifically bound thereafter (Figure 5.30). Binding of the biotinylated target peptide should occur with the immobilized phage which, therefore, cannot dimerize via intra-*stem loop coiled coil* formation. The quantification of the bound target peptide is accomplished via HRP reaction. The enzyme is introduced as a streptavidin conjugate which specifically binds biotin.

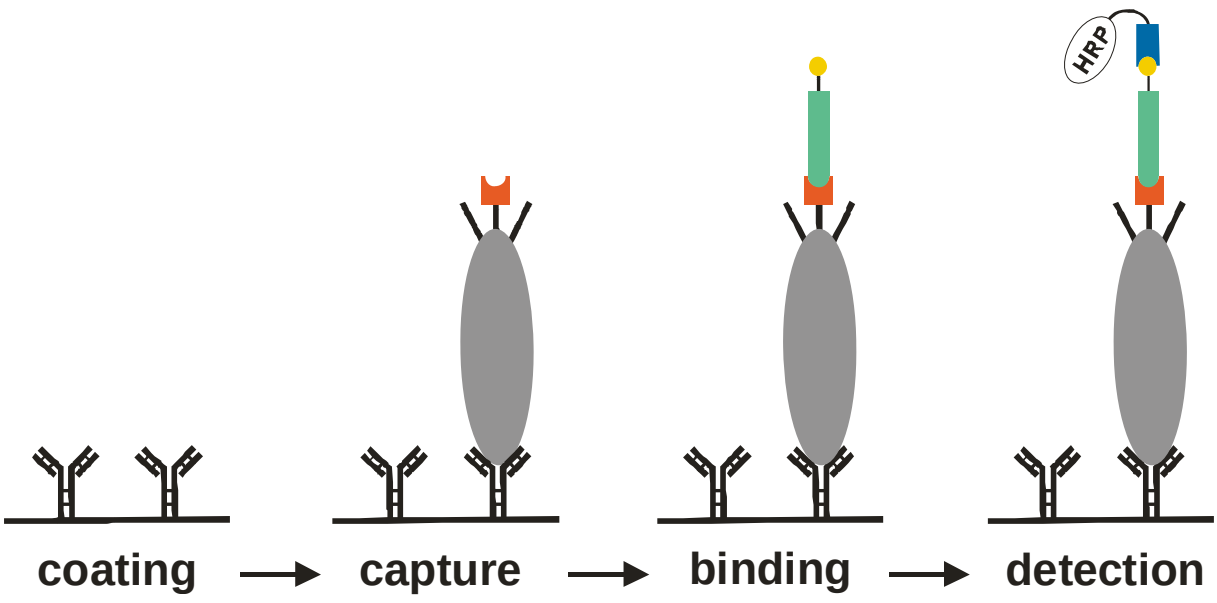


Figure 5.30: *Phage ELISA with solid phase binding of phage and target molecule. Phage is represented in grey, stem loop part of the pIII-fusion protein in red, screening peptide in green, biotin in yellow, streptavidin in blue, and antibodies in black.*

Since high amounts of guanidine hydrochloride may disturb the capture of the phage, this binding reaction was performed in different GdnHCl concentrations (0.25M – 6M). Again, higher signals for the (t/E) and (o/E) binding reactions, compared to (t/P) and (o/P) could be observed. It seemed that, once *stem loop* dimerization occurred, it could not be reversed.

For further binding tests, different amounts of target peptide (10nM – 10µM) were given directly into the bacteria cell culture during phage production and release. Thus, the electrophilic peptide fragment could bind and ligate to the *stem loop* fragment before homodimerization occurred. Either the supernatant of cell culture or precipitated and resolved phage have been used in *reversed phage ELISA* experiments with *solution phase binding*. Although different concentrations of detergent in the washing steps after binding reaction were used, no specific phage/target binding was detected.

It was concluded that either homodimerization via *coiled coil* formation was predominant over *stem loop*-E_o binding, even when the electrophilic peptide fragment was present during phage release, or the native ligation could not proceed for any other reason. Since the N-terminal cysteine residue of the *stem loop* is crucial for this reaction, *ELISA* tests have been performed to detect the thiol group of this amino acid on phage surface.

Therefore, maleimide was used to selectively bind the cysteine side chain instead of the electrophilic peptide fragment. In addition to the template phagemid clone (t), the original phagemid clone with two stop codons was used as a negative control. Two different strategies were applied for the maleimide-cysteine reaction: a *reversed phage ELISA* and a *phage ELISA* (Figure 5.31). In the *reversed phage ELISA* experiments, plate wells are coated with maleimide-conjugated BSA. After reaction of the free thiol group in the pIII-fusion protein, captured phage is detected via HRP-conjugated anti-M13-antibody and enzyme-catalyzed color reaction.

The *phage ELISA* tests started with the coating of the solid phase with anti-M13-antibody. After capture of phage, the cysteine side chain was covalently linked to HRP-conjugated maleimide. An enzymatic color reaction was used for quantification of phage surface-displayed free thiol groups. For the experiments phage were used that were prepared in 6M guanidine hydrochloride solution. Maleimide-phage reaction in the *reversed phage ELISA* and phage capture in the *phage ELISA* were performed in different concentrations of GdnHCl (0M – 6M) in order to provide access to *stem loop* monomers for sterically unhindered maleimide-thiol reactions. Neither of the *ELISA*-strategies were able to detect the free thiol group of the N-terminal cysteine residue in the pIII-*stem loop* fusion protein.

Since (t)- as well as (o)-phages have been reamplified several times for the binding tests that were performed, it had to be tested if the failure in *coiled coil* formation and thiol detection was caused by deletions or mutations within the (t)-phagemid genome that would lead to wild type phage production. Therefore, test PCRs with primers that flank the *stem loop* gene as

well as sequencing of both (t) and (o) phage clones were performed. The results could prove that all phage that were used in binding tests possessed the correct *stem loop*-encoding DNA.

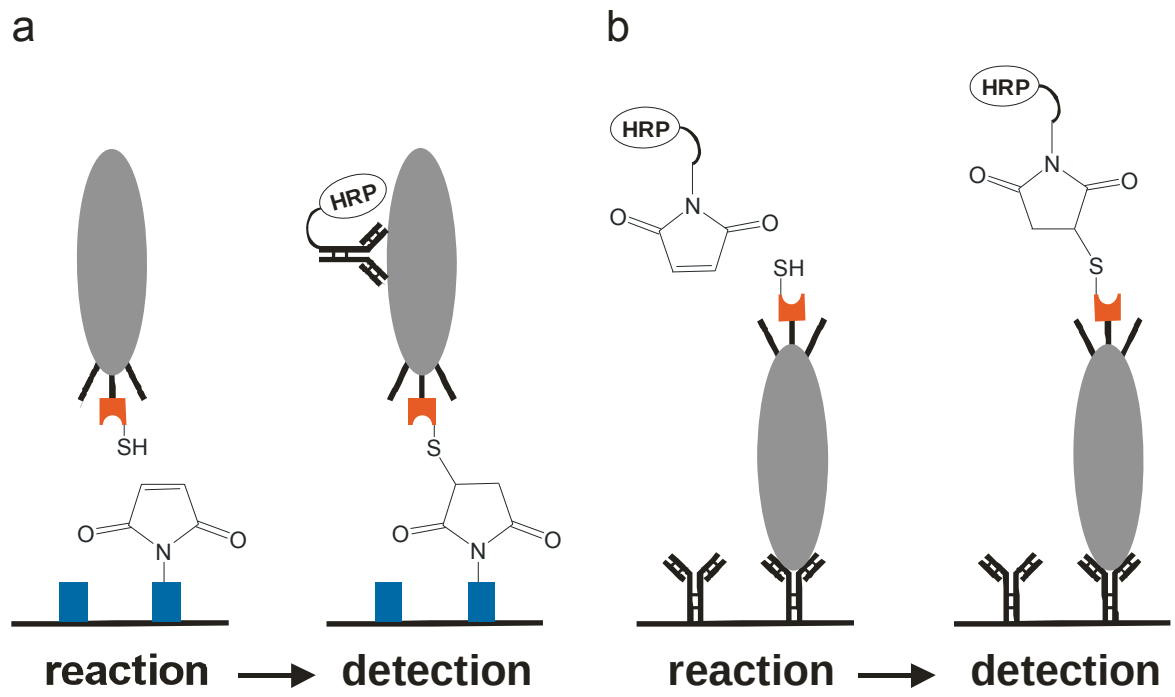


Figure 5.31: *Reversed phage ELISA* ***(a)*** *and phage ELISA* ***(b)*** *for the detection of free thiol groups using maleimide conjugates. Phage is represented in grey, stem loop part of pIII-fusion protein in red, BSA in blue, and antibodies in black.*

Two main reasons for the failure were taken into account: First, homodimerization of the *stem loop* protein fragment via *coiled coil* formation inhibits binding of the electrophilic peptide fragment and impedes thiol detection via steric hindrance. Second, the pIII-*stem loop* fusion protein is not presented on phage surface. During the phage assembly process, the same homodimerization of two *stem loop* fragments may prohibit the incorporation of the fusion protein into the phage coat. Instead, only wild type pIII is presented on phage surface. In both cases, the *stem loop* has to be redesigned in order to prevent the fusion protein from homodimerization.

5.3.4.2 Redesign of the *Stem Loop* Protein Fragment

In the original *stem loop* protein, the nucleophilic peptide fragment interacts with the N-terminal part of the template fragment via *coiled coil* formation. Thereby, negatively charged amino acid side chains in the polar *coiled coil* interaction domain of the nucleophilic fragment

form salt bridges with positively charged groups of the polar *coiled coil* interface in the template part (Figure 5.32a).

