

Studying the Interaction Profiles of Nonnatural Amino Acids – Towards Predicting their Specific Applications at α -Helical Interfaces



Dissertation zur Erlangung des akademischen Grades des
Doktors der Naturwissenschaften (Dr. rer. nat.)

eingereicht im Fachbereich Biologie, Chemie, Pharmazie
der Freien Universität Berlin

vorgelegt von

Elisabeth K. Nyakatura, MSc

aus Berlin, Deutschland

August 2013

1. Gutachterin: Prof. Dr. Beate Koksch (Freie Universität Berlin)

2. Gutachter: Prof. Dr. Nediljko Budisa (Technische Universität Berlin)

Disputation am: 12. November 2013

Erklärung

Die vorliegende Arbeit wurde auf Anregung und unter Anleitung von Frau Prof. Dr. Beate Kokschi in der Zeit von Oktober 2008 bis Juli 2013 an dem Institut für Chemie und Biochemie des Fachbereichs Biologie, Chemie, Pharmazie der Freien Universität Berlin angefertigt.

Hiermit versichere ich, dass ich die vorliegende Arbeit mit dem Titel "*Studying the Interaction Profiles of Nonnatural Amino Acids –Towards Predicting their Specific Applications at α -Helical Interfaces*", ohne Benutzung anderer als der zugelassenen Hilfsmittel selbstständig angefertigt habe. Alle angeführten Zitate sind als solche kenntlich gemacht. Die vorliegende Arbeit wurde in keinem früheren Promotionsverfahren angenommen oder als ungenügend beurteilt.

Berlin im August 2013

Elisabeth K. Nyakatura

Die Dissertation wurde in englischer Sprache verfasst.

PUBLIKATIONSLISTE*

- (7) **E.K. Nyakatura**,[‡] R.R. Araghi,[‡] J. Mortier, S. Wieczorek, C. Baldauf, G. Wolber, B. Kokschi, *An unusual interstrand H-bond stabilizes the hetero-assembly of helical $\alpha\beta$ -chimeras with natural peptides*, ACS Chem. Biol., 2013, in revision.
([‡]equal contributors)
- (6) H. Erdbrink,[‡] **E.K. Nyakatura**,[‡] S. Huhmann, U.I.M. Gerling, D. Lenz, B. Kokschi, C. Czekelius, *Synthesis of enantiomerically pure (2S,3S)-5,5,5-trifluoroisoleucine and (2R,3S)-5,5,5-trifluoro-allo-isoleucine*, Beilstein J. Org. Chem., 2013, 9, 2009-2014.
([‡]equal contributors)
- (5) **E.K. Nyakatura**, O. Reimann, T. Vagt, M. Salwiczek, B. Kokschi, *Accommodating fluorinated amino acids in a helical peptide environment*, RSC Adv., 2013, 3(18), 6319-6322.
- (4) M. Salwiczek, **E.K. Nyakatura**, U.I.M. Gerling, S. Ye and B. Kokschi, *Fluorinated amino acids: compatibility with native protein structures and effects on protein–protein interactions*, Chem. Soc. Rev., 2012, 41, 2135-2171.
- (3) T. Vagt, **E. Nyakatura**, M. Salwiczek, C. Jäckel, B. Kokschi, *Towards identifying preferred interaction partners of fluorinated amino acids within the hydrophobic environment of a dimeric coiled coil peptide*, Org. Biomol. Chem., 2010, 8, 1382-1386.
- (2) L. Vozzolo, B. Loh, P.J. Gane, M. Tribak, L. Zhou, I. Anderson, **E. Nyakatura**, R. G. Jenner, D. Selwood, A. Fassati, *Gyrase B inhibitor impairs HIV-1 replication by targeting Hsp90 and the capsid protein*, J. Biol. Chem., 2010, 285(50), 39314-28.
- (1) M. Salwiczek, S. Samsonov, T. Vagt, **E. Nyakatura**, E. Fleige, J. Numata, H. Cölfen, M.T. Pisabarro, B. Kokschi, *Position-dependent effects of fluorinated amino acids on the hydrophobic core formation of a heterodimeric coiled coil*, Chem. Eur. J., 2009, 15, 31, 7628-7636.

* Aus dieser Dissertation gingen bisher die Veröffentlichungen 3-7 hervor.

DANKSAGUNG

Frau Prof. Dr. Beate Koksche danke ich für die interessante und abwechslungsreiche Themenstellung dieser Arbeit. Insbesondere möchte ich Ihnen für die mir stets gewährte wissenschaftliche Freiheit und das Vertrauen danken, die es mir ermöglichten eigene Ideen selbstständig zu entwickeln und umzusetzen. Auch möchte ich mich bei Ihnen für die Ermöglichung zahlreicher Reisen zu nationalen und internationalen Konferenzen bedanken.

Weiterhin möchte ich mich bei Prof. Dr. Nediljko Budisa (Technische Universität Berlin) für die Übernahme des Zweitgutachtens meiner Dissertation bedanken.

Für die rasche und dennoch sehr sorgfältige Korrekturlesung des Manuskriptes dieser Arbeit, sowie für ihren stets sehr motivierenden Zuspruch möchte ich mich bei Dr. Allison Berger bedanken.

Herrn Prof. Dr. Wolfram Saenger und Prof. Dr. Markus Wahl danke ich für die Bereitstellung bester Arbeitsbedingungen für molekularbiologische Untersuchungen in Ihren Laboratorien. Den Mitgliedern Ihrer Arbeitsgruppen danke ich für die sehr angenehme Arbeitsatmosphäre und stete Hilfsbereitschaft.

Für die sehr umfangreiche fachliche Unterstützung, sowie für die vielen wertvollen Diskussionen bezüglich des *phage displays* möchte ich mich bei Dr. Toni Vagt bedanken. Dr. Jérémie Mortier möchte ich für die Durchführung der Molekulardynamik-Simulationen und die sehr fruchtbare Zusammenarbeit danken. Dr. Raheleh Rezaei Araghi danke ich ebenfalls für die Zusammenarbeit, sowie für die Bereitstellung der Daten, welche sie im Rahmen unseres gemeinsamen Projektes generierte. Auch danke ich Oliver Reimann, Sebastian Wieczorek und Susanne Huhmann, die während der Bearbeitung Ihrer Masterarbeiten intensiv an diesem Forschungsprojekt mitgearbeitet haben. Meiner Freundin Jessica Gazdiç möchte ich für die graphische Unterstützung danken. Mein spezieller Dank gilt meinem Freund und Mentor Dr. Mario Salwiczek. Ich bedanke mich für die vielen intensiven Diskussionen und Anregungen, die schon früh mein besonderes Interesse für das *coiled-coil* Faltungsmotiv weckten.

Den Mitgliedern der AG Koksche danke ich für das Messen unzähliger Masseproben, die ständige Bereitstellung von frischem Spülole, die Instandhaltung unseres Geräteparks, und die stets sehr freundliche Arbeitsatmosphäre. Herzlich bedanke ich mich bei Ulla Gerling, Dr. Jessica Falenski und Dr. Aileen Justies für ihre kameradschaftliche Hilfsbereitschaft und die vielen anregenden Gespräche.

Abschließend gebührt mein Dank meinem Mann, meinen Freunden und meiner Familie, insbesondere meinen Eltern. Ich danke für die grenzenlose Unterstützung, für den absoluten Rückhalt, das entgegengebrachte Verständnis und für das bedingungslose Vertrauen in mich und meine Fähigkeiten. Ohne den konstanten Zuspruch sowie die aufbauenden Worte meiner Lieben wäre die Realisierung dieser Promotion nicht möglich gewesen.

KURZZUSAMMENFASSUNG

Nichtnatürliche Aminosäuren werden häufig in peptidbasierte Pharmazeutika eingebaut, um deren Bioverfügbarkeit und Spezifität zu erhöhen. Die systematische Untersuchung der Interaktionsprofile solcher Aminosäuren mit ihren natürlichen Analoga in nativen Proteinumgebungen liefert wertvolle Erkenntnisse, die Vorhersagen über die Eigenschaften neuartiger nichtnatürlicher peptidbasierter Therapeutika und Biomaterialien ermöglichen. In dieser Arbeit sollten mithilfe von *phage display* Wechselwirkungspartner identifiziert werden, die bevorzugt mit fluoralkylsubstituierten Aminosäuren oder mit Motiven alternierender β - und γ -Aminosäuren, im Kontakt helikaler Strukturen interagieren.

Die Untersuchung von Peptiden mit unterschiedlichen fluorierten Analoga von (S)-2-Aminobuttersäure zeigte, dass der Einbau solcher Aminosäuren in den hydrophoben Kern eines parallelen, heterodimeren *coiled coil*-Faltungsmotivs ähnliche Wechselwirkungspartner hervorbringt. Trotz deutlicher Unterschiede in Größe und Hydrophobie der fluorierten Seitenketten, bevorzugen alle hier untersuchten Aminobuttersäureanaloga die aliphatischen Aminosäuren Leucin oder Isoleucin als ihre natürlichen Wechselwirkungspartner. Die Selektion führte zu einer optimierten Packung um die fluorierten Seitenketten, was sich in einer thermischen Stabilitätserhöhung äußerte. Diese Studie verifiziert, dass *coiled coil*-Strukturen bereitwillig verschiedene fluoralkylsubstituierte Aminosäuren in ihrem hydrophoben Kern als nichtnatürliche Bausteine akzeptieren.

Im Gegensatz dazu wurden mittels *phage display* diverse *heptad repeat*-Muster selektiert, die Erkennungsspezifitäten für $\beta\gamma$ -Foldamere an der Grenzfläche von parallelen und antiparallelen α -helikalen Anordnungen aufweisen. Mit den selektierten Sequenzen konnte gezeigt werden, dass durch die Einführung von ausgewählten, polaren Wasserstoffbrückendonoren die Stabilität von Interaktionen zwischen rückgraterweiternden β - und γ -Aminosäuren und α -peptidischen Motiven erhöht wird. Diese Donoren sind in der Lage Wasserstoffbrücken mit freien Rückgratcarbonylgruppen von $\alpha\beta\gamma$ -Chimären auszubilden. Zusätzlich können auch Mutationen, die die Oberfläche des hydrophoben Kerns vergrößern zu einer thermischen Stabilisierung beitragen.

Schließlich wurden einleitende Studien zur rationalen Entwicklung von proteasestabilen, peptidbasierten Inhibitoren der gp41 Hüllproteinuntereinheit von HIV durchgeführt. Es wird angenommen, dass gp41 während der Infektion eine Schlüsselrolle einnimmt, indem es ein Bündel aus sechs Helices ausbildet, das sich aus drei N-terminalen Heptad-Repeat (NHR) Segmenten und drei C-terminalen Heptad-Repeat (CHR) Segmenten zusammensetzt. Mehrere C-Peptid Analoga unterschiedlicher Länge wurden intrahelikal verbrückt, um ihre α -Helizität zu erhöhen. Die Auswirkungen dieser Modifikation auf die Bindungsaffinität zu NHR Segmenten wurden untersucht. Außerdem wurde eine CHR Sequenz generiert, die bindungsaffin zu einem NHR Segment ist und somit möglicherweise die Ausbildung des hexameren Bündels inhibieren kann. Das hochkonservierte Trp-Trp-Ile Motiv, das für die Bündelausbildung ausschlaggebend ist, befindet sich in der Mitte dieser Sequenz und wird in künftigen Untersuchungen als Ausgangspunkt für Substitutionsstudien mit fluorierten Aminosäuren fungieren.

ABSTRACT

Nonnatural amino acids are frequently incorporated into peptide-based pharmaceuticals to improve bioavailability and specificity. Systematic studies of how such residues interact with natural amino acids in the context of native protein folds may yield important information that can facilitate the prediction of the properties of novel nonnatural peptide therapeutics or biomaterials. To this end, the current study made use of phage display to screen large numbers of helical microenvironments to identify, based on binding affinity, natural peptide sequences that preferably interact with sequences containing either fluoroalkyl substituted amino acids or alternating sets of β - and γ -amino acids.

The use of peptides that contain different analogues of (S)-2-aminobutyric acid with different numbers of fluorine atoms, and thus different side-chain volumes, showed that the incorporation of these amino acids into the hydrophobic core of a parallel heterodimeric coiled coil leads to similar pairing characteristics. Despite their differences in hydrophobicity and size, all investigated amino acids prefer to interact with the aliphatic amino acids leucine and isoleucine. However, the selection led to an optimized side chain packing which is expressed in thermal stability enhancements. It was verified that coiled coils readily accommodate diverse fluorinated aliphatic amino acids as nonnatural building blocks within their hydrophobic cores.

In contrast, diverse α -amino acid patterns were found with recognition specificity for $\beta\gamma$ -foldameric sequences, at the interface of parallel or antiparallel helical assemblies. Here, the sequences selected by phage display show that the incorporation of a polar H-bond donor functionality can significantly improve helical interactions involving backbone extended amino acids, because these donors are able to engage a free backbone carbonyl of $\alpha\beta\gamma$ -chimeras in an interstrand H-bond. A mutation leading to an increase of the surface area of the hydrophobic core had a similar effect on the thermal stability.

Finally, initial studies were carried out towards the rational design of a protease resistant peptide-based inhibitor of HIV's gp41 envelope protein subunit. Gp41 is thought to play a key role in infection by facilitating host cell entry via the assembly of a bundle of six α -helices composed of three N-terminal heptad repeat (NHR) segments and three C-terminal heptad repeat (CHR) segments. Several short C-peptide analogues of differing length were intrahelically crosslinked to increase α -helicity, and tested for binding affinity to NHR segments. A CHR derived peptide sequence was generated that shows affinity for a NHR segment, and may therefore inhibit bundle formation. The highly conserved Trp-Trp-Ile motif, which is crucial for tight helix alignment, constitutes the center of this CHR peptide and will serve as a starting point for systematic substitution studies with fluorinated aliphatic amino acid analogues.

ABBREVIATIONS

| | |
|-----------------------|----------------------------------------------------------|
| ACN | Acetonitrile |
| Boc | <i>tert</i> -Butyloxycarbonyl |
| bp | Base pairs |
| Bio | Biotin |
| BSA | Bovine serum albumin |
| Carb | Carbenicillin |
| CD | Circular dichroism |
| Cfu | Colony forming units |
| CHR | Carboxy-terminal heptad repeat |
| DBU | 1,8-Diazabicyclo[5.4.0]undec-7-ene |
| DCM | Dichloromethane |
| DIC | Diisopropylcarbodiimide |
| DIPEA | Diisopropylethylamin |
| DMF | Dimethylformamide |
| DNA | Deoxyribonucleic acid |
| DTT | Dithiothreitol |
| dNTP | Deoxyribonucleotide |
| dsDNA | Double stranded deoxyribonucleic acid |
| <i>E. coli</i> | <i>Escherichia coli</i> |
| EDT | 1,2-Ethanedithiol |
| ESI-ToF-MS | Electrospray ionization Time-of-Flight Mass Spectrometry |
| Fmoc | 9-Flourenylmethoxycarbonyl |
| FRET | Fluorescence resonance energy transfer |
| gp41 | glycoprotein 41 |
| GndHCl | Guanidine hydrochloride |
| HBS | Hydrogen bond suggurate |
| HIV | Human immunodeficiency virus |
| HOAt | 1-Hydroxy-7-aza-benzotriazole |
| HOBt | 1-Hydroxybenzotriazole |
| HPLC | High Performance Liquid Chromatography |
| LB | Lysogeny broth |
| MD | Molecular dynamics |
| MOPS | 3-(N-morpholino)propanesulfonic acid |
| mRNA | Messenger Ribonucleic acid |
| Mtt | N-Methyltrityl |
| NMP | 1-Methyl-2-pyrrolidone |

| | |
|----------------------|----------------------------------------------------------------------|
| NHR | Amino-terminal heptad repeat |
| OD | Optical density |
| PBD | Pocket binding domain |
| PBS | Phosphate-buffered saline |
| PCR | Polymerase chain reaction |
| PEG | Polyethylene glycol |
| PIP | Piperidine |
| PS | Packaging signal |
| PTFE | Polytetrafluoroethylene |
| rpm | Rounds per minute |
| RT | Room temperature |
| SB | Super broth |
| SEC | Size exclusion chromatography |
| SD | Supplementary data |
| SLS | Static Light Scattering |
| SOC | Super Optimal broth with Catabolite repression |
| SPPS | Solid Phase Peptide Synthesis |
| SPR | Surface plasmon resonance |
| ssDNA | Single stranded deoxyribonucleic acid |
| <i>t</i>-Bu | <i>tert</i> -Butyl |
| TAE | Tris-acetate-EDTA |
| TBS | Tris-buffered saline |
| TBTU | O-(Benzotriazol-1-yl)-N,N,N',N'-tetramethyluronium tetrafluoroborate |
| TCEP | <i>tris</i> -(2-carboxyethyl)phosphine |
| TFA | Trifluoroacetic Acid |
| TIS | Triisopropylsilane |
| T_M | Melting temperature |
| Tris | Tris-(hydroxymethyl)aminomethane |
| Trt | Trityl / Triphenylmethyl |
| UV/VIS | Ultraviolet and visible spectrum |
| vdW | van der Waals |

The three- and one-letter codes for amino acids are consistent with the biochemical nomenclature proposed by the IUPAC-IUB commission (*Eur. J. Biochem.* **1984**, *138*, 9-37). Abbreviations of noncoded amino acids relevant to the present dissertation are given below. The abbreviations correspond to the L-amino acids.

| | |
|------------------------------------------------------|----------------------------------------------------------------------------------------------------|
| 3³-F₃Ile | 3',3',3'-Trifluoroisoleucine; (S)-2-Amino-3,3,3-(trifluoromethyl)pentanoic acid |
| 5³-F₃Ile | 5,5,5-Trifluoroisoleucine; (S)-2-amino-5,5,5-trifluoro-3-(methyl)pentanoic acid |
| 5³,5'³-F₆Leu | 5,5,5,5',5',5'-Hexafluoroleucine; (S)-2-amino-5,5,5-trifluoro-4-(trifluoromethyl)pentanoic acid |
| 4³-F₃Val | 4,4,4-Trifluorovaline; (S)-2-amino-3-(trifluoromethyl)butanoic acid |
| 2,3,4,5-F₄Phe | 2,3,4,5-Tetrafluorophenylalanine; (S)-2-amino-3-(2,3,4,5-tetrafluoro phenyl)propanoic acid |
| 2,3,5,6-F₄Phe | 2,3,5,6-Tetrafluorophenylalanine; (S)-2-amino-3-(2,3,5,6-tetrafluorophenyl)propanoic acid |
| 2,3,4,5,6-F₅Phe | Perfluorophenylalanine; (S)-2-amino-3-(perfluorophenyl)propanoic acid |
| Abu | Aminobutyric acid; (S)-2-aminobutyric acid |
| Abz | 2-Aminobenzoic acid |
| Aha | Aminoheptanoic acid; (S)-2-aminoheptanoic acid |
| DfeGly | Difluoroethylglycine; (S)-2-amino-4,4-difluorobutanoic acid |
| DfpGly | Difluoropropylglycine; (S)-2-amino-4,4-difluoropentanoic acid |
| MfeGly | Monofluoroethylglycine; (S)-2-amino-4-fluorobutanoic acid |
| Nle | Norleucine; (S)-2-aminohexanoic acid |
| TfeGly | Trifluoroethylglycine; (S)-2-amino-4,4,4-trifluorobutanoic acid |
| TfmAla | Trifluoromethylalanine, 2-amino-2-(trifluoromethyl)propanoic acid |
| Y(NO₂) | 3-Nitrotyrosine |

CONTENTS

| | | |
|-----------|---------------------------------------------------------------------------------------------------|------------|
| 1 | Introduction..... | 1 |
| 2 | Theoretical Background and Scientific Context..... | 3 |
| 2.1 | The α -Helical Coiled-Coil Folding Motif..... | 3 |
| 2.2 | Modifying Helical Structures with Nonnatural Amino Acids..... | 12 |
| 2.3 | Phage Display as an Example for Protein Evolution..... | 29 |
| 3 | Aim..... | 35 |
| 4 | Concept and Previous Studies..... | 37 |
| 4.1 | Applied Screening System..... | 37 |
| 4.2 | Previous Phage-Display Experiments with VPE/VPK..... | 40 |
| 5 | Results and Conclusions..... | 43 |
| 5.1 | Accommodating Fluorinated Amino Acids in Parallel Coiled-Coil Dimers..... | 43 |
| 5.2 | Stabilizing a Coiled Coil that Contains a Set of Alternating β -and γ -Amino Acids.. | 53 |
| 5.3 | Towards Protease Stable Fluorinated HIV Entry Inhibitors..... | 63 |
| 6 | Summary and Outlook..... | 75 |
| 7 | Experimental Procedures and Analytical Methods..... | 79 |
| 7.1 | Peptide Synthesis and Characterization..... | 79 |
| 7.2 | Structural Analysis..... | 89 |
| 7.3 | Phage Display..... | 95 |
| 9 | Supplementary Data..... | 107 |
| 9.1 | Phage Display with Fluorinated Amino Acids..... | 107 |
| 9.2 | Phage Display with $\beta\gamma$ -Foldamers..... | 114 |
| 10 | Literature..... | 123 |

1 Introduction

The fidelity of nearly all life functions relies on specific interactions involving proteins. To fulfill their respective tasks in nature, proteins have evolved by means of selective pressure.¹ The particular functionality of peptides and proteins is created by the physical and chemical properties of their building blocks, i.e. amino acids, which are precisely arranged in a three-dimensional structure through folding.² Folded protein units can serve as modules building up large assemblies such as virus particles or muscle fibers. They also provide specific binding sites, as found in enzymes or proteins that carry oxygen and they regulate the function of DNA. The omnipresence of peptides and proteins in biological systems includes a wide variety of processes responsible for biological disorders and diseases.^{3,4} These can be either caused by defective proteins or mediated by protein-protein interactions during pathogen infection.

Studying the relationship between the amino acid sequence of a protein and its biologically active three dimensional structure, not to mention function, remains one of the greatest challenges in molecular biology.¹ α -Helices constitute the most abundant structural motif, and are thus found at the interface of several important intracellular protein-protein interactions mediating a number of critical signaling events and pathogenic processes.⁵ For example, the entry of viruses into host cells is frequently accomplished via helical bundle formation of surface exposed proteins. Since protein-protein interactions generally involve large interfacial surface areas, the α -helical folding motif has been of extensive interest as a minimal unit in the development of interfering peptides. Precisely folded peptides can be utilized to engage in numerous energetically favorable contacts over a relatively large surface, a property that can enhance binding affinity and selectivity.^{6,7} Therefore, as a drug class, peptide based inhibitors offer numerous advantages over small molecules.

Unfortunately, the clinical utility of peptide based inhibitors is limited by their low bioavailability, which results from their rapid degradation by proteases, slow passive transport in blood, and generally low membrane penetration.⁸ Moreover, when isolated from the rest of the protein, small α -helical segments of a protein usually show little or no helical content.⁶ One approach to improving the pharmacokinetic properties of peptides is to introduce nonnatural building blocks with distinct functionalities. Recent advances in solid phase peptide synthesis, chemoselective peptide ligation,⁹ protein semi-synthesis,¹⁰ and nonnatural protein expression via aminoacylated suppressor-tRNA or via *sense* codon reassignment using auxotrophic host strains¹¹⁻¹³ have enabled the site-specific or global incorporation of nonnatural amino acids into peptides and proteins.¹⁴

Artificial amino acid surrogates, in particular those containing heteroatoms which are not found in the natural amino acids pool, can be used to introduce new functional groups, that may change the nature of the protein-protein interactions in which the protein of interest

engages. In this respect, the substitution of one or more hydrogen atoms by fluorine seems opportune, since, for a number of reasons, the fluorination of organic molecules has been proven to be beneficial in medicinal chemistry.¹⁵ While C-F substitutions for C-H are considered to be “shape conservative”, the presence of a single fluorine atom can significantly alter the biophysical and chemical properties of amino acids, such as hydrophobicity, acidity/basicity, reactivity, and conformation.¹⁴ Although several studies have demonstrated that fluorocarbon containing amino acids can substantially increase peptide and protein stability,¹⁶⁻²² recent studies of the Kokschi group have shown that the impact of size and polarity of the fluorinated building blocks highly depends on the immediate environment of the substitution, a fact that makes the effects of such substitutions difficult to predict.²³⁻²⁵

Another approach that addresses the disadvantages that accompany the use of peptide inhibitors is the design of unnatural oligomers that closely resemble both, protein folding and function (commonly referred to as foldamers).²⁶⁻³⁰ Foldamers which are composed of amino acids that possess one or two additional backbone methyl groups, denoted β - and γ -amino acids, respectively, have been shown to mimic the self-assembly of native peptides, and thus present functional motifs in an ordered manner. Since proteases usually do not recognize β - and γ -amino acids, proteolytic stability can be achieved by substituting them for native amino acids. However, despite the growing number of such systems, the precise mimicry of side chain topology and recognition properties of natural α -helical sequences remains challenging.

Systematic studies of the interaction profiles of fluoroalkyl substituted amino acids and foldameric sequences in helical protein environments are needed to provide a knowledge base, and perhaps even predictors of behavior, for the rational design of modified peptide based pharmaceuticals targeting helical protein interfaces.

2 Theoretical Background and Scientific Context

2.1 The α -Helical Coiled-Coil Folding Motif

Due to unfavorable solvent interactions of hydrophobic amino acid residues, isolated α -helices are only marginally stable in aqueous medium.¹ Therefore, in proteins, they are usually stabilized by being packed together through hydrophobic side chains. As early as 1952 Crick showed that if two assembling α -helices are not straight rods but rather wound around each other in a superhelix, a so called coiled-coil arrangement, side-chain interactions are maximized.³¹ In general, coiled coils are composed of two to six α -helices that are wrapped around each other in a left handed manner. Recently, also a stable heptamer could be generated.³²

An estimate of 5-10% of all proteins possess coiled-coil domains,³³ whose biological function seems to relate to their length and architecture.³⁴⁻³⁶ Coiled-coil domains that stretch out over several hundred amino acids generally undergo structural and mechanical functions; they build up the cytoskeleton (intermediate filament proteins) or assemble muscle fibers (myosin).³⁷ Many short coiled coils have been shown to serve as structural components or recognition domains of DNA binding proteins. Prominent representatives of the latter are the transcription factors Jun, Fos, and GCN4.

Their diverse functions and their involvement in pathogenic events, for example viral fusion, make them attractive candidates for drug development. Moreover, since their basic structural features are well understood, their folding is predictable, and the interaction partners within the assembly are clearly defined, coiled coils are commonly applied as model systems to study unnatural amino acids in natural polypeptide environments.

2.1.1 Structure

As mentioned above, the coiled coil folding motif was first hypothesized in 1952 by Crick in the course of the investigation of the X-ray diffraction pattern of α -keratin.³¹ Furthermore, Pauling and Corey postulated that the primary structure of α -keratin must follow a repetitive amino acid sequence that comprises seven residues, of which every third and fourth amino acid is hydrophobic.³⁸ It took 20 years to confirm this repetition pattern (denoted heptad repeat $(abcdefg)_n$) through sequencing tropomyosin,³⁹ and another 20 years to solve the first high resolution structure of an isolated coiled coil.⁴⁰

Each helical turn in coiled coils consists of 3.5 amino acids. Thus, one heptad comprises two turns. As proposed by Pauling and Corey, positions *a* and *d* typically harbor nonpolar amino acids (Leu, Ile, Val, Ala, Met),⁴¹ which form the first recognition domain through hydrophobic core packing ('knobs-into-holes') - the key driving force for folding and assembly.⁴² Charged amino acids at positions *e* and *g* form the second recognition motif by

engaging in interhelical electrostatic interactions. The solvent exposed positions *b*, *c*, and *f* are usually occupied by polar residues.

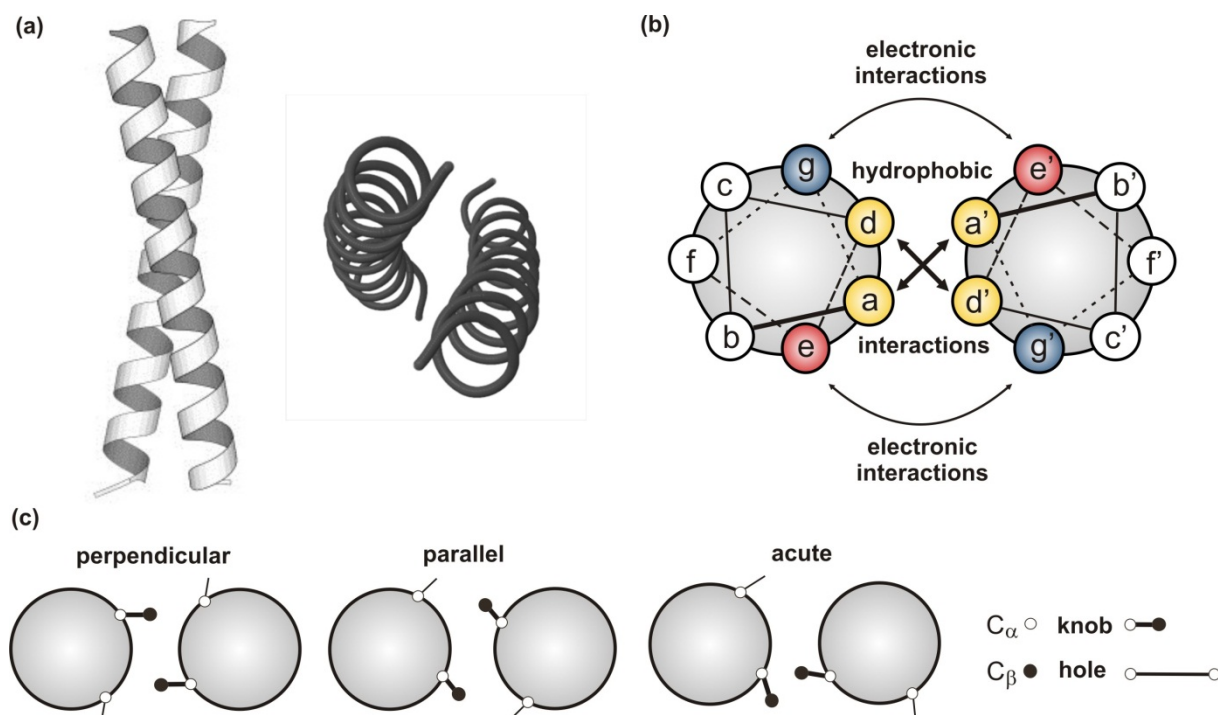


Figure 2.1: (a) backbone secondary structure of a dimeric coiled coil: frontal view (left) and along the superhelical axis (right). (b) Helical Wheel Representation that shows where each residue side chain is approximately located in the top view (heptad repeat pattern) looking down the α -helix of two helices aligned in parallel. (c) different core-packing geometries (based on P. Harbury et al.)⁴³

Efficient packing of *a*- and *d*-positions within the hydrophobic core is the main stability providing factor. The packing geometries of residues at *a*- and *d*-positions can differ fundamentally (Figure 2.1).^{43,44} Antiparallel helix orientation leads to the packing of *a* against *d'* and *d* against *a'*, resulting in two equally packed layers within the hydrophobic core.⁴⁵ In a parallel alignment of the helices, however, the amino acid composition within the core can substantially influence the oligomerization state.⁴³ In parallel dimers residues in *a*-positions are packed side-by-side with residues in *a'*-positions of the opposite strand. In this type of packing, termed *parallel packing*, the α - β bond vector points out from the helical interface. As a consequence, β -branched residues (Ile or Val) are favored at *a*-positions because they project their branched γ -methyl group back into the helical interface.⁴² Residues in *d*-positions pack against amino acids in *d'*-positions and the C_{α} - C_{β} bond vectors point towards each other. This type of geometry, called *perpendicular packing*, precludes β -branched residues from occupying these sites and Leu is favored.⁴³ From the latter originates the name of the famous '*leucine zipper*'; the hallmark of DNA transcription factors, such as GCN4.⁴² In parallel coiled coil tetramers this geometry is reversed. The packing at *a* is perpendicular, while the packing at *d* is parallel. In the case of parallel trimers, the packing at *a* and *d* closely resembles each other (intermediate between perpendicular and parallel) and is

therefore referred to as *acute packing* (Figure 2.1c). Both sites tolerate branched and unbranched hydrophobic residues to the same degree.