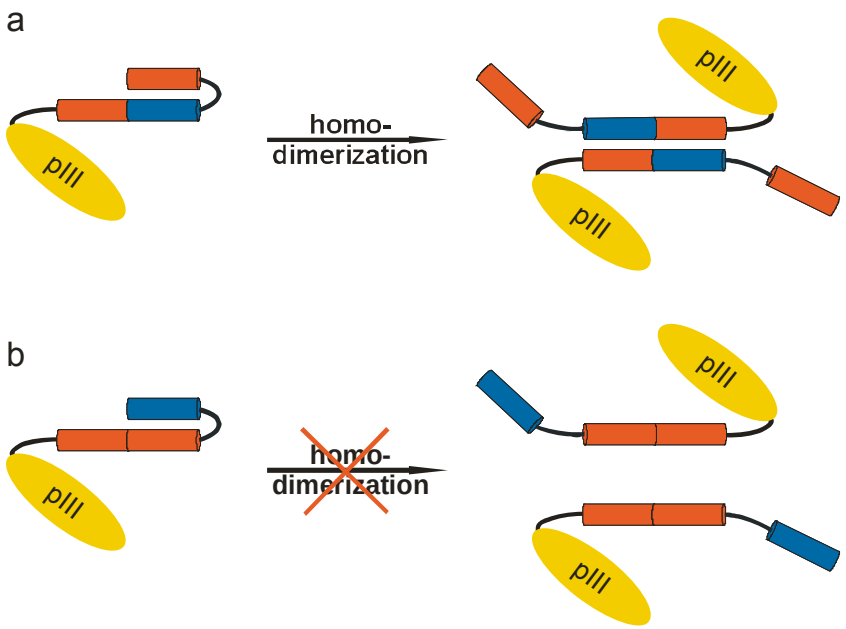


Figure 5.32: *Schematic representation of the design of the polar coiled coil interface and homodimerization behavior in pIII-stem loop fusion proteins of* ***a)*** *the original design and* ***b)*** *the redesign. Stem loop parts with negatively charged interface are represented in red, parts with positively charged interface are shown in blue. Minor coat protein pIII is shown in yellow.*

This *stem loop* can form an antiparallel *coiled coil* that is comparable in sequence and structure to the ligation product of N and E_o. This *coiled coil* dimer could be proven to be extremely stable and did not denaturate until 90 °C (Section 5.2.1.1). The polar interface of the C-terminal part of the template strand in the *stem loop* was not changed, since this domain interacts with the electrophilic peptide fragments, which are already synthesized. Consequently, the amino acid pattern of the polar *coiled coil* interface of the nuceophilic fragment part and the N-terminal portion of the template strand were varied. Therefore, negatively charged residues in **e** and **g** positions were exchanged by positively charged amino acids and vice versa (Figure 5.33). In the resultant *stem loop*, the entire template part possessed a negatively charged polar interface that is formed by positions **e** and **g** (Figure 5.32b). Due to a repulsion of the negative charges, a homodimerization of the pIII-fusion proteins via formation of hyper stable *coiled coil* structures was avoided.

Since both phage types (t) and (o) are needed for binding tests with the redesigned *coiled coil*, two *stem loop*-encoding DNA fragments were constructed and cloned into pComb3H at the 5'-end of geneIII (encoding for the C-terminal part of pIII). One DNA construct encoded for the correct *stem loop* sequence (denoted as "nt" for new temp) and the other possessed

two stop codons in amino acid positions Glu29 and Leu30 (denoted as “no” for new orig). Bacteria that were transfected with a (no) phagemid produce wild type phage.

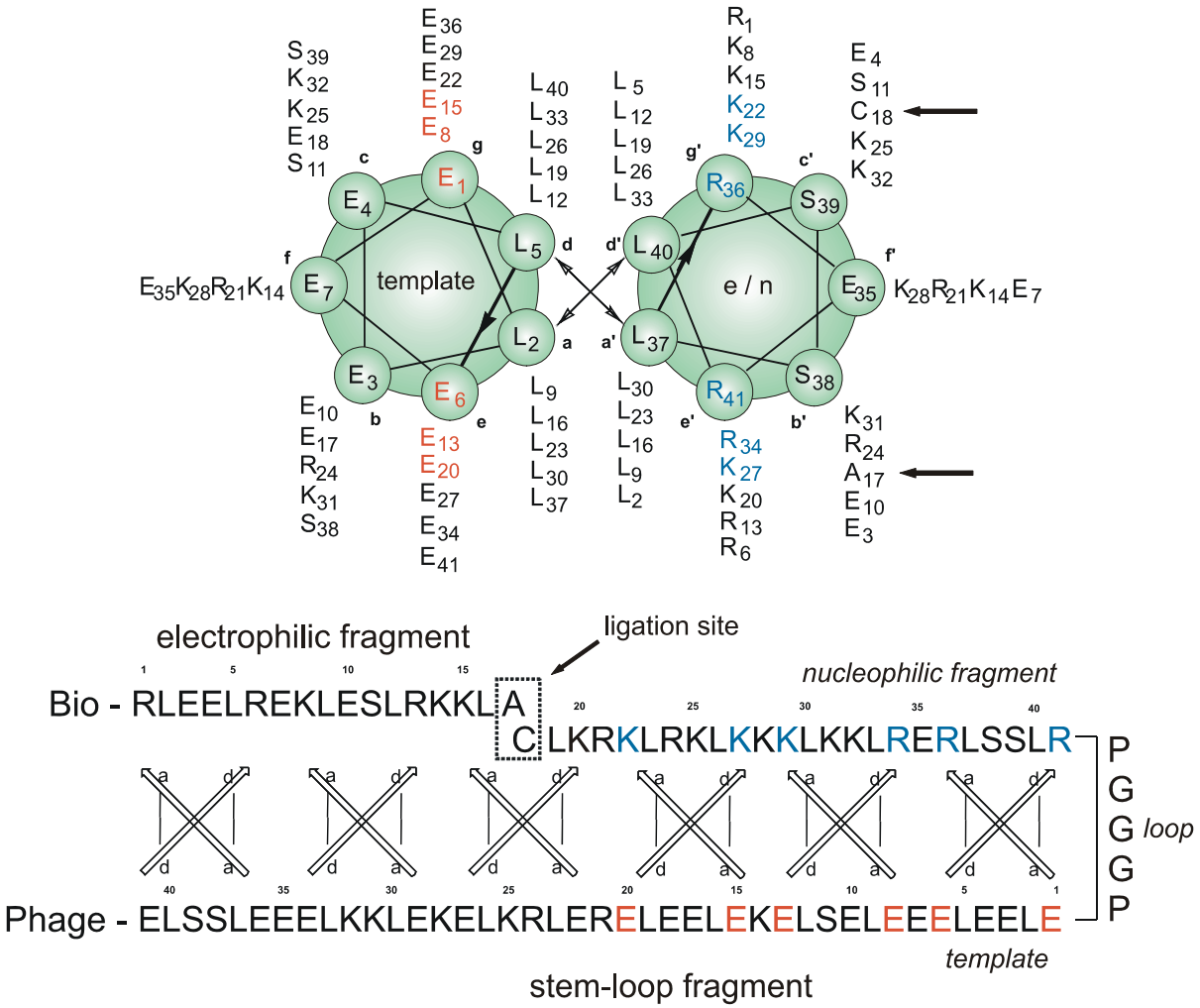


Figure 5.33: *Helical-wheel and sequence representation of the redesigned α-helical coiled coil dimer. Altered amino acid positions within the polar interface are highlighted in blue (nucleophilic fragment) and red (template), respectively.*

Unexpected difficulties occurred during DNA template construction. Comparable to the DNA construction of the former *stem loop* design, overlap extension PCR, using four oligonucleotides as overlapping primers, was applied (Figure 5.24). However, neither standard one-step strategies with all four oligonucleotides in one PCR reaction, nor two-step strategies with separate construction of the 5’ and the 3’ half of the template and subsequent full-length synthesis in a second PCR step resulted in DNA fragments of the right size. Even the optimization of annealing temperature and addition of DMSO to the PCR reactions to prevent mispriming and secondary structure formation could not solve the problem.

A three-step PCR strategy provided the *stem loop* DNA molecules showing the right size. The first step comprised the construction of both the 5' and the 3' half of the template, followed by the filling of small amounts of the products to reach the full-length DNA fragment in a second step. The addition of large amounts of the two outer primers for amplification in a last step comprised the main portion of reaction cycles. However, a high mutation rate was observed after cloning into the phagemid vector and transfection. Finally, the DNA construction was successful in applying a one-step PCR strategy with only trace amounts of both middle primers to minimize cycle rounds and the use of a DNA-polymerase with proofreading activity to lower the mutation rate.

The binding of the electrophilic peptide fragment with the new *stem loop* protein presented on phage surface was evaluated in experiments that were comparable to the binding tests with the original *stem loop* fusion protein. The phagemid clone with stop codons (no) and the biotinylated target peptide of non-*coiled coil* amino acid sequence were used as negative controls. Thus, in each binding test, phage/target combinations (nt/E), (nt/P), (no/E), and (no/P) were tested. *Reversed phage ELISA* experiments with *solution phase binding* (Figure 5.28) as well as *solid phase binding* (Figure 5.29) were performed with variation of several parameters and conditions, such as GdnHCl and Tween concentrations in *coiled coil* binding reactions and washing steps, temperature and the presence of DTT. In the *solution phase* protocol, the concentration of target peptide during binding and the dilution of the binding reaction for capture have been varied.