Moreover, it was found that placement of polar Asn residues in the hydrophobic environment favors dimer formation,⁴⁶ while hydrogen bonding between two asparagine side chains may also direct helix orientation.^{47,48} Folding specificity can be further dictated by interhelical coulomb interactions between *e*- and *g*-positions.⁴⁹ Depending on whether these interactions are attractive or repulsive, a specific orientation is stabilized.⁵⁰⁻⁵² Charged residues in core flanking positions may aid coiled-coil stability by contributing to hydrophobic interactions in their nonionic form.⁵¹

2.1.2 HIV's gp41 – a Coiled Coil Facilitates Viral Entry

To infect cells, viruses need to overcome the cell membrane barrier. Enveloped viruses achieve this by membrane fusion, generally initiated by so called fusion proteins that are expressed as precursor proteins.^{53,54} Upon receptor binding at the cell membrane these proteins undergo dramatic conformational transitions which result in viral and cellular membrane merger. Knowledge about the molecular and biophysical events of the process is a prerequisite for a thorough understanding of this essential step in the virus life cycle, as well as for the rational design of intervention methods. The so called 'class I' fusion mechanism used by the human deficiency virus (HIV) which involves transmembrane coiled-coil domains is common to a variety of enveloped viruses including those responsible for influenza, Ebola, and SARS.^{54,55}

Cell Entry

HIV's envelope glycoprotein (gp160) is proteolytically cleaved into two noncovalently associated subunits, gp120 and gp41.^{56,57} CD4 cell receptor binding of gp120 induces a cascade of conformational changes in the envelope protein which lead to cytokine coreceptor (CCR5 or CXCR4) binding and exposure of the gp41 fusion peptide (Figure 2.2).^{55,56,58,59}

Current models suggest that gp41's amino terminal region (denoted fusion peptide) interacts with the target cell membrane resulting in an intermediate pre-hairpin state of gp41 that bridges the viral and the host cell membrane.⁶⁰ This pre-hairpin possesses a relatively long half-life and thus constitutes the target for several inhibitory compounds.^{57,61,62} The prehairpin then refolds into the six-helix bundle core structure, and it is this transition that catalyzes membrane fusion.⁶³ Experimental evidence suggests that fusion proceeds via a gp41 membrane stalk which opens a fusion pore that is subsequently expanded.^{61,63,64} The energy released during gp41 refolding is used to overcome the kinetic barrier,⁵⁵ which is emphasized by the high thermostability of gp41's core structure.^{65,66}

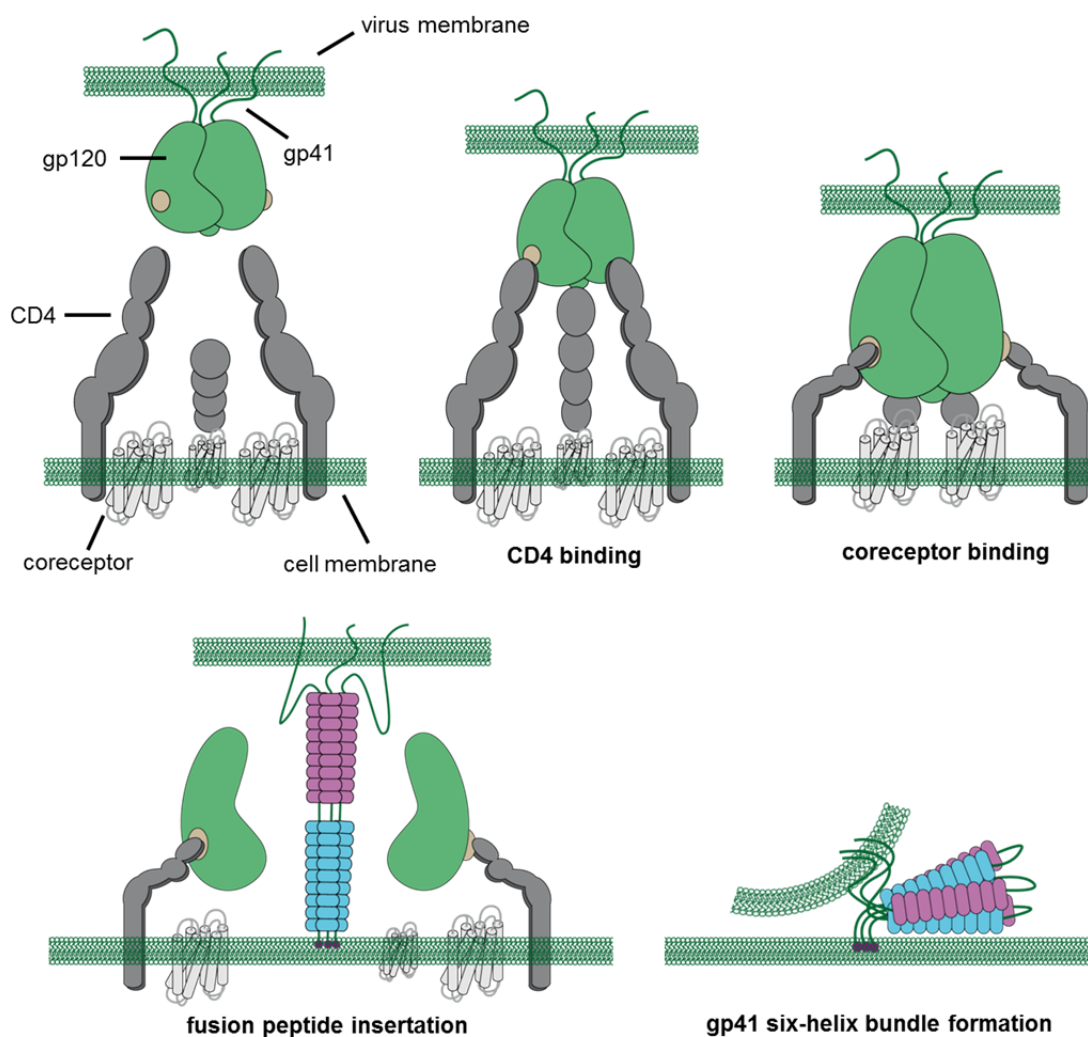


Figure 2.2: Mechanism of HIV entry. On CD4 binding (binding site for CD4 is shown in ocher), gp120 undergoes conformational changes. CD4-induced epitopes can then bind to chemokine receptors. Thereafter, gp41 is released into a fusogenic conformation and its N-terminal (blue) and C-terminal (purple) helices form a hairpin structure, leading to the approximation of viral and cellular membranes, which results in membrane fusion. Adapted from Esté et al.⁶⁷

Core Structure of HIV-1 Gp41

The gp41 molecule is a transmembrane protein with two distinct features within its ectodomain. First, the amino terminus, which possesses the hydrophobic, glycine-rich fusion peptide, that is essential for host cell membrane attachment. Second, two heptad repeat regions that are separated through a loop region, which contains two disulfide bridged cysteine residues. The two heptad repeat stretches, termed N (amino terminal) and C (carboxy-terminal) peptides, or NHR and CHR respectively, have been crystallized and studied extensively.^{56,60} It was shown that the N- and C-terminal peptides fold into the six-helix bundle which represents the fusogenic state of gp41 (Figure 2.2).⁵⁶ The center of this bundle consists of a parallel, trimeric coiled coil of three N-peptide helices exhibiting acute knobs-into-holes packing geometry similar to that of a trimeric GCN4 isomer (Figure 2.3).^{44,56} Three C-peptides wrap antiparallel to the N-helices in a left-handed direction around the outside of the central trimer.⁵⁶ The overall dimensions of this complex comprise a cylinder

measuring ~ 55 Å in height and ~ 35 Å in diameter. The interior residues at the *a*- and *d*-positions of the NHR are predominantly hydrophobic, while sporadic buried polar interactions are present. A sequence comparison of HIV-1 and SIV gp41 showed that the residues at *a*- and *d*-positions are highly conserved (Figure 2.3b). An electrostatic potential map of the N-peptide coiled-coil trimer revealed that its surface is largely uncharged. The grooves that enable C-peptide interaction are lined with predominantly hydrophobic residues (Leu, Ile, Val, Trp). As a result isolated N-peptides tend to aggregate in polar solvents.⁶⁵ In contrast, the complex of N- and C-peptides, where C-peptides are packed against the outside of the N-peptide trimer in an antiparallel fashion, exhibits a highly charged surface due to solvent exposed acidic residues of C-peptide helices.⁵⁶ As a consequence the complex shows greater solubility than the homomeric internal trimer.

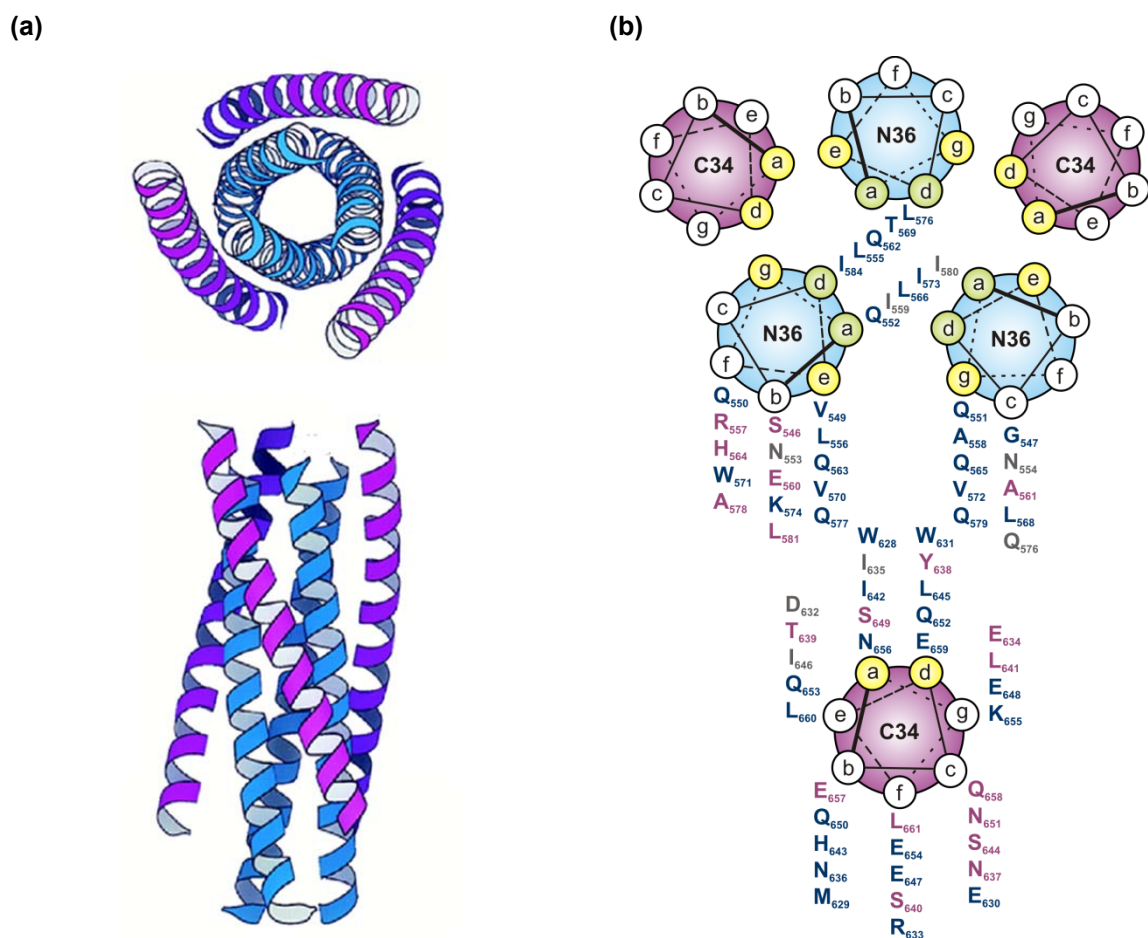


Figure 2.3 (a) A model of the HIV type-1 (HIV-1) gp41 CHR region interaction with the NHR region to form a hairpin structure (six-stranded helix bundle). The NHR helices are depicted in blue; the CHR helices are depicted in purple. (b) Helical Wheel Representation of a 36 amino acid peptide derived from the NHR (N36) and a 34 amino acid peptide derived from the CHR (C34). The residues at each position are colored according to their conservation between HIV-1 (HXB2) and SIV (Mac239): blue, identity; black, conservative change; purple, nonconservative change. The NHR helices interact through “knobs-into-holes” packing interactions at the *a* –and *d*-positions (colored green). Positions of the NHR and CHR helices that occupy the interhelical space between two NHR helices and a CHR helix are shown in yellow. (Adapted from Chan et al.)⁵⁶

The interaction between the C- and N-peptides occurs mainly via hydrophobic residues in positions *e* and *g* of three grooves on the surface of the central N-peptide trimer, denoted

hydrophobic pocket, and positions *a* and *d* of the buttressing C-peptide. Sequence comparisons between HIV and SIV show that residues lining these grooves as well as residues involved in binding to them are highly conserved. Each of the grooves has a particularly deep cavity (~16 Å long, ~7 Å wide, and ~5-6 Å deep) that accommodates three hydrophobic residues from the interacting C-helix: Ile₆₃₅, Trp₆₃₁, and Trp₆₂₈. N-peptide residues that flank CHR helices (positions *b*, *c*, and *f*) are predominantly divergent between HIV and SIV.

Since the hydrophobic pocket is highly conserved among strains, it is considered to be an ideal target for the development of HIV fusion inhibitors; however, the intermolecular binding interface of the gp41 fusion complex comprises a large surface area. Due to their precise folding, α -helical peptide inhibitors can be utilized to present functional groups in a multivalent manner over a relatively large surface. As a consequence, peptide based HIV-1 fusion inhibitors generally exhibit higher binding affinities and selectivities when compared to their small molecule analogues.^{68,69} Towards this end, several chemically modified CHR derived peptide sequences have been generated in the course of the current study.

2.1.3 Coiled Coils as Model Systems to Study Nonnatural Amino Acids

With respect to the rational design of proteins that adopt a desired fold, which can be time-consuming and laborious, coiled coils have emerged as suitable model systems for the systematic study of nonnatural amino acids within native protein environments.^{14,42,43} Based on the sequence-structure relationship outlined above, they fulfill the following key requirements of model systems: (i) a predictable fold that defines interacting residues; (ii) structural and thermodynamic stability towards substitutions; and (iii) sensitivity to detect even subtle modifications. Moreover, automated solid phase peptide synthesis allows for quick and efficient generation of coiled-coil peptides. Several coiled coil based model systems have been generated to study the impact of fluorination on helical protein-protein interactions,^{16,18-20,22,23,70-72} as well as to investigate the structural consequences of the incorporation of backbone extended amino acids into helical assemblies.⁷³⁻⁷⁶ In the following, the two model systems employed in this study to identify the preferred interaction profiles of these two types of nonnatural amino acids are explained in detail.

VPE/VPK

The VPE/VPK model system (Figure 2.4) was *de novo* designed and extensively characterized in the Kocsch group. It has been applied to study the impact of amino acid side chain fluorination in a native helical protein environment (*vide infra*; see chapter 2.2.1).²⁵ Its design is based on the amino acid composition of the transcription factor GCN4 as well as on studies of Hodges *et al.*^{40,77}

To guarantee that observed effects can be traced back to a single fluoroamino acid substitution, specific heterodimerization of the model is of crucial importance. This condition was accomplished by applying two key design strategies:²⁵

- (i) Introduction of $e-g'$ and $g-e'$ pairs that engage in favorable electrostatic interactions in heteromer formation, but would repel one another in either homomeric assembly.⁷⁸
- (ii) Positioning valine in all a -positions, since β -branched amino acids in α -positions are known to stabilize parallel dimers,^{42,79} and leucine, as the most frequently found aliphatic amino acid in hydrophobic cores,⁸⁰ in all d -positions.

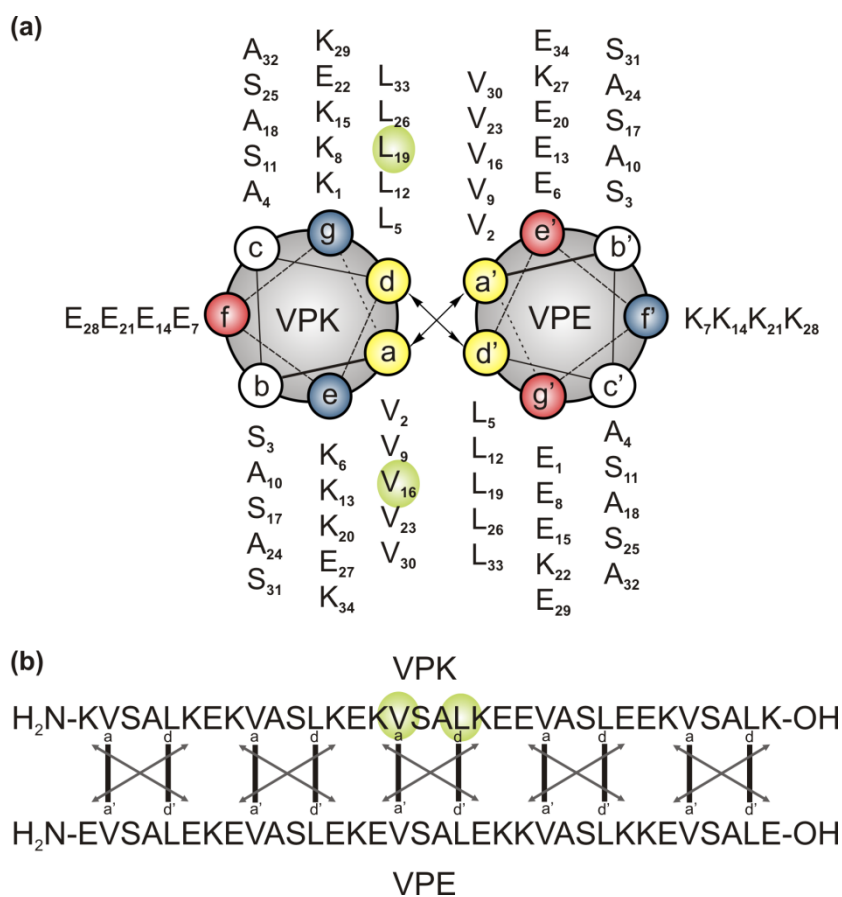


Figure 2.4: (a) Helical wheel representation of the parallel VPE/VPK heterodimer. (b) Amino acid sequence of the VPE/VPK system. Substitution positions highlighted in green.

An additional stabilization element for heteromeric parallel orientation was provided through the exchange of lysine/glutamic acid against glutamic acid/lysine in positions g_{22} and e_{27} of VPK and g'_{22} and e'_{27} in VPE, respectively. To prevent interhelical electrostatic repulsion positions a and b were filled with alanine and serine in an alternating manner. Alanine serves as a helix inducer and serine provides for solubility in aqueous medium.^{81,82} The f -positions of VPK are filled with glutamic acid, and with lysine in VPE, to reduce their overall charge at neutral pH. Parallel heterodimerization was verified by conducting FRET assays in combination with sedimentation velocity and equilibrium experiments.²⁵

Acid-pp/Base-pp

The α -helical coiled coil forming Acid-pp/Base-pp model system was established in the Koksch group to investigate the effects of substitutions with backbone extended amino acids. ^{83,84} In this system, one peptide carries exclusively lysine residues in *e*- and *g*-positions (denoted Base-pp), and a second peptide, which possesses exclusively glutamic acid residues in *e*- and *g*-positions (Acid-pp), serves as the interaction partner for Base-pp (Figure 2.5). Each peptide is 35 amino acids in length and therefore includes five heptad repeats. Heterooligomerization is driven by the burial of hydrophobic surface area contributed by Leu residues in *a*- and *d*-positions, and is directed by electrostatic interactions between oppositely charged amino acids that flank the resulting hydrophobic core.

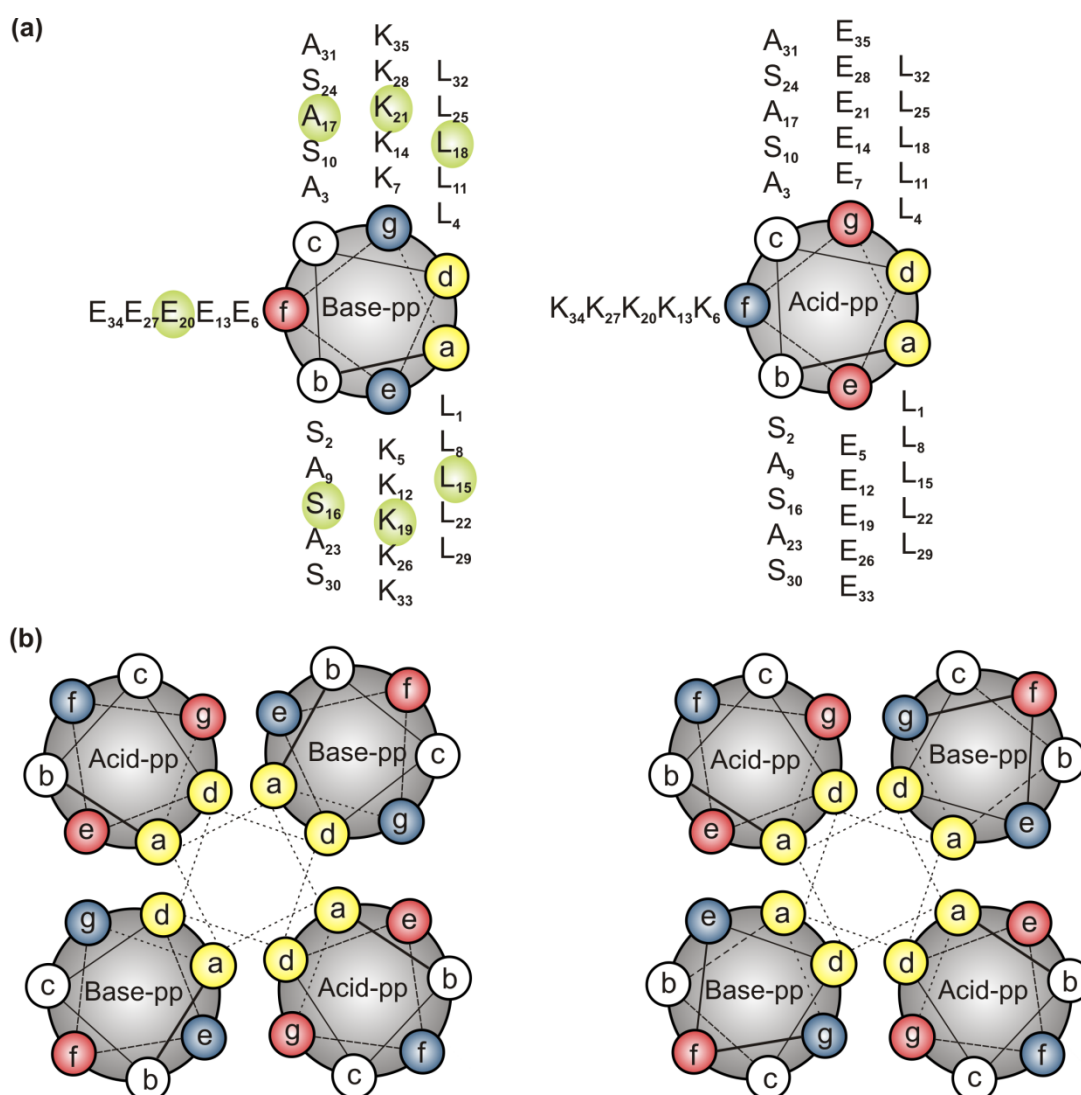


Figure 2.5: (a) Helical wheel representation of Acid-pp/Base-pp system. Substitution positions highlighted in green. (b) tetrameric coiled coil arranged parallel (left) and antiparallel (right), respectively. (Figure 4b is adapted from Merkel et al.) ⁸⁵

While CD spectroscopy indicates mostly unfolded confirmations for the individual peptides, an equimolar mixture of Acid-pp and Base-pp shows significant α -helical structure formation. ⁸⁴ Moreover, size exclusion chromatography revealed the formation of four-helix-

bundles. In this arrangement, the side-chain methylene groups of residues at *e*- and *g*-positions pack against hydrophobic residues at *a*- and *d*-positions broadening the hydrophobic interface, and thus contributing substantially to the stability of the bundle. Molecular dynamics simulations indicate that the formation of heterotetrameric system is possible in a parallel as well as an antiparallel helix alignment. Moreover, Rezaei Araghi *et al.* could show that this system tolerates the substitution of a central heptad repeat unit of Base-pp with a pentad of alternating β - and γ -amino acids while retaining the global conformation of the fold (*vide infra*; see chapter 2.2.2).

2.2 Modifying Helical Structures with Nonnatural Amino Acids

Since the α -helix comprises the largest class of protein secondary structures and plays an integral role in many important protein-protein interactions (PPIs), it has been of extensive interest as the minimal unit toward the development of interfering compounds.^{6,7,86,87} Short helical segments of a protein usually show little or no helical character when isolated from the stabilizing protein context, and these peptides are in general inherently unstable, suffering from storage and stability problems *in vitro* and degradation *in vivo*. One major goal of the chemical biology community has been to develop efficient means of mimicking this protein secondary structure, such that the resulting product resists proteolytic degradation, penetrates cell membranes, and presents functional groups in a multivalent manner over a large surface area. In this regard, the incorporation of nonnatural amino acids into peptides and proteins, as well as the development of unnatural oligomers with strong and predictable conformation propensities, denoted foldamers, has enhanced the power and versatility of bioactive peptides and peptidomimetics.

2.2.1 Fluorinated Hydrophobic Amino Acids

Due to its unique stereoelectronic properties, namely its small size, extremely low polarizability, and the strongest inductive effect among all chemical elements,⁸⁸ fluorine has become a powerful tool for modulating the properties of pharmaceuticals and biologically active compounds. In this respect, the incorporation of side chain fluorinated amino acids has been established as an efficient strategy to alter distinct functionalities of peptides and proteins, such as hydrophobicity, acidity/basicity, and conformation.¹⁴ Since a full review of fluorinated amino acids and their applications is beyond the scope of this thesis, properties of hydrophobic fluorinated amino acids that are of crucial importance for the understanding as well as some representative studies that demonstrate the diversity of fluorine's effects in helical peptide structures will be outlined in the following.*

General Aspects of Side-Chain Fluorination in Hydrophobic Amino Acids

A hydrocarbon chain or a similar nonpolar molecule is incapable of forming hydrogen bonds with water, and as a consequence its introduction into polar solvents results in the formation of a structured "cage" of water molecules (solvation shell) around its surface.⁸⁹ This phenomenon, denoted hydrophobic effect, is the reason for the burial of hydrophobic amino acids in the core of a protein. It is considered to be the major driving force during protein folding,⁹⁰ although the formation of hydrogen bonds between polar side chains as well as aromatic interactions are known to contribute considerably to protein stability.⁹¹⁻⁹³

After hydrogen, fluorine bears the second smallest atomic radius among all chemical elements, and therefore, the substitution of a hydrogen atom with fluorine is prevalently

*adapted from M. Salwiczek, E. K. Nyakatura, U. I. M. Gerling, S. Ye and B. Kokschi, *Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions*, *Chem. Soc. Rev.*, 2012, 41, 2135-2171.

considered to be nearly conservative or bioisosteric. As this holds true for single fluorine substitution, the volume of fluorocarbon groups seems to unproportionally increase with the number of fluorine atoms. Due to difficulties arising from interpreting differences between steric and electrostatic contribution, rotation and shape, estimations of the molar volume of such groups vary in literature. Our current understanding suggests that a CF_3 group approximates twice the van der Waals volume of a CH_3 group (Figure 2.6a), while its steric effects are close to that of an isopropyl group or even larger substituents such as sec-butyl or cyclohexyl groups.^{94,95}

Size, or rather shape, and hydrophobicity of an amino acid's side chain determine its ability to engage in packing interactions with other residues. By plotting side chain van der Waals volumes of amino acids versus their retention time of a RP-HPLC experiment, Samsonov *et al.* investigated the relationship between size and hydrophobicity of various amino acids.^{96,97} In this experiment the non-polar phase of a reversed-phase column serves as a mimic of a biological membrane or the kind of hydrophobic interactions, that would be present in hydrophobic cores of proteins and in ligand–receptor binding.⁹⁸ Using analogues of α -L-aminobutyric acid, they showed that the impact of an increasing fluorine content on the hydrophobicity of aliphatic amino acids is indeed larger than the mere increase in steric size by side-chain elongation or branching (Figure 2.6b).⁹⁶

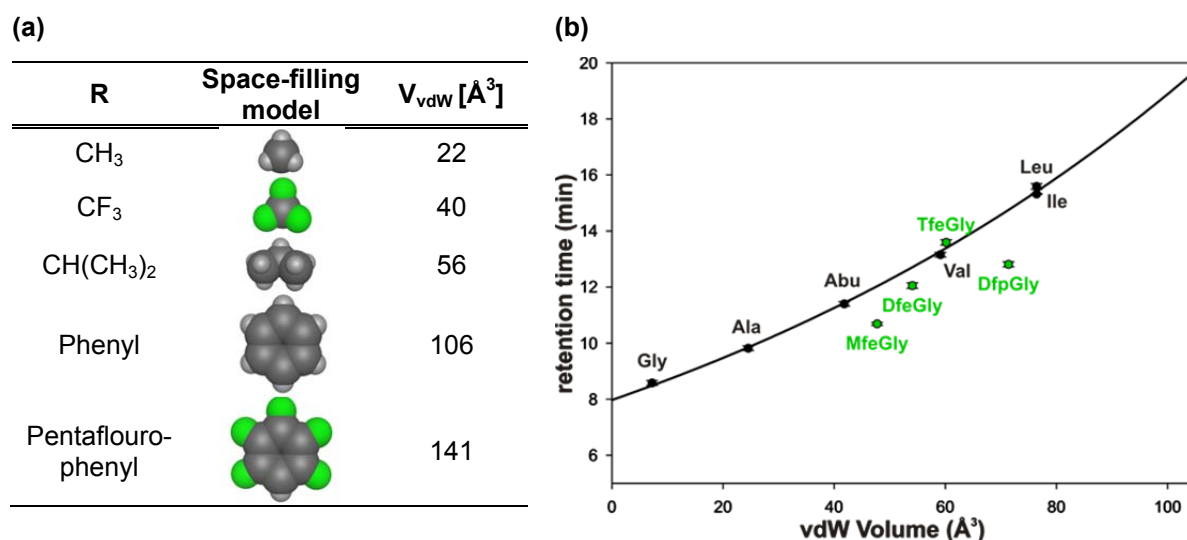


Figure 2.6: (a) comparison of van der Waals volumes.^{94,99,100} (b) Retention times of Fmoc amino acids plotted against the van der Waals volume of their side chains.⁹⁶ Nonfluorinated amino acids are depicted in black and their correlation is shown as a black line. Fluorinated amino acids are represented in green. (S)-2 aminobutyric acid (Abu), (S)-4-monofluoroethylglycine (MfeGly), (S)-4,4-difluoroethylglycine (DfeGly), (S)-4,4,4-trifluoroethylglycine (TfeGly), and (S)-4,4-difluoropropyl-glycine (DfpGly).