While higher signals were observed for assays with E compared to P, no differences could be found between both phage types. This means that, again, only unspecific binding of the electrophilic peptide fragment to the phage coat was detected, which could not be prevented or washed away with the detergents used in the assays. Also the detection of the thiol group in the N-terminal cysteine side chain of the *stem loop*, applying *maleimide-phage ELISA* experiments (Figure 5.31b) with different binding cascade strategies and varying reagents for disulfide reduction, failed. The possible explanation that the *stem loop* protein fragment on phage surface is degradated in cell culture before phage preparation could be precluded by *reversed phage ELISA* experiments with capture of a binding reaction that occurred in cell culture during phage production and release. After the *coiled coil* is formed and both fragments are ligated, proteolysis should not be an issue.

Although *stem loop* homodimerization via cystine formation and *coiled coil* folding could not occur, binding and ligation of the *stem loop* and the electrophilic peptide fragment could not be accomplished. It was concluded that either both fragments do not bind tightly enough or the *stem loop* is not presented as pIII-fusion protein on phage surface for any other reason than homodimer formation. In order to investigate both issues separately, the binding of *stem*

loop and E_o had to be evaluated without phage. Therefore, the *coiled coil* fragment was expressed as fusion with maltose binding protein (MBP).

5.3.4.3 Evaluation of *Coiled Coil* Formation and Native Ligation on MBP

Maltose Binding Protein is commonly applied for the evaluation of binding activities of phage-borne peptides in a protein milieu where it can fold properly. Therefore, the peptide is displayed at the N-terminal end of the soluble MBP of *E. coli.* The maltose binding protein is a very good candidate for a protein to be conjugated with the test peptide or protein. The protein is accumulated in the periplasmic space of bacteria, from where it can be purified. It is a ~50kD monomeric protein without any cysteine residues. Therefore, misfolding with cysteine-containing peptide conjugates upon addition and removal of disulfide-reducing reagents is precluded. Furthermore, enzyme-conjugated anti-MBP antibodies are commercially available for the application of the MBP-fusion in *ELISA* binding assays.
Since the MBP-*stem loop* fusion protein was to be expressed by *E. coli*, the *stem loop*-encoding DNA was transferred from the 5'-end of recombinant geneIII in pComb3H to the 5'-end of the MBP-gene in the plasmid vector pC3MBP via identical Sfi I cloning sites. Both of the redesigned *stem loop* DNA templates, (nt) and (no) with two stop codons, were cloned into the MBP vector. The *coiled coil* binding was tested with (nt-MBP) and the electrophilic peptide fragment without amino acid substitutions (E). The proteins (no-MBP), an MBP with conjugated peptide of non-*coiled coil* sequence (ncc-MBP), and pure unconjugated MBP as well as peptide P were used as negative controls. The following MBP/target combinations were tested in binding assays:

(nt-MBP/E) = positive sample; (no-MBP/E), (ncc-MBP/E), (MBP/E), (nt-MBP/P), (no-MBP/P), (ncc-MBP/P), and (MBP/P) = negative controls

After vector transfection and expression of the fusion proteins in XL1-blue cells, the supernatants of cell lysates were directly used in *reversed MBP binding ELISA* experiments without affinity purification of the MBP-fusion proteins (Figure 5.32). The assay started with coating of the plate wells with streptavidin, followed by immobilization of the biotinylated target peptides. The MBP/target binding occurred on solid phase. For the detection and quantification of *coiled coil* formation, an alkaline phosphatase (AP)-conjugated anti-MBP-antibody was bound to the fusion protein, and an enzyme-catalyzed color reaction was measured.
The *ELISA* binding tests of each sample were perfomed with and without disulfide reducing DTT, respectively. A dilution series of MBP solution was used for *coiled coil* binding and

native ligation. No signals have been found for any of the samples with the non-*coiled coil* target peptide. Furthermore, the electrophilic peptide fragment E_o bound neither of the cell lysates without MBP (no-MBP/E), wildtype (MBP/E), and MBP conjugated with a non-*coiled coil* peptide (ncc-MBP/E) in significant amounts. Only samples with the *stem loop*-MBP fusion protein and E_o gave strong *ELISA* signals that decreased along the MBP-dilution series (Figure 5.35). Consequently, *coiled coil* formation occurred between the *stem loop* fragment and the electrophilic peptide fragment. A further indication for successful fragment binding and ligation was the odor of benzylmercaptan, a side product of the native ligation that could be detected in the samples without DTT.

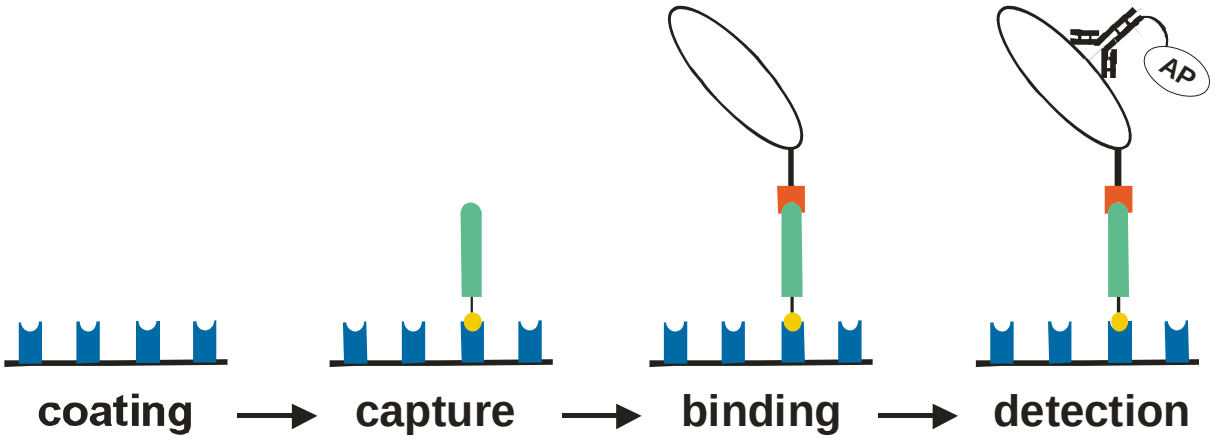


Figure 5.34: *Reversed MBP ELISA with solid phase binding of MBP and target molecule. MBP is represented in white, stem loop protein in red, screening peptide in green, biotin in yellow, streptavidin in blue, and the antibody in black.*

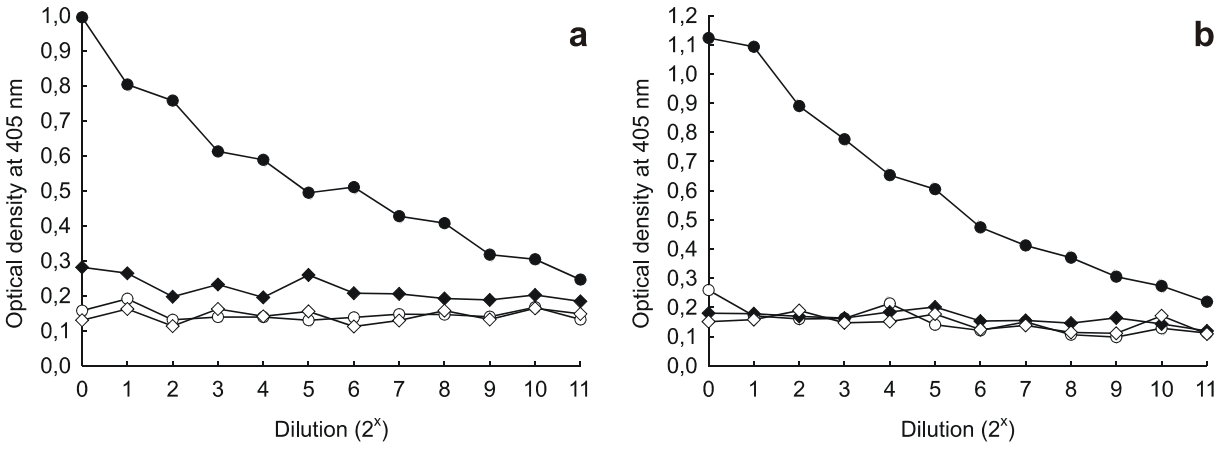


Figure 5.35: *Results of reversed MBP ELISA binding tests of the stem loop (circles) with the electrophilic peptide fragment E_o (closed symbols) without **(a)** and with DTT **(b)** in the coiled coil binding reaction; ncc-MBP (diamonds) and target peptide P (open symbols) were used as negative controls.*

No significant differences in signal intensities could be observed between samples with and without DTT. Since a free thiol group of the N-terminal cysteine residue in *stem loop* is crucial for native ligation, it was concluded that the *stem loop* did not dimerize via disulfide linkage, neither in the periplasm of *E. coli* nor in the cell lysate.

These results proved the design concept of the new *stem loop* protein fragment. The electrophilic peptide fragment can bind via *coiled coil* formation and subsequent native ligation occurs. However, this reaction does not proceed on phage surface. It was concluded that the pIII-*stem loop* fusion protein is not inserted into the phage coat during phage production. As a result, only wild type pIII is present on the surface. To prove this conclusion, the protein composition of the phage particle has been invastigated.

5.3.4.4 Evaluation of the Display of the *Stem Loop* Protein on Phage Coat

To evaluate the protein composition of the phage particles, phage preparations were boiled to destroy the structure of the particles and obtain all phage proteins solubilized (Figure 5.36). A subsequent sodium dodecyl sulfate-polyacrylamide gelelectrophoresis (SDS-PAGE) separated the proteins according to their molecular mass. After transfering the proteins to a membrane (Western Blot), an antibody that specifically binds to the C-terminal region of the minor coat protein pIII of filamentous phage M13 (anti-pIII-mouse antibody / primary antibody) was used to flag the pIII proteins on the membrane. After the specific binding of an anti-mouse-antibody that is HRP-conjugated (secondary antibody) to the primary antibody, an HRP-catalyzed reaction was recorded on an audioradiography film via chemiluminescent detection. This technique should selectively visualize pIII molecules. To obtain acceptable signal-to-noise ratios, both antibodies had to be titered in test experiments.
A phage preparation based on the (nt) phagemid clone was tested. Two bands on the audioradiography film were expected, if the pIII-*stem loop* fusion protein was present on phage surface. A (no)-phage preparation served as negative control. Due to the stop codons in the *stem loop*-encoding DNA sequence upstream of geneIII, only one band, representing the wild type pIII protein, should be seen for this phage preparation.
Furthermore, SDS-PAGE and blotting analysis was performed for pIII molecules derived from (nt) and (no) phagemid clones that had not been obtained from phage preparation but overexpression in *E. coli* without helper phage infection. This IPTG-induced protein expression yields only pIII-*stem loop* fusion protein in case of (nt) and no protein in case of (no), since wild type geneIII is not present in the phagemid vector.

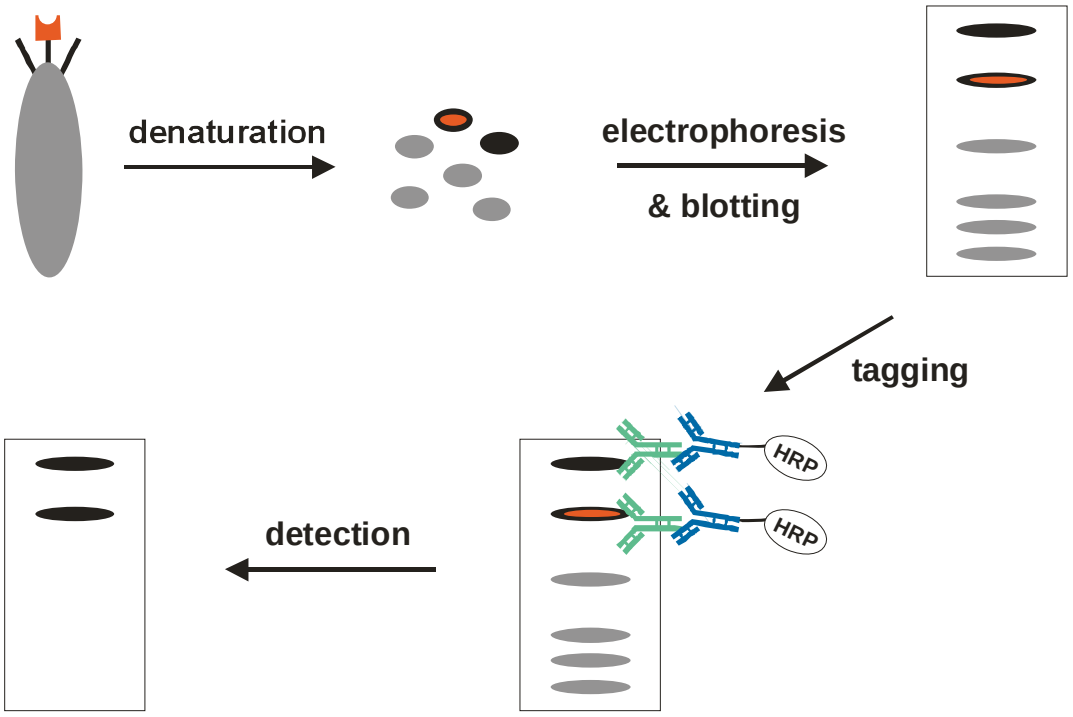


Figure 5.36: *Evaluation of the protein composition of phage particles. Wild type pIII molecules are shown in black, pIII-stem loop fusion proteins in red and black, M13 coat proteins pVI, pVII, pVIII, and pIX in grey, primary antibody in green, and secondary antibody in blue.*

The results of this study proved the suspicion to be true (Figure 5.37). In both the (nt) and the (no) phage preparation samples only one protein, the wild type minor coat protein pIII, was detected. Although this protein has a molecular mass of 42.5kD, the corresponding bands on the SDS-gel appeared at around 70kD. This well-known shift[220,221] may be caused by the glycine-rich linkers between the domains of the protein. The pIII-*stem loop* fusion protein appeared as a double band at >40kD (calculated molecular weight: 24kD) in the (nt)-sample from *E. coli* overexpression but could not be found in the (nt)-phage preparation. The second band may be the result of proteolytic degradation of the *stem loop* fragment within the cell lysate. As expected, no pIII protein was found in the (no)-sample from overexpression in *E. coli*.

Obviously, the pIII-*stem loop* fusion protein is not inserted into the phage coat during assembly of the virion, although this protein can efficiently be produced by *E. coli* and is not toxic to the bacteria cell. A possible explaination was that the *stem loop* fragment interacts with the pIII-portion in the fusion protein and, thus, disturbs correct folding of the C-terminal part of the phage coat protein. Since this domain is crucial for the correct formation of the protein coat of the virion, phage particles with this fusion protein may be unstable. A solution to this problem was the insertion of a linker between both domains in the fusion protein.

Since the *stem loop* fragment could be shown to be fully functional when fused to the N-terminus of maltose binding protein (Section 5.3.4.3), a protein construct was created with MBP between the truncated pIII (amino acids 230-406) and the *stem loop*.

Figure 5.37: *Western Blot of denaturated phage particles and pIII-fusion protein preparations of (nt) and (no) phagemid clones. Lanes 1 and 10: protein mass standard (values given in kD), lane 2: (nt)-phage preparation, lane 3: (no)-phage preparation, lane 4: (nt)-pIII fusion protein from overexpression, lane 5: (no)-pIII-fusion protein from overexpression, lanes 6 and 9: blank, lanes 7and 8: 1:10 dilutions of lane 4 and 5 samples.*

5.3.4.5 Display of the MBP-*Stem Loop* Fusion Protein on Phage Surface

For the construction of a (nt)-phagemid clone that encodes for the pIII-MBP-*stem loop* fusion protein (denoted as “ft” for fusion template), the DNA segment of the pC3MBP vector that encodes for the MBP-*stem loop* conjugated protein had to be amplified via PCR and subsequently ligated to the 5’-end of the pIII-gene in the pComb3H phagemid vector. The Sfi I cloning sites of the phagemid vector were used for the insertion of the MBP-*stem loop* DNA. Both vectors pC3MBP and pComb3H possess equal Sfi I cloning sites for the insertion of DNA that encodes for the peptide or protein to be fused to MBP and pIII, respectively (Figure 5.38a). The palindromic Sfi I recognition sequence at the 5’-end of the MBP-*stem loop* gene was left unchanged. In contrast, the Sfi I recognition site between the MBP and the *stem loop* genes on pC3MBP had to be mutated, otherwise Sfi I digest of the amplification product that had to be performed in order to obtain the matching sticky ends for cloning into pComb3H would have destroyed the DNA construct. Instead, a new Sfi I cloning site was inserted at the

3'-end of the MBP gene that provides matching sticky ends for the ligation of the fusion gene to the 5'-end of the truncated geneIII in pComb3H.