It is a frequently produced argument that hydrogen-to-fluorine substitutions provide for increasing hydrophobicity of aliphatic hydrocarbons and thus may be used to stabilize the protein structure.¹⁴ However, due to an extremely high electronegativity of the fluorine atom (with a value of 4 highest in Pauling scale)¹⁰¹, the C–F bond exhibits a large dipole moment which is reversed in orientation compared to the rather small C–H dipole. As a consequence,

fluoroalkyl groups possess two seemingly contrary physicochemical properties, polarity *and* hydrophobicity. Single aliphatic fluorine substitutions actually reduce rather than increase the overall hydrophobicity (MfeGly in comparison to Abu), and a critical number of fluorine atoms per alkyl chain is required to achieve what could be called ‘hyper-hydrophobicity’ (e.g TfeGly; Figure 2.6b).^{14,96} One plausible reason for the extreme hydrophobicity of highly fluorinated aliphatic hydrocarbons is fluorine’s inherently low polarizability, which makes it reluctant to engage in dispersive interactions with both water and other hydrocarbons. Therefore, perfluorocarbons tend to segregate from hydrophilic *and* lipophilic environments (referred to as ‘fluorous effect’)¹⁰². Despite its strong dipole, the C–F bond hardly ever acts as a hydrogen bond acceptor. However, it has been found that orthogonal polar interactions with carbonyl groups frequently occur in small molecule–protein co-crystals.¹⁰³

Substituting hydrogen with fluorine on the aromatic ring of phenylalanine leads to a rearrangement of the electrostatic potential and increases the hydrophobicity of the aryl side chain. While single hydrogen substitutions cause negligible perturbation,¹⁰⁴ the increased size of pentafluorophenylalanine may lead to destabilization of proteins owing to steric repulsion (molecular volumes of C₆H₆ and C₆F₅H are 106Å³ and 141Å³).¹⁰⁰

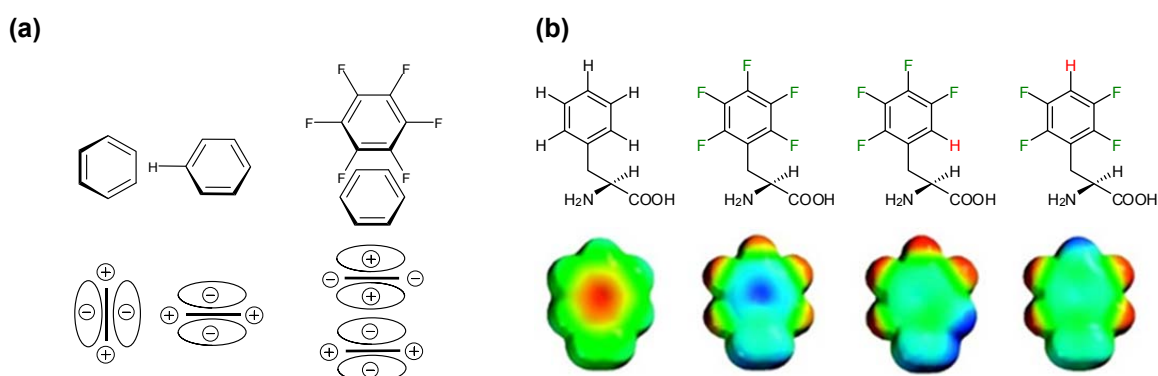


Figure 2.7: (a) π -Stacking (b) structures and space filling models depicting partial charges and electrostatic potentials of penta- and tetrafluorinated phenylalanine analoges. Negative and positive electrostatic potentials are indicated in red and blue, respectively. Adapted from Zheng et al.¹⁰⁵

Aromatic compounds possess a quadrupole moment which arises from the positively charged σ -framework between two regions of π -electron density on the faces of the ring (Figure 2.7).¹⁰⁶ Thus, the electrostatic potential of aromatic compounds is usually negative inside the aromatic ring, while the hydrogen atoms on the outside are partially positive charged. As a result, a face-to-face stacking arrangement of two hydrocarbon aromatic rings is repulsive, whereas edge-to-face interactions are favored (Figure 2.7a). Owing to the electron withdrawing effect of the fluorine atoms, the quadrupole moment of perfluoroaromatic molecules is of opposite polarity. Consequently, the net electrostatic attraction between parallel π -faces of phenylalanine and perfluorophenylalanine side chains leads to a fairly strong noncovalent interaction.^{93,107-110}

Incorporating Aliphatic Fluorinated Amino Acids in Coiled Coils

The intrinsic propensity of amino acids to promote certain secondary structures is an important factor in peptide and protein folding. Cheng *et al.* showed that when an amino acid of interest is introduced at a guest position of a monomeric α -helical alanine based model peptide, the α -helix propensity can be calculated from CD data. They found that fluorinated analogues of leucine and 2-aminobutyric acid exhibit lower α -helix propensities than their hydrocarbon analogues.^{111,112} However, while the reduced helix propensity may in part be attributed to unfavorable solvent interactions at exposed positions of the monomeric helix, it is not clear as to what extent helix propensity affects protein stability at buried positions, for example, within hydrophobic cores of proteins.⁹⁷

Since coiled coil formation is directed by the hydrophobic effect, i.e. the segregation of nonpolar side chains from water resulting in a hydrophobic helical interface, it is conceivable that substitution of hydrophobic core residues by fluorinated analogues might still stabilize these structures. Thus, coiled coil peptides have gained widespread attention as model systems for studying the impact of fluorination on hydrophobic protein-protein interactions.¹⁴

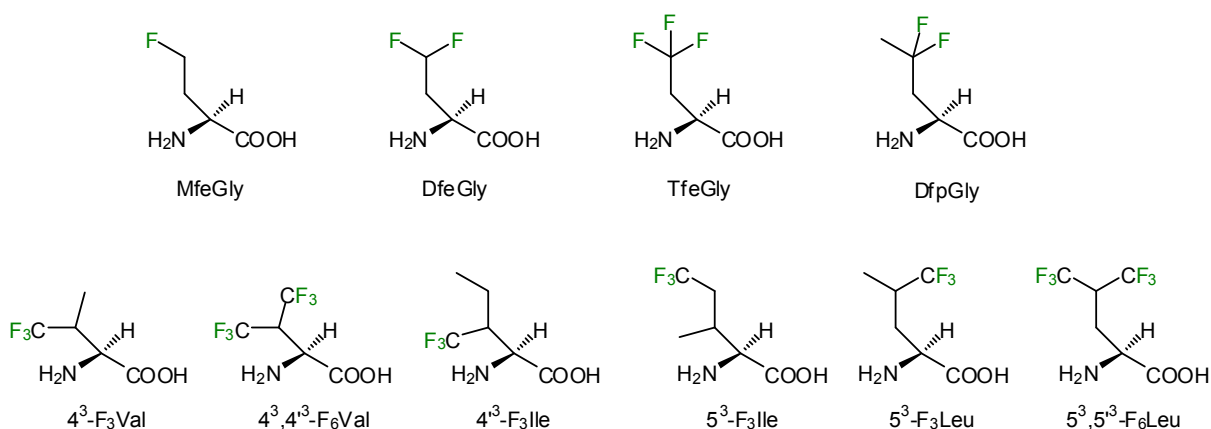


Figure 2.8: Side chain fluorinated analogues of here discussed aliphatic L-amino acids.

Tirrell and colleagues investigated the effects of substituting all leucine residues in *d*-positions of a GCN4 subdomain with Trifluoroleucine (5³-F₃Leu),⁷² and all valine residues in *a*-positions in a slightly adapted sequence of the same DNA binding protein with trifluoroisoleucine (5³-F₃Ile) or trifluorovaline (3*R*)-4³-F₃Val),¹¹³ respectively. Fluorination of this DNA-transcription factor left its DNA-binding affinities and specificities unchanged, while resulting in coiled-coil structures that are highly resistant to both increasing temperature and denaturant. Remarkably, when comparing Ile and Val analogues, the replacement of the δ -CH₃ group of Ile by CF₃ showed an approximately eight-fold higher stabilization than the replacement of the γ -CH₃ group of Val.¹¹³ This finding might be explained with a substantial loss of side-chain entropy of Val due to unfavorable steric interactions of the large γ -CF₃ group with the helix backbone.

When Kumar and colleagues compared the introduction of diastereomeric $5^3\text{-F}_3\text{Leu}$ and $5^3,5'^3\text{-F}_6\text{Leu}$ at d -positions of a dimeric coiled-coil model peptide,^{114,115} a substantial increase in thermal stability resulted from substituting both $\delta\text{-CH}_3$ groups of leucine with CF_3 . This indicates that stability and fluorine content closely correlate. In light of these findings, the observation made in follow up studies that complete fluorination of a GCN4 hydrophobic core (α -positions with $4^3\text{-F}_3\text{Val}$ and simultaneously d -position with $5^3\text{-F}_3\text{Leu}$) leads to increased stabilities,²⁰ is not surprising. It is conceivable that the additional stability is derived from sequestering the more hydrophobic trifluoromethyl groups from aqueous solvent. However, another explanation for the stabilization of above mentioned hydrophobic domains might be given in the fluorous effect, i.e. the phenomenon that highly fluorinated alkyl groups possess increased hydrophobicities and lipophobicities likewise. To test this hypothesis, the same group of scientists generated two helical assemblies, one that carried Leu in all α - and d -positions (peptide HH) and one that carried its hexafluorinated analogue in all core positions (peptide FF).^{21,70} The incubation of disulfide bridged HF-dimers led to a preferential reorganization of FF- and HH-homodimers, suggesting a strong impact of the 'fluorous effect', since mixed hydrocarbon-fluorocarbon cores seem to be disfavored.

The latter findings have been challenged by studies undertaken in the group of N. Marsh. Utilizing a tetrameric, antiparallel coiled-coil model assembly, this group investigated the effects of gradual substitutions of hydrophobic core leucine residues with $5^3,5'^3\text{-F}_6\text{Leu}$.¹⁹ The increase in stability that was observed upon fluorination of the coiled coils two central layers almost exactly matches the partitioning of $5^3,5'^3\text{-F}_6\text{Leu}$ into organic solvents. Consequently, the authors argued that stability improvements result from increased hydrophobicities of fluorinated amino acids rather than specific 'fluorine-fluorine interactions'. Moreover, the substitution of two additional layers within the hydrophobic core resulted in a less pronounced increase of stability than observed for the first two layers.¹⁸ NMR studies indicated that an increasing degree of fluorination results in a more rigid backbone and less dynamic hydrophobic core. Hence, interactions between the fluorinated residues do not seem to add stability, while global substitution of leucine residues with $5^3,5'^3\text{-F}_6\text{Leu}$ results in crowding of fluorinated analogues within the core.

Additionally, the authors studied tetrameric bundles, in which leucine and $5^3,5'^3\text{-F}_6\text{Leu}$ pack against each other in an alternating fashion.¹⁶ They found that this packing arrangement leads to greater stabilities than in the fully fluorinated analogue, and explained their observations with a more efficient packing of the fluorinated amino acids in the hydrophobic core. Their recently solved X-ray structures of three tetrameric coiled coil peptides exhibiting hydrophobic cores which are packed with either fluorocarbon or hydrocarbon side chains strengthen this hypothesis.¹⁷ It was shown that $5^3,5'^3\text{-F}_6\text{Leu}$, though larger, is well accommodated within the protein with minimal structural perturbation, because it closely resembles the shape of the hydrocarbon side chain that it substitutes. Obtained

crystal structures reveal no evidence for preferential fluororous interactions between fluorinated residues. Instead, the same principles that underpin the stability of natural proteins (efficient packing of side chains and conventional hydrophobic effects) seem to be responsible for the enhanced stability of the $5^3,5^3$ -F₆Leu containing assembly. Moreover, this study showed that global fluorination of the hydrophobic core does not stabilize coiled coils *per se*, because a peptide possessing a fluorinated core which is constituted of TfeGly turned out to be slightly less stable than its natural counterpart that harbors leucine at substitution positions. Thus, it seems that if changes in residue size are controlled for, extensively fluorinated proteins can be designed that closely preserve the shape of side chains, enabling accurate packing while simultaneously increasing size and hydrophobicity.

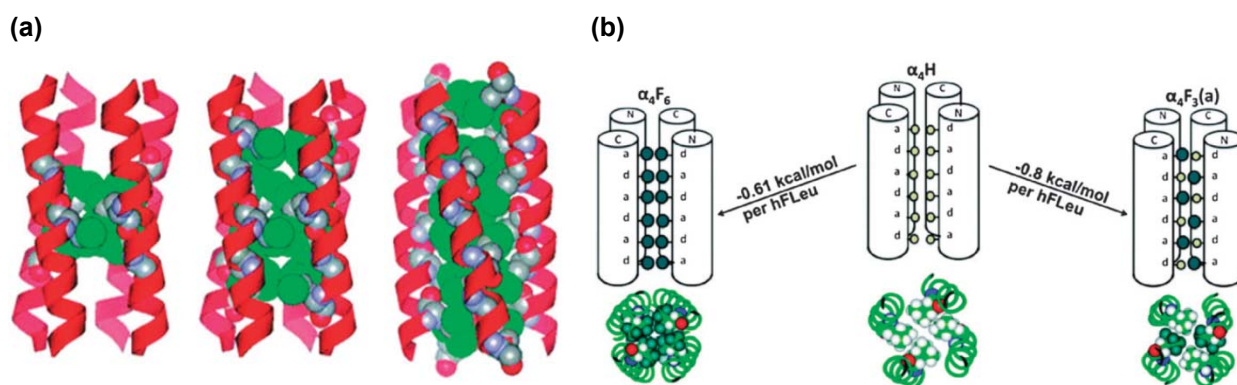


Figure 2.9: (a) Models of the four-helix bundle containing $5^3,5^3$ -F₆Leu in two, four and six layers of the hydrophobic core. Adapted from Lee et al.¹⁸ (b) Schematic representation of the tetrameric helical bundle showing the different substitution patterns and the per $5^3,5^3$ -F₆Leu residue stabilization involved. Reproduced with permission from Buer et al. (2009).¹⁶

All above mentioned studies address the partial and global fluorination of the hydrophobic core of coiled-coil peptides. In contrast to this, the Koksich group investigates single substitutions within the hydrophobic core with different analogues of (S)-2-aminobutyric acid that possess increasing fluorine stoichiometry and side-chain volume. To this end, two different coiled coil based model systems were *de novo* designed. An antiparallel homodimer,^{23,24} and the parallel heterodimeric VPE/VPK bundle (see Figure 2.4).^{25,97} In both models the substituted amino acid interacts exclusively with native hydrophobic amino acids. Thus, these systems can be utilized to investigate the impact of a gradual fluorine content increase on the stability of this folding motif. Single substitutions of difluoroethylglycine (DfeGly), trifluoroethylglycine (TfeGly), or difluoropropylglycine (DfpGly) for valine or leucine residues at central hydrophobic core positions yielded less stable assemblies than their native coiled-coil analogues. This might be attributed to generally lower hydrophobicities (except for TfeGly when substituted for valine Figure 2.6b), and different shapes as well as reduced helix propensities in comparison to substituted amino acids.

Moreover, the effects of single amino acid substitutions were shown to be strongly position dependent (Figure 2.10). Most pronounced differences were observed for DfpGly.²⁵

When incorporated at an *a*-position of the antiparallel assembly, this fluorinated building block is strongly destabilizing, while it yields the most stable fluorinated parallel heterodimer. These findings suggest that the orientation and flexibility of fluorinated side chains within a certain protein environment determine the impact of fluorine-induced polarity. MD studies by Pendley *et al.*,¹¹⁶ which also revealed an important role of electrostatics in the stability of parallel coiled coil systems containing fluorinated amino acid residues $5^3,5'^3$ -F₆Leu in the hydrophobic core, support these conclusions.

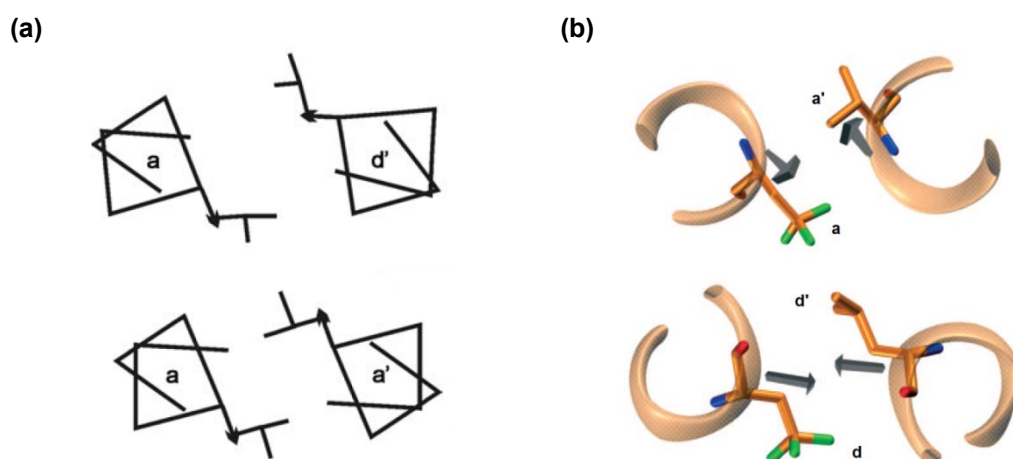


Figure 2.10: (a) Differences in packing of *a*-position in antiparallel (top) and parallel (bottom) coiled coil dimers. (b) Snapshots from MD simulations show the different packings of a fluorinated side chain (TfeGly) with its directly interacting partner within the parallel coiled-coil model system at *a*-position (top) and *d*-position (bottom). Reproduced with permission from Salwiczek *et al.* (2009).²⁵

However, at *a*-positions of parallel coiled-coil heterodimers stability largely correlates with the size of the side chain, while this correlation could not be observed when substituting in *d*-positions of the parallel heterodimer.²⁵ Due to a fundamentally different packing arrangement at the two hydrophobic positions of parallel coiled coils (see chapter 2.1.1), the through fluorine substitutions strongly polarized β -methylene groups point away from the hydrophobic core at the *a*-position of the parallel dimer (Figure 2.10b). At the *d*-position they point directly into the core and thereby disturb hydrophobic interactions. Hence, fluorine induced polarity has a stronger destabilizing effect at the *d*-position while at the *a*-position the size of the side chain is the prevailing factor that determines stability.