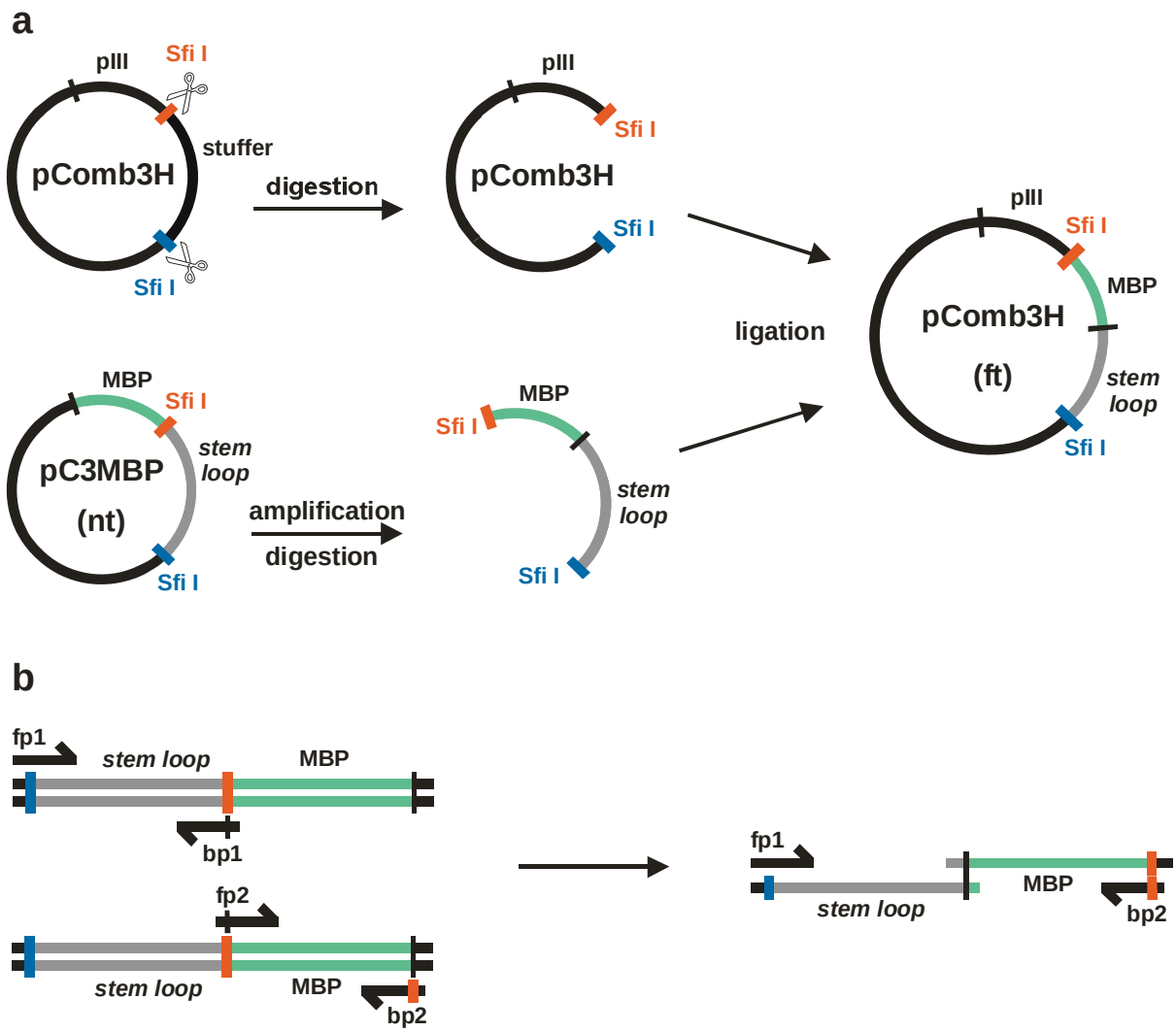


Figure 5.38: *Schematic representation of **a)** the cloning strategy for the construction of the pIII-MBP-stem loop (ft) phagemid clone and **b)** the two-step-PCR strategy for the synthesis of the corresponding DNA-construct; 5'- Sfi I cloning site is represented in blue, the 3' - Sfi I site in red; fp: forward primer; bp: backward primer.*

To accomplish this approach, a two-step-PCR strategy was applied (Figure 5.38b). In a first step, the MBP and the *stem loop* half of the DNA construct were amplified using primers for the mutation of the cloning sites. In a second step, both mutation products have been elongated and amplified to yield the full-length DNA construct that encodes for the *MBP-stem loop* fusion protein.

Since binding tests of the *stem loop* with the electrophilic peptide fragment as MBP-fusion proved that the thioester peptide does not bind to the MBP protein (Section 5.3.4.3), it was

not necessary to create a pIII-MBP-fusion construct that is conjugated to a non-*coiled coil* peptide as negative control in further binding tests. Instead, the (no)-phagemid clone, that produces wild type phage, and the biotinylated target peptide P served as negative controls. The (ft)-phagemid was transfected in *E. coli* for phage preparation. Binding tests were performed with the following phage/target combinations:

(ft/E) = positive sample; (ft/P), (no/E), and (no/P) = negative controls

The assay parameters of the MBP *ELISA* experiments were used since these conditions had been proven to be applicable for the evaluation of *coiled coil* formation and native ligation. The binding was tested with *solution phase* and *solid phase stem loop* binding strategies in *reversed phage ELISA* experiments. Neither of both tests resulted in specific binding of the *coiled coil* fragments. Since native ligation between MBP-*stem loop* and E_o could be demonstrated, it was concluded that the pIII-MBP-*stem loop* protein is, comparable to pIII-*stem loop*, not incorporated into the phage coat during assembly of the virion. This suspicion was confirmed by an *ELISA* experiment for the detection of the MBP protein on phage surface via a specific antibody (Figure 5.39).

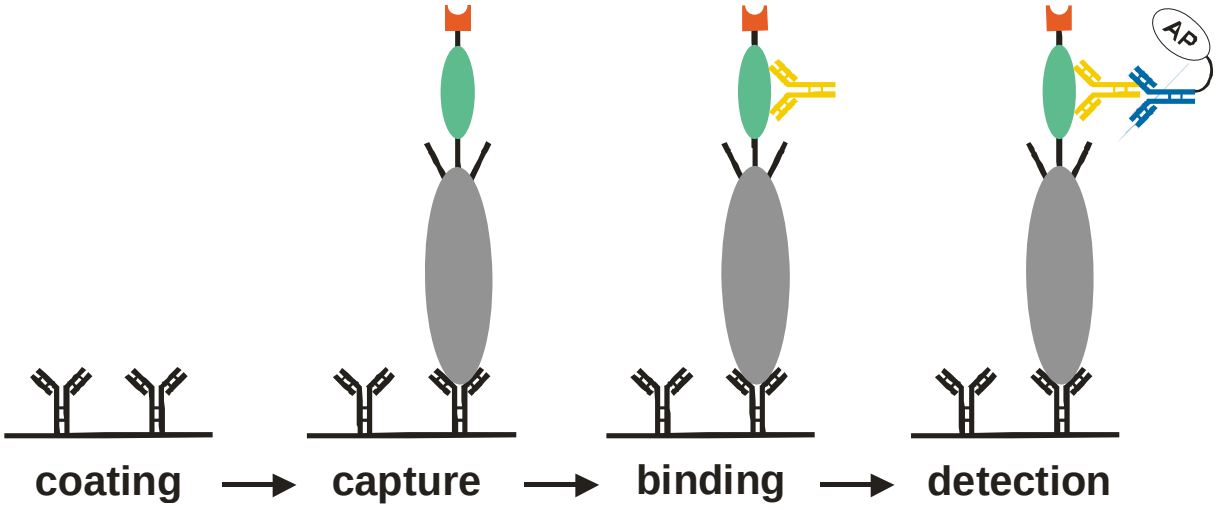


Figure 5.39: *ELISA for the evaluation of the display of the pIII-MBP-stem loop fusion protein on phage surface via specific MBP detection. Phage is shown in grey, MBP in green, stem loop in red, primary antibody in yellow, and secondary antibody in blue.*

In this assay, plate wells were coated with an anti-M13-antibody for the immobilization of phage. An anti-MBP-mouse-antibody was bound to MBP (primary antibody) with subsequent binding of an anti-mouse-antibody (secondary antibody) that was conjugated with alkaline phosphatase (AP). Enzyme-catalyzed color reaction was used for the quantification of

detected MBP protein. The (no)-phage served as negative control. In the binding assay, no differences between the wild type-like (no) and the (ft)-phage could be detected.
Since it could be shown that pIII-homodimerization via *coiled coil* formation and *stem loop*-pIII interactions that disturb the folding of the coat protein are not the problem, another reason had to be found for the issue that pIII molecules, which are fused to the *stem loop* protein fragment, are not integrated into phage coat during assembly in bacteria and release. It was described in literature that the phage display of cysteine-containing peptides as pVIII-fusion led to the incorporation of disulfide-bridged pVIII homodimers into the protein coat of the virion particles.[222] Thereby, pVIII-dimers are formed in the periplasm of the bacteria prior to phage assembly.

Although the phage display of cysteine-containing proteins as pIII-fusion has already been described, even with the cysteine as the N-terminal residue, other cysteine containing pIII-fusion proteins may form disulfide-linked dimers that cannot be inserted into phage. The differences in dimerization behavior between proteins in the periplasmic space may be caused by their folding. If cysteine residues are buried within the protein core, they cannot be oxidized by thioredoxin-containing proteins that are anchored in the cytoplasmic membrane. In contrast, the N-terminal Cys of the *stem loop* is solvent exposed and, therefore, accessible to oxidizing reagents. In a next step, the *coiled coil* formation of the *stem loop* with the electrophilic peptide fragment was tested without the cysteine residue.

5.3.4.6 Investigation of the *Coiled Coil* Formation without Native Ligation

To perform the *stem loop*-E_o binding reaction on phage surface without presentation of a cysteine-containing pIII-fusion protein, the N-terminal Cys residue of the redesigned *stem loop* protein fragment was exchanged by alanine. This amino acid was chosen because it reveals a high propensity for helix formation and introduces a small and unpolar side chain that would not interfere with binding and *coiled coil* formation.
New *stem loop*-encoding DNA constructs that were based on the sequence of (nt) have been designed, produced, and ligated to the 5'-end of the truncated geneIII (amino acids 206-430 of pIII) into the pComb3H vector. One construct comprised the complete *stem loop* sequence, and bacteria that were transfected with the resulting phagemid clone (denoted as "at" for alanine template clone) produced the pIII-fusion protein. As described for the (t) and (nt) clones, a DNA construct was designed with two stop codons in amino acid positions Glu29 and Leu30 (denoted as "ao" for alanine original clone). Since bacteria that were transfected with (ao) produce wild type phage after superinfection with helper phage, the construction of this phagemid was not necessary for binding tests. However, in the case of

successful *coiled coil* formation on phage surface, this clone had to be applied for library construction in order to avoid high phage populations with non-randomized *stem loop* fragment. In contrast to the originally designed *stem loop* with an N-terminal cysteine residue, binding of the electrophilic peptide fragment E_o to the (ao)-phage will not result in the formation of a covalent bond via native ligation (Figure 5.40).