These results demonstrate that the effect of partly fluorinated side chains strongly depends on their packing/orientation within the hydrophobic core. Neither hydrophobicity nor side-chain volume alone dictate stability of coiled-coil interactions and within the framework of packing characteristics it may be more adequately described as a cooperative effect of both. Moreover, these studies highlight that already a single mutation in the hydrophobic core of a well packed protein can have a significant effect on stability, depending on the microenvironment surrounding the substitution site.

Fluorinated Phenylalanine Influencing Quadrupole Interactions in Helical Assemblies

Interactions between aromatic side chains involve hydrophobic, van der Waals, and electrostatic forces.^{109,117} Hydrophobic and van der Waals forces lead to strong interactions with other hydrophobic or aromatic side chains, while the electrostatic component is assumed to contribute geometric preference to the interaction. One example involving quadrupole interactions of amino acid side chains is given by the villin headpiece subdomain (Figure 2.11a), a 35-residue three helix bundle structure comprising a hydrophobic core, which features a cluster of three edge-to-face oriented phenylalanines.

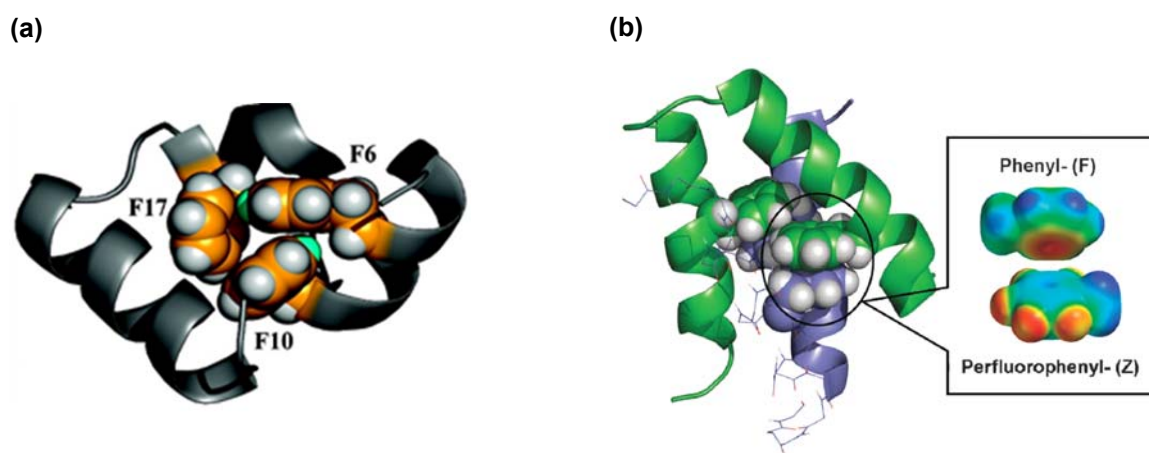


Figure 2.11: (a) Structure of villin head piece (PDB: 1YRF) highlighting the interaction of phenylalanines in the hydrophobic core. Reproduced with permission from Zheng et al.(2009).¹⁰⁵ (b) space-filling model highlighting electrostatic potentials of face-to-face stacked phenylalanine and perfluorophenylalanine side chains in model peptide. Blue indicates a positive and red a negative electrostatic potential. Reproduced with permission from Zheng et al. (2010).¹⁰⁷

Zheng *et al.* used two tetrafluorinated phenylalanine isomers (2,3,4,5-F₄Phe and 2,3,5,6-F₄Phe) to substitute Phe residues of VHP, thereby aiming at an enhanced hydrophobicity as well as a retained possibility to undergo aromatic interactions with the remaining hydrogen atom of phenylalanines aromatic side chain.¹⁰⁵ Indeed, calculated partial charges of the remaining hydrogen atoms indicate a substantial increase (Figure 2.7), and as hypothesized investigated mutants adopt stable and native-like structures. Moreover, gained results indicate that tetrafluorophenylalanines contribute more favorably to VHP's stability than pentafluoro-phenylalanine, which points to retained edge-to-face quadrupole interactions.

Additionally, Zheng and Gao recently investigated a model peptide that folds into a homodimer comprising a hydrophobic core with one phenylalanine residue from one monomer stacking face-to-face against a phenylalanine side chain of the other monomer. With the help of this model, they were able to demonstrate that an accurate positioning of aromatic and perfluoroaromatic rings can enable selective protein-protein interactions.^{107,118} A double mutant in which both phenylalanines are replaced by its perfluorinated analogue exhibits a tremendous increase in thermal stability ($T_M = 80$ °C as opposed to 30 °C of parental peptide), which is assumed to result from the increased hydrophobicity of the

fluorine atoms in comparison to hydrogen atoms. Moreover, when the two dimers are mixed they undergo a complete transition from homodimers to heterodimers.¹⁰⁷ The specificity of the heterodimer is expected to be an effect of the quadrupole interaction of stacked aromatic and perfluoroaromatic rings, whose electrostatic attraction was calculated to contribute as much as -1.0kcal/mol to protein structural stability.

The aromatic stacking properties of pentafluorophenylalanine have also been used to enhance Phe/Phe interactions in a collagen model peptide to facilitate self-assembly into fibrils.¹¹⁹ Collagens constitute a ubiquitous protein family that can aid preventing bleeding and promote tissue repair in injured vessel walls by binding to the collagen binding receptor on platelets, which in turn induces platelet aggregation and thus promotes blood clotting.¹²⁰ They possess a rope-like triple helix fibril structure of proline-rich polypeptide strands comprising a multiple glycine-proline-hydroxyproline repeat (G-P-O)_n in their core. By attaching F₅Phe and Phe to the N- and C-termini of a (G-P-O)₁₀ collagen model peptide, respectively, Cejas *et al.* were able to demonstrate that aromatic stacking interactions can be used to encourage propagation by end-to-end stacking into fibrils similar to that of collagen.¹¹⁹ Calculating the energetics for head-to-tail stacking revealed significantly higher binding energies for triple helices comprising fluorinated residues than for their hydrocarbon analogues, and their ability to mimic collagen's biological function was confirmed.

Taken together these studies demonstrate the necessity of tertiary constructs where the static geometry provides for the correct stacking configuration of aromatic side chains, and furthermore show that the quadrupole interaction of the phenyl-pentafluorophenyl pair can be strong enough to direct protein-protein interactions and fibril formation of collagen peptide mimics.

2.2.2 Peptidomimetic Foldamers Adopting Helical Structures

The term foldamer was first proposed by Gellman in 1996 to describe "any polymer with a strong tendency to adopt a specific compact conformation".^{27,121} It thus comprises all unnatural oligomers that display conformational propensities similar to those of proteins and nucleic acids.¹²² Since precise global conformations are crucial for the distinct biological function of proteins, their folding with high accuracy is indispensable. A great deal of effort has been made to extend this relationship between folding and function to unnatural oligomers, with the aim of not only achieving biopolymer-like shape but also activity.^{27,28}

Peptidomimetics constitute the majority of foldameric oligomers and can be generated through backbone modifications of peptides, such as an isosteric or isoelectronic amid bond substitution or the introduction of additional fragments. The synthesis of peptidomimetics possessing homologated amino acids, i. e. residues in which a variable number of CH₂ units is introduced between the amino and carboxylic acid groups involved in the peptide bond, has afforded molecules that are capable to adopt well-defined helical conformations and

exhibit proteolytic stability.^{29,30,123} While homogeneous systems, e.g. exclusively composed of β -amino acids (β -peptides)^{26,124} or γ -amino acids (γ -peptides)¹²⁵, determined early foldamer studies, recently a number of foldamers with heterogeneous backbones, that is foldamers that contain more than one type of subunit, have been reported.¹²²

Conformational Aspects of Backbone Homologated Amino Acids

Folded polypeptide structures of α -amino acids are conventionally defined using backbone torsion angles as descriptors of local conformations.³⁰ Thus, each α -amino acid in a peptide chain possesses two degrees of torsional freedom about the N-C $^{\alpha}$ (Φ) and C $^{\alpha}$ -CO (Ψ) bonds, with the peptide unit largely restricted to a trans, planar ($\omega \sim 180^\circ$) conformation. Omega (ω) amino acids are defined as residues in which a variable number of backbone atoms is introduced between the flanking peptide units. This backbone elongation enhances the number of degrees of torsional freedom, and hence results in an expansion of energetically accessible conformational space. Consequently, in β -amino acids the local backbone conformations are determined by three torsional variables (Φ, Θ, Ψ) and in γ -residues to number of torsional variables is extended to four ($\Phi, \Theta_1, \Theta_2, \Psi$). In α -amino acids the sterically allowed region of space diminishes drastically when going from C $^{\alpha}$ unsubstituted glycine via monosubstituted alanine to disubstituted α -aminoisobutyric acid (Aib).¹²³ Accordingly, unsubstituted ω -amino acids are expected to be able to exhibit a larger extent of conformational space than their substituted counterparts. In fact, helical β -peptides have been long thought to be unable to acquire ordered conformations, due to the entropic effect of the additional single bond.¹²⁶

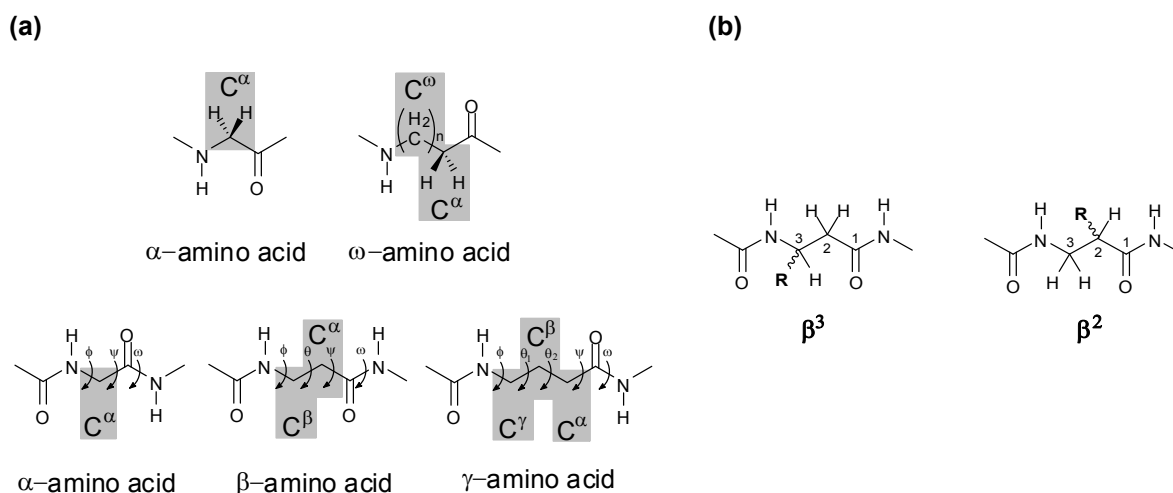


Figure 2.12: (a) chemical structures of α - and ω -amino acid residues (top) and definition of backbone torsion angles of a, b, and γ residues (bottom) (b) Substitution possibilities for monosubstitution of β -amino acids. Numbering of C atoms from the C-terminus is indicated. The substitution can be on C2 (β^2) or C3 (β^3) and in each case, depending on the configuration at the C2 or C3 atoms, two different possibilities (2S and 2R or 3S and 3R) arise.

However, substituting at the backbone carbon atom in ω -amino acids generally results in the creation of new chiral centers, which can restrict the torsional freedom and hence influence

the handedness of folded backbone conformations (Figure 2.12b).³⁰ To further constrain conformational space, different strategies including di- and germinal substitutions as well as the introduction of cyclic amino acids can be applied. First proof for folded structures in β -peptides were derived from the work of Seebach and co-workers, in which they generated folded intramolecularly hydrogen bonded structures of short sequences containing monosubstituted β^3 -residues.¹²⁴

Today numerous monosubstituted β - and γ -residues, especially those with bulky substituents, are known to adopt gauche conformation (Θ and $\Theta_2 = 60^\circ$, respectively).³⁰ These residues are generally observed in helices. The respective trans rotamers (Θ or $\Theta_2 = 180^\circ$) lead to fully extended backbone conformations and are thus readily accommodated in β -strand structures. However, several ω -amino acids have been characterized in both, peptide sequences that adopt helical or hairpin conformations in the crystalline state, so that conformational choices are likely to be influenced by the nature of the cooperative hydrogen bond interactions formed during folding.

Structural Implications of Introducing β - and γ -Amino Acids

Similar to secondary structures formed by α -polypeptide chains, the adopted secondary structure of poly β - and γ -amino acids are principally determined by intramolecular hydrogen bonding.¹²² In conventional peptides, the main conformations are 3_{10} - and α -helices, with 10 and 13 atoms in the hydrogen bonded rings, respectively, while there is a facile interconversion between these two on a relatively flat energy landscape.¹²⁷ Hydrogen bonds in α -helices exhibit the directionality $\text{CO}(i)\cdots\text{NH}(i+4)$, resulting in a hydrogen bonded ring 13 atoms (C_{13}) in length. The acceptor lies towards the N-terminus of the sequence, while the donor NH group lies towards the C-terminal end.

From structural investigations dealing with oligomers of homologated L-amino acids that fold into helices in solution, three main conclusions can be drawn (Figure 2.13):¹²⁸

- (i) On progressing from α - to β - to γ -peptides, the stability of the helices increases, although the number of hydrogen bond donors and acceptors per chain atom decreases.
- (ii) The helicity reverses upon each homologation step: right handed for C_{13} - α -, left handed for C_{14} - β -, and again right-handed for C_{14} - γ -helical peptides.
- (iii) The direction of the helix macrodipole also reverses with homologation: it points from N- to C- in α -, from C- to N- in β -, and again from N- to C-terminus in γ -peptides.

Among documented secondary structures of β -peptides the 14-helix (C_{14}) has emerged as the most frequently observed conformation.¹²⁸ It is stabilized by hydrogen bonded rings between $\text{N-H}(i)$ and $\text{C=O}(i+2)$ in a three residue repeat arrangement. Deviating from an

idealized left handed C_{14} -helix, substituents in the (i) and ($i+3$) positions are not exactly positions on top of each other, but offset by ca 15° , resulting in a '3.1 helix'.

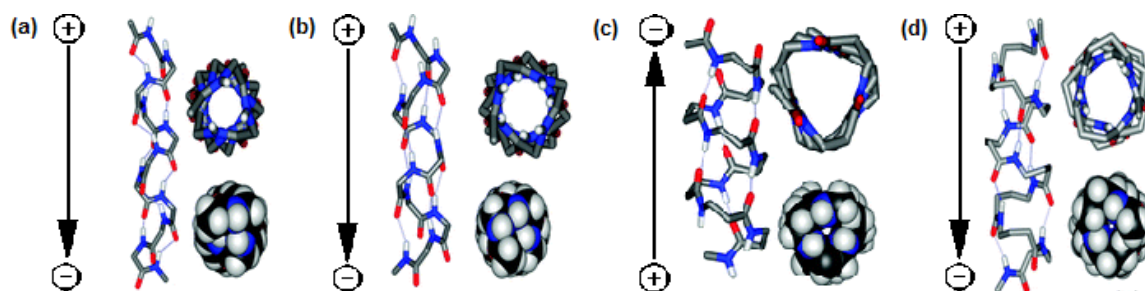


Figure 2.13: (a) 3_{10} helix; (b) C_{13} helix (α -peptide); (c) C_{14} helix (β -peptide); (d) C_{14} helix (γ -peptide)
Adapted from Vasudev et al..³⁰

A crystal structure of a β -peptide revealed that hydrogen atoms on unsubstituted C2 backbone methylene groups closely approach main chain carbonyls, so that $\text{CH}\cdots\text{O}=\text{C}$ hydrogen bonds are feasible.¹²⁹ Although these hydrogen bonds are weaker than those involving amide protons, they might contribute to stability and thus partially explain the greater conformational stability observed for short β -peptides vs α -peptides.^{127,129} For this secondary structure to be observed in solution a chain length of only six residues is required.¹²⁸ Homogenous double homologation of proteinogenic amino acids in peptides leads to γ -peptides. These oligomers most commonly form C_{14} -helices in solution, which result from hydrogen bonds between $\text{N-H}(i)$ and $\text{C}=\text{O}(i+3)$. In γ - C_{14} -helices 2.6 residues comprise one turn and as little as four amino acid units are required to be able to observe a helical structure in NMR.