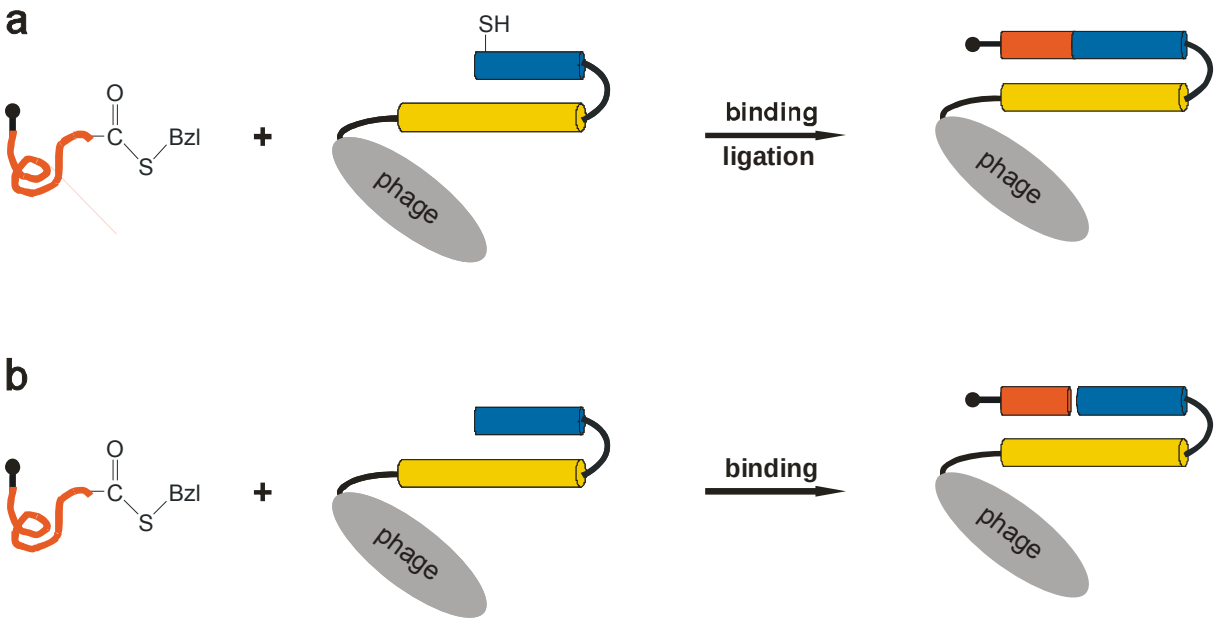


Figure 5.40: *Schematic representation of the binding of the biotinylated electrophilic target peptide to the phage-displayed stem loop protein fragment with* ***(a)*** *and without N-terminal cysteine residue* ***(b)****. Electrophilic peptide fragments are shown in red, the nucleophilic part and the template part of the stem loop in blue and yellow, respectively. The biotin anchor is marked with a closed circle.*

The (ao) and (at) DNA templates were synthesized via one-step overlap extension PCR from four oligonucleotides, comparable to the syntheses of the (nt) and (no) constructs (Section 5.3.4.2). Although a polymerase with proofreading activity was used for the PCR experiments, a high error rate was observed for the constructions of both the (at) and the (ao) DNA templates. No acceptable clone could be found for (ao). Since this phagemid clone does not necessarily have to be used in the binding tests that had to be performed before library construction, this issue was skipped at this point. In case of (at), no clone was found with the sequence entirely errorless. However, a clone could be isolated that had the codon of Lys31 mutated to arginine. As this mutation results in a substitution of lysine by arginine in a *coiled coil* **b** position, no structure destabilizing impact or interference with *coiled coil* formation was expected as consequence. Since the newly inserted guanidino group can form stronger salt bridges with Glu35 than the amino group of lysine, this alteration should

stabilize the structure of the *stem loop* protein fragment. This effect was expected to be advantageous for the binding of the electrophilic peptide fragment, since it causes a higher propensity for helix formation in the C-terminal portion of the template part that binds to the thioester peptide. This structure-stabilizing impact may compensate the *coiled coil* destabilization that is caused by the lack of native ligation. Consequently, the isolated phagemid clone, containing the described mutation, was chosen as (at)-phage.
For technical reasons, *ELISA* experiments could not be performed in order to test the binding of the electrophilic peptide fragment to the alanine-substituted *stem loop*. Instead, the bound phage were removed from solid phase by proteolysis and quantified by reinfection of bacteria and titration of input and output. Therefore, the biotinylated target peptide was immobilized on streptavidin-coated magnetic particles and phage/target binding reaction occurred on solid phase. Phage that were produced from (no)-transfected bacteria (wild type phage) were used as negative control. The output titers for both the (at) and the (no) phage were very low and did not differ significantly. Since *stem loop*-E_o binding without native ligation is much weaker than that of the full-length dimer, washing steps after binding reaction were reduced from 10 to 1 in a further test. However, no increase in output/input ratio could be achieved for neither of both samples.

It was taken into account that the C-terminal benzylthioester of the electrophilic peptide fragment may strongly destabilize the *coiled coil* structure due to steric hindrance and, thus, weakens the binding of both peptide fragments. Consequently, binding tests were repeated with a variant of the non-substituted electrophilic peptide fragment (E_o) that possessed a native peptide-like C-terminal carboxyl group instead of the benzylthioester (denoted as E_C). The hydrolysis of the ester group failed. While the moiety was stable under weak basic conditions (pH 8-9), the peptide was destroyed when stronger bases were used. Instead, E_C was newly synthesized. Binding tests were perfomed with different numbers of washing steps. The results revealed that both the output values and the output/input ratios of both samples decreased with increasing number of washing steps (Figure 5.41).
Thereby, no significant differences between the (at)-phage clone, that is supposed to display the alanine-containing *stem loop* on surface, and the wild type phage could be observed. One possible explanation for the failure was again that the pIII-fusion protein is not inserted into phage coat. Since the display of *coiled coil*-based *stem loop* proteins without cysteine residues on minor coat protein pIII had already been described in literature,[7] this reason was relatively unlikely. Alternatively, the E_C-peptide was too short to form a stable *coiled coil* with the *stem loop* fragment without ligation to the full-length dimer. Obviously, native ligation is crucial for stable *coiled coil* formation on phage surface and, thus, the N-terminal cysteine residue of the *stem loop* is needed.

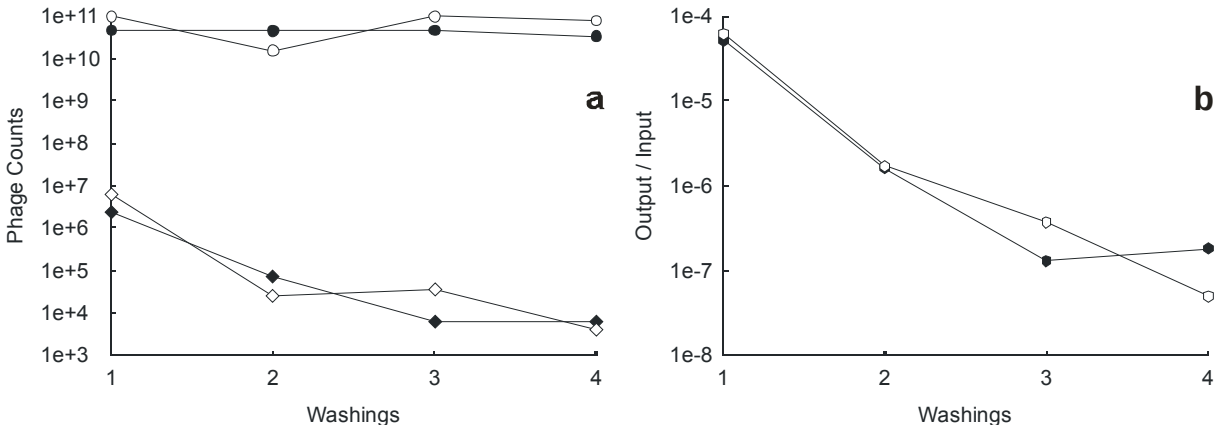


Figure 5.41: *Phage titers **(a)** and output/input ratios **(b)** of binding tests of E_C with the alanine-containing stem loop fragment (closed symbols) and the wild type phage (open symbols). Phage input values are shown by circles, output values by diamonds.*

On the one hand, display of cysteine-containing phage coat-fusion proteins can lead to dimerization via disulfide linkage, which may inhibit incorporation into the particle. On the other hand, affinity selection from phage-displayed cysteine-containing randomized proteins, e.g., antibody libraries,' and even native ligation on N-terminal cysteine residues, can be successfully accomplished. One strategy to get the cysteine-containing *stem loop* fragment displayed on phage surface was, therefore, changing the relative position of the Cys residue in the protein. Since the structure of the *stem loop* could not significantly be altered in order to permit *coiled coil* formation and native ligation, the protein had to be expressed and inserted into phage coat possessing an N-terminal tag. This additional peptide fragment had to be removed prior to binding reaction.

5.3.4.7 Display of the *Stem Loop* on Phage Surface with an N-Terminal Tag

A peptide tag, the so-called FLAG tag, was chosen, because it offers two main features: a recognition sequence of the enteropeptidase enterokinase for cleavage from the *stem loop* protein and an epitope for a commercially available antibody (Figure 5.42).