Two distinct approaches to heterogeneous backbone foldamers are currently applied.¹²² One approach involves oligomers composed of multiple 'blocks', each containing different subunits and thus resulting in chimeras of α -peptide and foldamer, and the other approach is to examine backbones with a regular pattern of subunit alteration. Both approaches can be combined in various ways, facilitating access to hybrid foldamer systems with broad side chain diversity. The extension from homogeneous to heterogeneous backbones, comprising different types of homologated amino acids in the same strand, dramatically increases the number of possible H-bond networks in resulting helices. However, many of these structures may be anticipated as simple modifications of the common 3_{10} - or α -helices of conventional peptides; the 3_{10} helix (with hydrogen bonding between the amide proton of residue i and the carbonyl of $i+3$) expands from repeating 10-atom hydrogen rings to 11-atom rings in $(\alpha\text{-}\beta)_n$ peptides, while the α -helix hydrogen bond pattern ($i+4 \rightarrow i$) expands from 13 to 14 atoms in $(\alpha\text{-}\alpha\text{-}\beta)_n$ peptides.^{127,130}

Of particular interest to the current work is the introduction of alternating β - and γ -amino acid residues in α -peptide chains. Regarding an equal number of backbone atoms, a dimer of one β - and one γ -residue can be substituted for a trimer of α -amino acids (Figure 2.14). Moreover, *ab initio* MO studies have shown that a β/γ -peptide can form a helix composed of

a 13-membered ring with an H-bond pattern ($i + 3 \rightarrow i$) that is similar to the one found in α -helices.¹³¹ Hofmann *et al.* predicted the possibility of this C_{13} -helix in aqueous environment via MO-calculations and highlighted its similarities to the native α -helix in terms of structure, H-bond pattern and dipole orientation. Thereafter, Gellmann *et al.*¹³² reported the finding of C_{13} -helices in crystal structures and NMR-analysis of short β/γ -amino acid tetra- to hexamers with constrained backbone due to cyclization. Obtained structures confirmed the theoretical finding of Hofmann *et. al* in terms of structure and H-bond pattern of this helix.

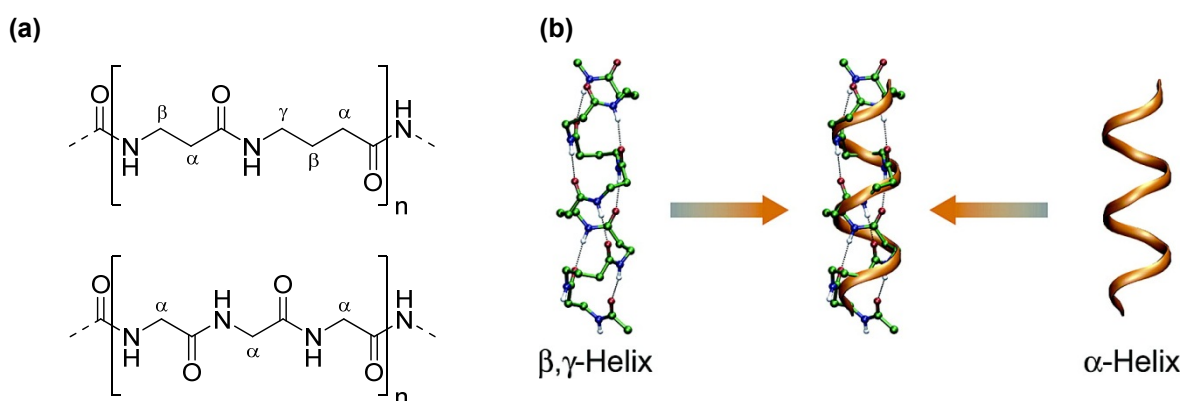


Figure 2.14: (a) an alternating sequence of α - and β -amino acid dimers (top) replaces α -amino acid trimers (bottom). (b) Helices of β,γ -hybrid peptides can mimic the α -helix formed by native α -peptides. Reproduced with permission from Baldauf *et al.* (2006)¹³¹

Helix-Bundle Formation Involving β/γ Amino Acids

While the design of foldamers that adopt well-defined secondary structures has been successfully demonstrated, many biological functions require the greater functional complexity available from tertiary and quaternary structures.¹²⁷ The design of oligomers that fold into unique three-dimensional structures is a great challenge in foldamer research, which probes our understanding of the mechanisms by which natural protein sequences fold while also laying the groundwork for the design of complex protein-like functions.

Inspired by the leucine zipper motif found in coiled-coil proteins, Shepartz *et al.* reported a β -dodecapeptide that self-assembles in aqueous solution into a highly thermostable octameric bundle and determined its structure by X-ray crystallography.¹²⁹ Residues along one helical face of this β -peptide promote the formation of a 14-atom helix through salt-bridge formation, while those on a second use β^3 -homoleucines to drive oligomerization. Self-complementary charges were introduced on a last phase to favor homo-oligomerization. The crystal structure revealed that the octamer bundle is composed of four copies of parallel helical dimers; two parallel dimers associate in an antiparallel fashion to create tetramers, which further associate into an octamer with a well-packed hydrophobic core akin to leucine zipper proteins. In cooperation with the group of Kumar *et al.* they could furthermore show that the substitution of one of four leucine residue with hexafluoroleucine in an analogue β -peptide results in a similarly well folded and thermostable octamer comprising a fluorous

core.¹³³ The incorporation of three hexafluoroleucine eliminated bundle formation, which might be attributed to overpacking of the hydrophobic core.

Gellman *et al* investigated the systematic substitution of β -amino acids into the dimeric coiled-coil subdomain of GCN4.¹³⁴ To minimize disruption within the hydrophobic core, substitutions were made along the polar side of the helix, resulting in a heptad repeat of $(\alpha_a\beta_b\alpha_c\alpha_d\alpha_e\beta_f\alpha_g)_n$. Although the substitutions destabilize the folding of the peptide in solution, it forms a parallel helix bundle in the crystal. Each bundle contains three foldameric peptides, instead of two as in the parental α -peptide coiled coil. Undertaking the same substitutions in a variant of GCN4 that forms tetrameric bundles resulted in trimeric foldamer assemblies in solution. But this α/β -peptide crystalized in the same tetrameric stoichiometry as its parent sequence. In both foldameric structures the β -amino acid residues are well accommodated, maintaining the overall helical geometry and hydrogen bonding pattern of the parental α -helices.^{127,134} They also show classical knobs-into-holes packing of the side chains between the helices and therefore retain the ability to form a discrete helix bundle quaternary structure. However, in each context, the substitutions subtly alter the angle of projection of the β -amino acid side chains and the helical packing geometries, thereby altering the specificity and stability of the peptides for a given association in solution. Encouraged by these findings Gellman *et al.* extended their studies to biological systems,^{53,135} including HIV's hexameric coiled coil envelope gp41 domain (see chapter 2.1.2). They could show that sequence-based $\alpha \rightarrow \beta^3$ replacements in combination with site-specific backbone rigidification via the introduction of cyclic β -residues within a gp41 C-peptide analogue, can lead to physical and biological mimicry of the critical gp41 subunit.⁵³ Physical studies in solution, crystallographic data, and results from cell-fusion and virus infectivity assays collectively indicate that the gp41-mimetic effectively blocks HIV-cell fusion, while being far less susceptible to proteolytic degradation than its analogous α -peptide.

As aforementioned, theoretical studies have shown that sequences composed of alternating β - and γ - amino acids are able to adopt helical structures similar in shape and H-bonding to the α -helix.¹³¹ Taking this into account and applying the equal backbone approach for substitutions with homologated amino acids, Koksche *et al.* substituted one coiled coil heptad repeat unit (abcdefg) of Base-pp (*cf.* chapter 2.1.3), comprising two 13-atom H-bonded turns of the α -helix, with a pentad of alternating β - and γ -amino acids.^{83,84} The assembly of the $\alpha\beta\gamma$ -chimera (denoted B3 β 2 γ) with a natural, oppositely charged Acid-pp α -peptide resulted in a mixed parallel- and antiparallel heterotetrameric leucine zipper albeit with reduced thermostability (Figure 2.15). Mutational analysis indicated that the specificity of the artificial coiled-coil assembly is driven by the interaction of β^3 -Leu residues in positions 15 and 17 of the chimera with side chains in positions a'_{15} and d'_{18} of Acid-pp. β^3 -Lys and γ^4 -Lys residues in positions 18 and 19 of the hybrid peptide form electrostatic interhelical interactions with negatively charged residues at positions e'_{19} and g'_{21} of the α -peptide.

tethered side chains on the same side of the helix. Filling these positions with nonnatural amino acids that possess an olefinic side chain and subsequent closure of a macrocyclic bridge of the nonnatural side chains via ruthenium-mediated ring-closing metathesis (RCM) has been successfully demonstrated to stabilize helices (Figure 2.16a).^{6,141,142} Verdine *et al.* developed a set of amino acids that capitalize two strategies to stabilize α -helical conformations at once.¹⁴² (i) Olefinic side chains of different length, which are able to crosslink via olefin metathesis if positioned in i and $i+4$ or $i+7$ distances. (ii) Substituting the C $^{\alpha}$ -hydrogen atom with a methyl group in order to exploit the helix stabilizing effect of α,α -disubstituted amino acids. The introduction of such macrocyclic so called ‘all-hydrocarbon staples’ into α -helical peptide scaffolds between two α -methyl, α -pentenyglycine residues at positions that lie on the same face of the helix via RCM can greatly increase α -helical content as well as chemical and thermal stability.^{7,143} Moreover, the incorporation of multiple staples can be utilized to rigidify a longer stretch of an α -helix and thereby to further increase protease stability.¹⁴⁴ Verdine *et al.* demonstrated that, if the stereochemical preference of the applied α,α -disubstituted amino acids is controlled for, even two $i,i+4$ staples can be formed in proximity on a single face of the helix (Figure 2.16a). Several biological active α -helical structures where stabilized by inserting one or two hydrocarbon staples within peptide sequences, that exhibit enhanced protease resistance as well as cell penetration, possibly owing to the lipophilic nature of the linker.^{68,145-148}

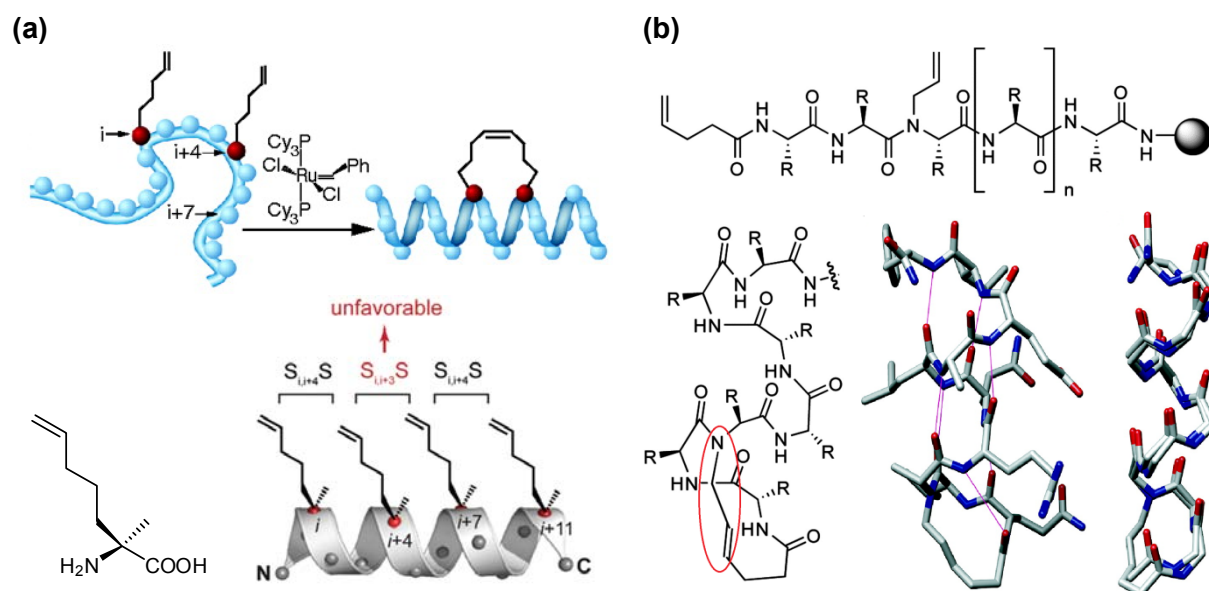


Figure 2.16: (a) Top: Ring-closing metathesis (RCM) of olefin-modified i and $i+4$ residues to generate hydrocarbon-stapled helices (top). Adapted from Walensky *et al.* (2004).^{6,146} Bottom: Structure of (S)- α -methyl, α -pentenyglycine unit (referred to as S5) used for hydrocarbon cross-linking (right) and schematic presentation of tandem ring-closing metathesis of a tetraolefinic peptide substrate (left). Since the RCM of amino acids in $S_{i,i+3}S$ relationship is stereochemical unfavorable, the double-RCM product is formed with 98% conversion. Adapted from Kim *et al.* (2010)¹⁴⁴ (b) Top: Exemplified structure of a peptide containing HBS residues on solid phase. Bottom: Putative i and $i+4$ hydrogen bonds (magenta lines) in crystal structure derived molecular model of HBS helix (middle), and overlay of crystal structure of a peptide containing HBS and a model of an idealized α -helix (right). Adapted from Henchey *et al.* (2008)⁵⁰

Another strategy for the stabilization of α -helices via RCM involves replacement of one of the main chain intramolecular hydrogen bonds with a covalent linkage (Figure 2.16b).^{7,149-151} High resolution crystal structures of a short peptide containing such a hydrogen bond surrogate (HBS) showed that this RCM-based macrocycle is capable of reproducing the conformation of a canonical α -helix. All i and $i + 4$ C=O and NH hydrogen-bonding partners fall within distances and angles expected for a fully hydrogen bonded α -helix. Thermal and chemical denaturation studies revealed that HBS stabilized peptides retain their conformation in aqueous buffers even at exceptionally high temperatures although they can be denatured with concentrated guanidinium chloride.⁷ Moreover, HBS helices are significantly more resistant to proteases than their unconstrained counterparts, which is not surprising considering their conformational stability. Arora *et al.* could show that HBS helices can stabilize biologically relevant peptides in helical conformations.^{150,151} These HBS α -helices can bind their expected protein receptor with high affinity while resisting trypsin-mediated proteolysis. An attractive feature of main chain hydrogen bond surrogates is that the placement of the crosslink inside the helix does not block solvent-exposed molecular recognition surfaces of the molecule, as compared to side chain crosslinking strategies. This property potentially allows HBS helices to target tight binding pockets on proteins.

2.3 Phage Display as an Example for Protein Evolution

Molecular recognition is essential to biology, and thus the discovery and characterization of interacting partners are critical endeavors for biological scientists.¹⁵² Directed evolution is commonly applied in protein engineering in order to harness the power of natural selection to evolve proteins with desirable properties, e. g., increased stability towards denaturing conditions, such as temperature extremes or organic solvents.^{153,154}

2.3.1 Concept of Protein Evolution

Protein evolution is based on the concept of linking the genotype to a certain phenotype, so that individual DNA can be directly translated into its protein sequence. This strategy enables the examination of protein properties, selection of the best phenotypes and finally the exploitation of the attached nucleic acids through sequencing. Thus, a typical directed evolution experiment involves three steps:

- (i) *diversification* – the gene encoding the protein of interest is mutated or randomized to create a library of gene variants;
- (ii) *selection* – the library is tested for the presence of mutants possessing the desired property;
- (iii) *amplification* – the identified variants are replicated manifold, enabling sequencing.

Usually the final two steps are repeated in iterative rounds, while amplifying selected variants and imposing selection pressures, so that evolution naturally selects library members with desired biological and physicochemical properties. Valuable techniques like cell expression display and cell surface display,¹⁵⁵ as well as ribosome display,^{156,157} mRNA display¹⁵⁸ and phage display¹⁵⁹ have emerged from this concept (Figure 2.17).