Enterokinase (EK) is the physiological activator of pancreatic trypsinogen in mammals.[223] While the transformation of trypsinogen to trypsin by proteolysis requires the entire protein,[224] hydrolytic cleavage of the recognition sequence-containing small peptidyl substrates from fusion proteins can be successfully performed with the single catalytic subunit of the enzyme (EK_L).[225,226] Although the often described high specificity for the tetraaspartyllysyl recognition sequence has made EK a widely used tool for the cleavage of expressed fusion proteins and has led to optimized production[227] and immobilization

protocols,[228] unwanted cleavage of fusion protein that was attributed to a broad specificity of enterokinase has also been described.[229] Minimum sequence requirements that have been concluded from structural data comprise a basic amino acid in P_1 and two acidic residues in P2 and P3 positions, respectively.[230] All canonical amino acids have been tested in EK substrates in P'_1 position, and it was found that this protease is relatively permissive of the amino acid residue downstream of the recognition sequence.[231]

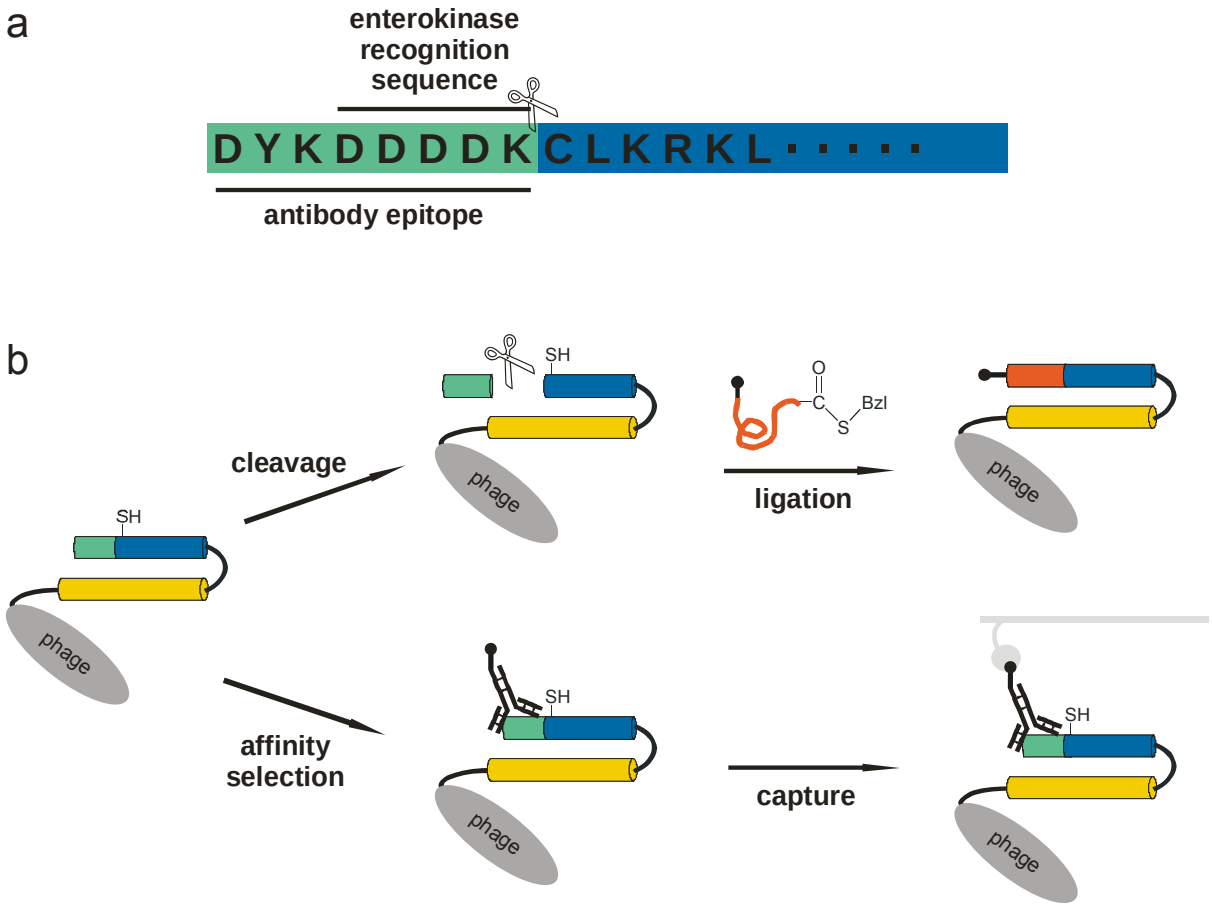


Figure 5.42: *Sequence of the FLAG tag **(a)** and schematic representation of the FLAG tag-stem loop fusion protein displayed on phage surface **(b)**. FLAG tag is shown in green, template fragment in yellow, nucleophilic fragment in red, labeled antibody in black.*

The FLAG tag-*stem loop* fusion protein was based on the sequence of the *stem loop* construct (nt) that has the polar interaction domain redesigned to prevent homodimerization via *coiled coil* formation (Section 5.3.4.2). The DNA encoding for the FLAG tag *stem loop* was produced via amplification of the (nt) gene on pComb3H, applying a forward primer that introduces the FLAG tag-encoding DNA segment as overhang (Figure 5.43). The Sfi I cloning site at the 5'-end of the stem loop had to be shifted to the 5'-end of the FLAG tag

DNA. Therefor, the original site was mutated and a new one was created at the end of the fusion gene.

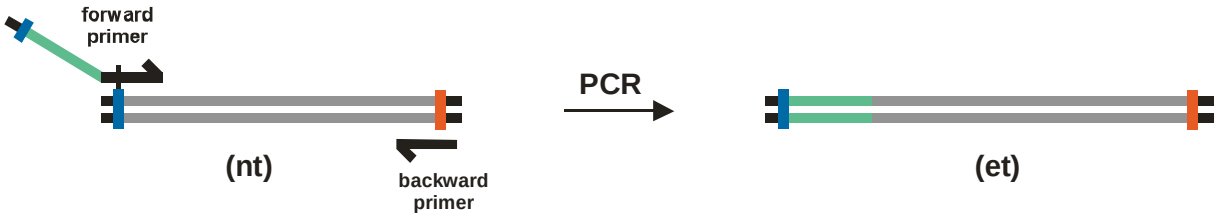


Figure 5.43: *Schematic representation of the construction of the FLAG tag-stem loop DNA fragment. Stem loop gene is shown in grey, FLAG tag gene in green, and Sfi I cloning sites in blue and red, respectively.*

The new DNA construct was cloned into pComb3H and the phagemid clone (denoted as “et” for enterokinase template) transfected into *E. coli* for phage preparation. Before evaluating EK_L-cleavage of the tag and subsequent binding of the free *stem loop* to the electrophilic peptide fragment, the presence of the FLAG tag-*stem loop* fusion was tested using an anti-FLAG-antibody (Figure 5.42b). In this binding test, the biotin-conjugated antibody was immobilized on the surface of streptavidin-coated magnetic particles. After capture, bound phage were removed from solid phase via proteolysis and quantified by reinfection and titration of input and output. Two phage served as negative control: the (no) clone that produced wild type phage and the (at) clone that produced phage with the Ala-substituted *stem loop* fusion protein on the surface.

The results revealed a significantly higher output/input ratio for the phage derived from (et) phagemid vector, compared to both negative controls (Figure 5.44). Consequently, the presence of the FLAG-stem loop-pIII fusion protein on phage surface could be proven by this experiment. However, the ratio is only about 5.3 times higher for (et) than for (at), and yet this value is about 20.6 times higher than for (no). Either this finding is an artefact or the applied antibody possessed little binding affinity to the *stem loop* protein fragment. The output values for all of the three tested phage clones were lower than expected, since the immobilized antibody should have a maximum phage-binding capacity of about 10^{12} particles per sample. One reason might have been the multiple biotin conjugation of the antibody molecules which resulted in visible clustering of the magnetic streptavidin-coated particles. This crosslinking obviously lowered the phage-binding capacity.

After *coiled coil* formation and native ligation of the *stem loop* with the electrophilic peptide fragment as well as the display of the FLAG-*stem loop* fusion protein on phage surface were proven, the tag had to be cleaved in order to perform the binding of both *coiled coil* fragments on phage surface.

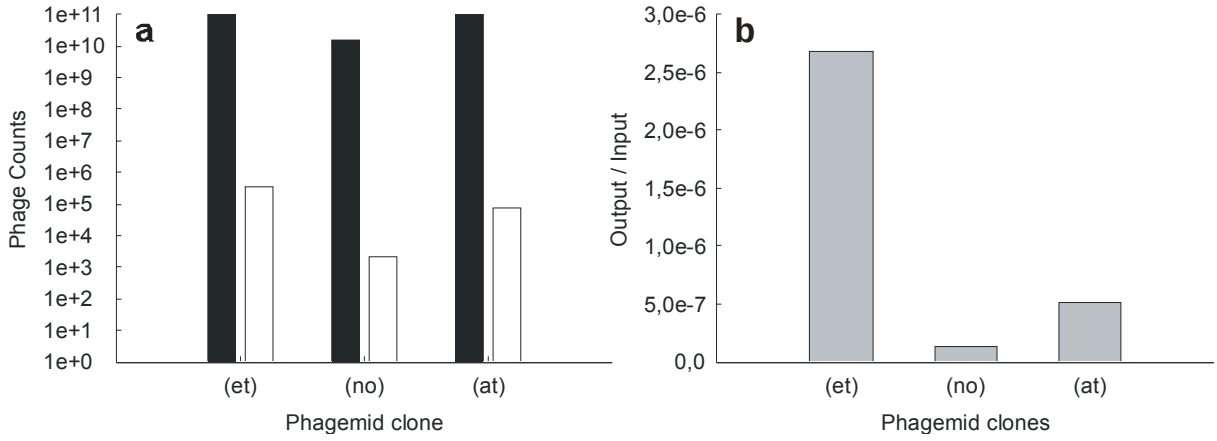


Figure 5.44: *Phage titers **(a)** and output/input ratios **(b)** of binding tests of an anti-FLAG-antibody to the FLAG-stem loop fusion on phage surface. Input values are represented by black bars, output values by white bars, and output/input ratios by grey bars.*

The major drawback at this point was the lack of knowledge about the enterokinase reaction conditions that had to be provided for the enzyme. Several factors had to be tested, such as reaction time, amount of enzyme, reaction temperature, and the presence of disulfide-reducing reagents. Commercial suppliers recommended reaction times of 16 hours and more. However, since enterokinase reaction should take place on the surface of virion particles and phage are precipitated from bacteria cultures without further purification, the reaction time should be as short as possible in order to avoid degradation of the phage-displayed *stem loop* protein fragment by bacterial proteases.

A further issue was the amount of enzyme needed for sufficient cleavage of the tag. This factor strongly depends on the application, which means the specific fusion protein. Values given from the supplier vary over several orders of magnitude from 0.0001% to 0.5% (w/w) of the protein to be cleaved. The use of disulfide reducing reagents was quite questionable since pIII-stem loop-FLAG proteins may be presented as disulfide-linked dimers on phage surface but the impact of such reagents on catalytic activity of the enterokinase was not known.

All these factors had to be systematically varied and optimized in order to accomplish quantitative removal of the FLAG tag. It was most important that the background of phage binding during library selection procedures is as low as possible, since it could be shown that amplification of unspecifically bound particles leads to enrichment of phage clones with vector deletions (Chapter 5.3.3). For technical reasons, *ELISA* tests could not be performed and captured phage in binding assays had to be quantified by bacteria infection and titration. Since a maximum of four samples per experiment could be tested, it was not possible to evaluate all of the critical factors of the enterokinase reaction in a combinatorial manner.

Reaction conditions had to be varied systematically instead in separate binding experiments. Samples without enterokinase served as negative controls.

The first experiments for the optimization of enterokinase reaction conditions were performed with 1µL of the commercially obatined enzyme solution (12.5U). Referring to the amount of enzyme needed for the cleavage of a test substrate (MBP-fusion protein), according to the supplier, 1µL of enzyme solution should convert 10^{13}-10^{14} phage-displayed fusion proteins. Since this value depends on the substrate, as mentioned above, the amount used was just a starting point. Each test reaction was perfomed with and without the incubation of phage with 25mM DTT, respectively, prior to preparation from cell culture. Cleavage reaction was performed at different temperatures (RT or 37°C) and in varying reaction times (1h – 20h). *Coiled coil* binding and native ligation proceeded with the immobilized, non-substituted electrophilic peptide fragment directly after enterokinase reaction.

A general observation was a higher phage output for samples that were treated with DTT, compared to that without the disulfide-reducing reagent. Since this effect was also found for negative controls, one possible explanation was that phage monomers are more efficient in infecting bacteria than disulfide-linked dimers and, thus, reach higher output values. Enterokinase reactions resulted in slightly higher phage output values, compared to the negative controls without addition of the enzyme. However, the phage counts of the EK_L-samples were mostly less than 1.5 times that of the negative controls. The cleavage reaction obviously proceeded but had to be further optimized to significantly lower the background in the subsequent binding reaction.

In the following cleavage and binding experiments the amount of enterokinase was varied (1.25 U – 125 U). Very surprisingly, the phage output of the binding reaction increased along with decreasing enzyme concentrations. Obviously, enterokinase unspecifically proteolyzed the *stem loop* protein fragment on phage surface. Conditions for the enterokinase reaction had to be found that minimized unspecific proteolysis. Therefore, the amount of EK_L, reaction time (1h / 20h), and the concentration of DTT for reduction of *stem loop* dimers prior to phage preparation were systematically varied in further binding tests.

The results revealed that the phage output increased, compared to the negative controls, along with a decreasing amount of enzyme but with an increasing reaction time and concentration of disulfide-reducing reagent (Figure 5.45). Longer reactions with less enzyme raised the specificity of enterokinase but the lowest background measured was still higher than 50%. The differences in output ratios between the samples of 10mM and 25mM DTT were somewhat surprising, however, since both concentrations should be high enough to quantitatively reduce phage dimers. It was concluded that the FLAG-*stem loop* fragment reoxidizes after phage preparation. This effect was not observed for the *stem loop*-MBP-fusion protein fragment without the tag (Section 5.3.4.3).

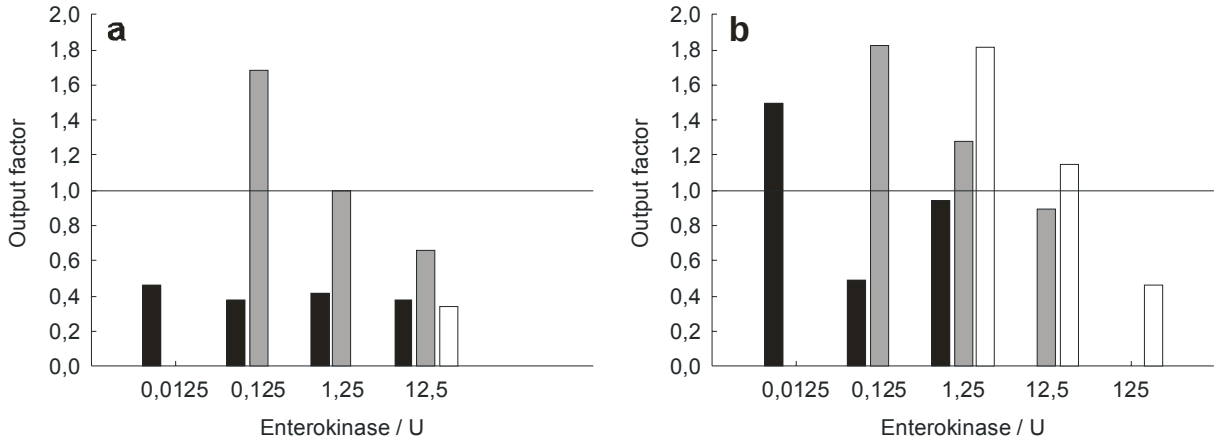


Figure 5.45: *Binding tests of the FLAG-stem loop fusion protein with the electrophilic peptide fragment (E_o) on phage surface. Phage were incubated with different amounts of enterokinase (EK_L) prior to binding reaction for 1h* ***(a)*** *or 20h* ***(b)****. Output factor is defined as the ratio of phage output of the sample with EK_L reaction and the corresponding negative control (same reaction conditions, but without EK_L reaction). Phage were incubated prior to preparation without DTT (black bars), with 10mM DTT (grey bars), and 25mM DTT (white bars), respectively.*

Since DTT from reduction prior to preparation is not completely removed by phage precipitation, the rate of reoxidation depends on the initial DTT concentration in phage culture. More test assays have been performed with phage that were resuspended in DTT-solution of different concentrations (5mM – 50mM) after precipitation. Thus, FLAG-*stem loop* protein fragments could not dimerize via disulfide formation during enterokinase reaction. The first tests showed that EK_L reactions with additional DTT could not be performed at temperatures higher than RT, over night, and with DTT concentrations higher than 10mM, respectively. Under these conditions, no differences in phage output between enterokinase samples and the corresponding negative controls were observed. Consequently, further tests were performed for 1h at room temperature in 10mM DTT.

For most samples tested, the output was about three times higher than the corresponding value of the negative control. Interestingly, this effect was completety independent from the amount of enzyme that was used (0.000125U – 1.25U) and, thus, attributed to an artefact. The disulfide-reducing reagent, which was necessary to prevent reoxidation of *stem loop* cysteine residues, obviously interfered with the structure or the catalytic mechanism of enterokinase. Test assays with the removal of the tag and native ligation of the electrophilic peptide fragment to the cleaved *stem loop* in a one-pot reaction failed too.

Unfortunately, successful cleavage of the FLAG tag and subsequent *coiled coil* formation followed by native ligation on phage surface could not be accomplished. It was not possible

to find conditions that are needed for the *stem loop* fragment to be present in a monomeric form and permit efficient enzymatic catalysis at the same time. Furthermore, unspecific degradation of the fusion protein on phage coat occurred, although the declared minimal sequence requirements for enterokinase recognition (two acidic residues in positions P_3 and P_2 as well as a basic residue in position P_1) are not present in the *stem loop* sequence. However, this might be the case for the truncated pIII portion of the fusion protein.

The newly developed *coiled coil* based screening system, which could be successfully applied for the investigation of the interaction properties of fluoroalkyl-substituted amino acids in native polypeptide environments (Section 5.2), should be used for the selection of preferred interaction partners of these non-natural building blocks via library screenings on phage surface. A *coiled coil* protein fragment was designed for display on phage surface and covalent binding to the target screening peptide could be shown on the surface of a soluble protein. Furthermore, a peptide library of high diversity was constructed. The construction procedure had been optimized and the obtained library size was sufficient for the screening of two separate randomized peptide pools: one for fluoroalkyl-substitutions in the polar interaction domain and another for alterations in the hydrophobic core. However, the display of the *coiled coil* protein fragment with free N-terminal cysteine residue for native ligation on the surface of filamentous phage M13 could not be accomplished. Since binding of the short target peptide to the phage displayed *coiled coil* fragment without formation of a covalent bond was too weak, a new strategy would comprise the display of one single *coiled coil* strand containing the amino acid positions to be randomized on phage surface. In this case, the screening had to be performed with the complementary helix strand via *coiled coil* formation. The main idea of this strategy for the selection from *coiled coil*-based libraries has meanwhile been described in literature.[232]