For cell surface display the proteins of interest are anchored to the surface of a living cell via integrated membrane proteins (Figure 2.17a). The gene encoding the fusion protein and the anchor protein is encoded in a plasmid harbored by cells. Ribosome display is based on the idea that ribosome complexes consisting of a nascent chain, ribosome, and mRNA can be generated, thereby establishing a linkage between the protein and the encoding gene (Figure 2.17b). mRNA display (Figure 2.17c) directly links the mRNA and the encoding protein via puromycin. The linkage is achieved by in vitro translation of the mRNA containing puromycin at the 3' end, as puromycin will be attached to the C-terminus of the nascent chain when it enters the ribosome. The most commonly applied system for phage display is that using M13 filamentous phage (Figure 2.17d). With this system, the protein of interest is displayed on the surface of a bacteriophage and the gene encoding the protein is encapsulated inside the phage.

Phage display is typically used for studying protein-protein interactions and screening for specific antibodies.^{152,160,161} It also has been successfully applied in the selection of specific

coiled-coil interaction partners,¹⁶²⁻¹⁶⁴ and was therefore used within the scope of this thesis. In the following section, the implemented bacteriophage strain as well as commonly applied strategies for protein display are outlined in detail.

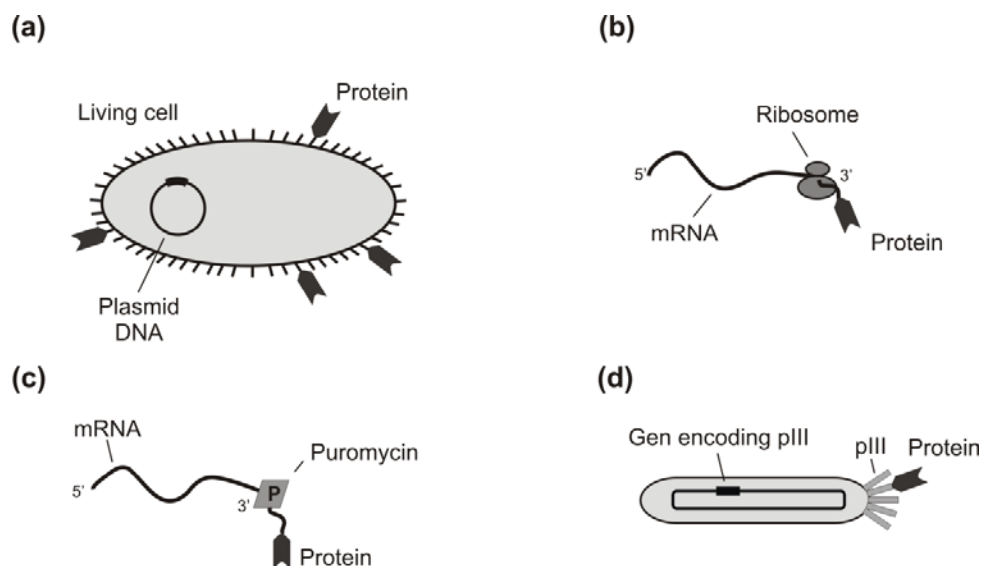


Figure 2.17: examples of strategies for genotype-phenotype linkage; (a) cell-surface display, (b) ribosome display, (c) mRNA display, (d) phage display. Adapted from T. Matsuura et al.¹⁶⁵

2.3.2 Phage M13 - Genome and Structure

While various types of phage, including lamda phage and T7 phage, have been implemented in phage display experiments, the most widely used system is that implementing M13 filamentous bacteriophage (phage).^{159,160} M13 is a Ff-class phage, i.e. a virus infecting bacteria carrying an F-plasmid (fertility factor) and recognizing the F-pilus of *E.coli* as the host cell receptor. The wild type particle is 65 Å in diameter and 9300 Å in length.¹⁶⁶ Its genome is a circular single stranded DNA (~6400 nucleotides, Figure 2.18a) which is encapsulated in a flexible protein cylinder.¹⁵² It encodes for 11 proteins that are grouped according to their function in the life cycle of the phage. One group (gene II, V and X) encodes proteins that are necessary for genome replication. A second group of genes codes for coat proteins (pVII, pIX, pVIII, pIII, and pVI), and a third for proteins that are required for particle assembly (pI, pXI, and pIV). The body of the cylindrical capsid consists of ca. 2700 copies of the predominantly α -helical major coat protein pVIII (Figure 2.18b). These 50 amino acid long proteins are arranged in an overlapping shingle-type array and tightly wrapped around the genome, which accounts for the high resistance of the virus capsule towards proteolysis.¹⁶⁷ 10-13 C-terminal residues of pVIII form the inside wall of the capsid. This region contains four positively charged lysine residues which interact with the negatively charged sugar phosphate backbone of the genome.¹⁵² Each end of the virion is capped with two different minor coat proteins. One end contains the hydrophobic 33 residue pVII and 32 residue pIX. The other end of the particle possesses the 112 residue pVI and the 406 residue pIII. The latter is made up of three domains (designated N1, N2 and CT) that are separated

by glycine rich regions (Figure 2.18c). N1 is essential for infection, the transfer of viral DNA into the cytoplasm as well as for the assembly of coat proteins. N2 is responsible for binding to the F-pilus. The C-terminal CT domain is crucial for formation of stable phage particles. X-ray crystallography showed that, when complexed with its bacterial co-receptor TolA,¹⁶⁸ the amino terminus of N1 extends into solution. This enables the presentation of N-terminal peptide or protein fusions without interfering with pIII function.

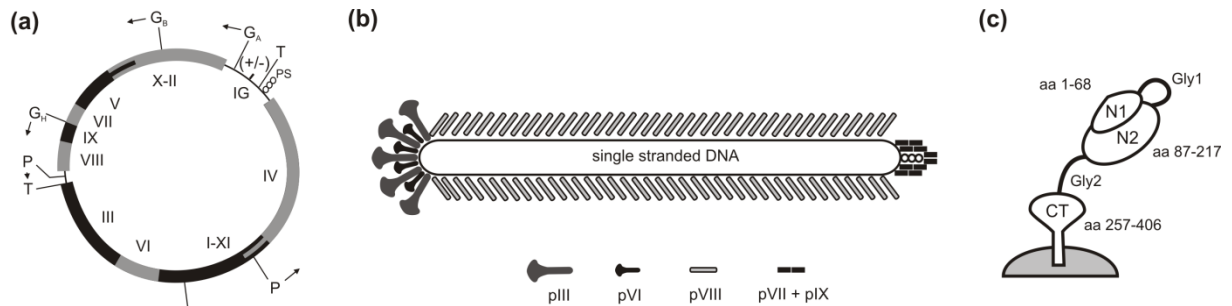


Figure 2.18: Schematic representations of (a) genome of M13 bacteriophage, (b) virion structure and (c) coat protein pIII. The diagram of the genome (a) shows the relative positions of the genes and the important terminators and promoters. IG refers to the intergenic region; T, the two strong terminators; t, the weak terminator in pl; GA, GB, and GH, the promoters of the frequently transcribed region; P, the promoters for the infrequently transcribed regions; PS, the packaging signal; and (+/-), the relative position of the origins of replication for the viral (+) and complementary (-) DNA strands. Adapted from Barbas et al.¹⁵²

2.3.3 Life cycle of Bacteriophage M13

Infection of the host cell is initiated by recognition of the bacterial F-pilus via the N2-domain of pIII, followed by structural rearrangements which induce coreceptor binding of the pIII N1-domain to the TolA membrane protein (Figure 2.19).^{161,166} Subsequently, the coat proteins are disassembled and embedded into the bacterial membrane through fusion. The viral (+)-strand DNA enters the cytoplasm. In the cytoplasm, the complementary (-)-strand is synthesized, both strands covalently linked, and supercoiled double-stranded DNA (dsDNA) generated by bacterial enzymes. The (-)-strand serves as the template for mRNA and thus translation into phage proteins. Post synthesis, the major and minor coat proteins (pIII, pVI, pVII, pVIII and pIX) are inserted into the cytoplasmic or outer membrane. Protein pV, as well as pII and pX, which control the replication of viral DNA, remain in the cytoplasm. Once a certain threshold pV-molecule concentration is reached, newly replicated ssDNA is coated by pV, which prevents conversion into dsDNA.^{169,170} The packaging signal (PS) of the genome remains uncoated and forms an imperfect hairpin that is crucial for the assembly of progeny virions.¹⁷¹ Particle assembly is initiated by five copies of each, pVII and pIX, which interact with this hairpin as well as with pVIII molecules to generate one end of the capsid. Afterwards the particle is elongated by gradual substitution of pV molecules with pVIII. Concurrently it is extruded through the host cell's membrane.¹⁵² During that process the positively charged C-terminus of pVIII molecules interacts with the phosphate backbone of viral DNA. Assembly is terminated by the insertion of pVI and pIII molecules as soon as the

end of the genome is reached. M13 is a nonlytic phage and thus infected host cells continue to grow and to divide indefinitely, albeit at half the rate of uninfected bacteria.

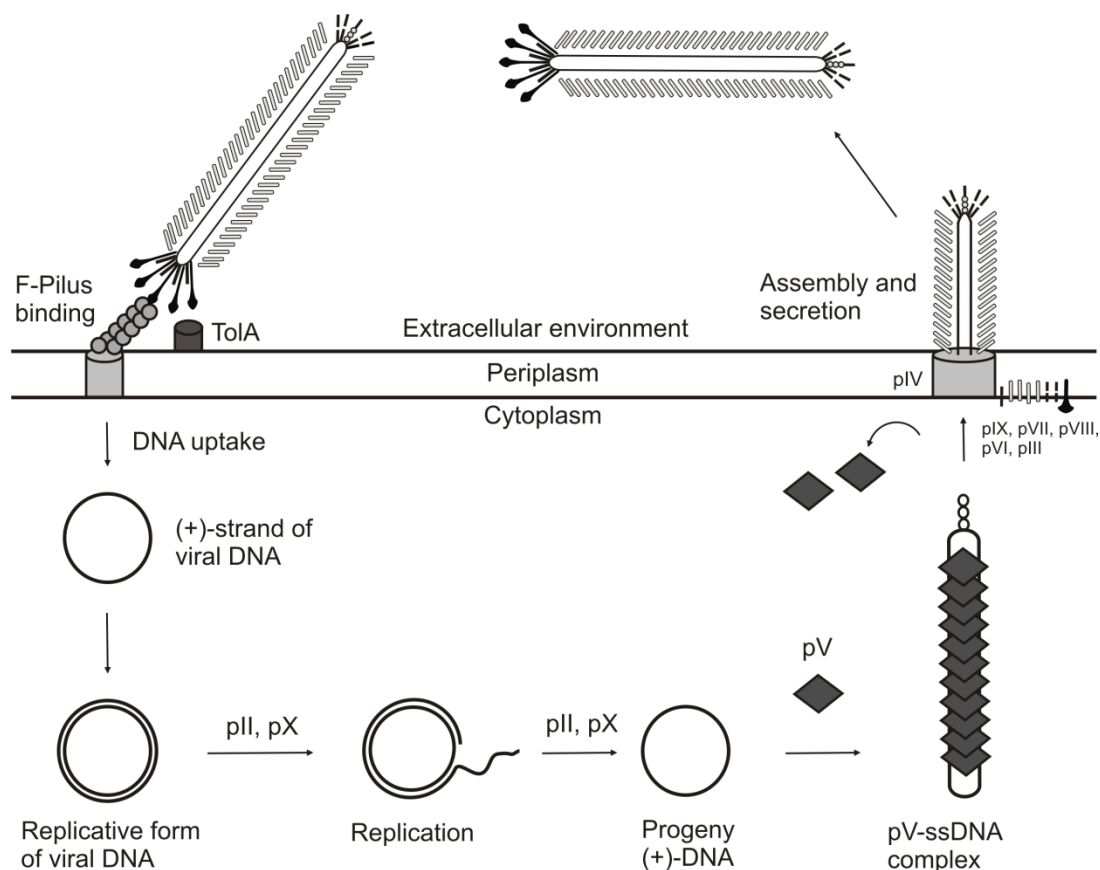


Figure 2.19: Schematic presentation of M13 life cycle. Adapted from Kehoe and Kay.¹⁶¹

2.3.4 Types of Peptide and Protein Display on Surface of M13

While peptide and protein fusions to the amino termini of pVII and pIX, as well as to the carboxy termini of pVI have been reported,¹⁷²⁻¹⁷⁴ common approaches for display are to fuse the foreign sequence to the amino terminus of pIII or pVIII.^{152,160,161} The number of presented peptides is determined by the protein that it is fused to, and on the applied strategy. Three different systems for presenting proteins on M13 surface are currently applied for different purposes (Figure 2.20).^{152,160,161}

The *phage system* is based on the phage genome and allows multivalent display of the recombinant protein. The corresponding DNA is genetically linked to the gene encoding one of the coat proteins. Thus, all copies of this protein are expressed as fusion products. In case of pVIII fusion, each of the ca. 2700 copies of the coat protein pVIII displays the desired peptide. Due to this high copy number pVIII-fusions are used to select ligands with low binding affinity. However, the stability of the viral capsule can be compromised through alterations of the major coat proteins. Although pIII protein fusions allow for display of large proteins due to less steric hindrance on the phage surface, the modification of all five copies of pIII results in reduced infectivity.

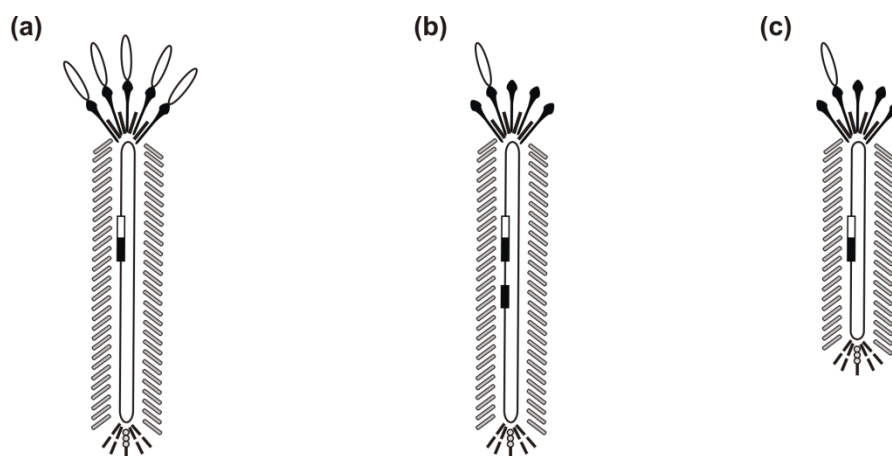


Figure 2.20: Types of phage display. (a) **phage system.** The phage genome contains only one copy of *pIII* (black), and the displayed sequence (white) is fused to it so that all copies of the coat protein carry the displayed sequence. In the **hybrid system** (b) as well as in the **phagemid system** (c) generally only one copy of the fusion protein is presented. Adapted from Kay et al.¹⁶¹

Problems arising from global modification can be circumvented by the introduction of *hybrid systems*, where just some or even only one copy of the fusion protein is incorporated into the viral coat. In this system an additional copy of the coat proteins gene is cloned into the genome and fused to the foreign DNA. If the gene encoding the fusion protein is equipped with a less potent promoter than the *wild-type* gene, the latter is transcribed and translated more frequently.^{175,176} Therefore, resulting phages harbour only one or low numbers of chimeric coat proteins, while the majority remains unmodified.

An alternative route to enable monovalent display is given with *phagemid systems*. In contrast to hybrid systems, two copies of the gene to be modified are located on separate genomes. The wild-type gene is located in the genome of a packaging-defective phage, commonly denoted helper phage, and the recombinant gene is placed on a phagemid. A phagemid is a plasmid that carries the packaging signal, an antibiotic resistance, as well as a viral and a bacterial origin of replication. Therefore, it can be amplified in bacteria and isolated similar to a regular plasmid. The library is cloned into the phagemid such that it will be expressed as a phage protein fusion. All *wild-type* proteins for phage production are encoded in the helper phage genome. Infection of phagemid carrying bacteria with helper phage initiates synthesis of all wild type proteins as well as the fusion protein. Due to the defect packaging signal on the helper phage genome, predominantly single-stranded phagemid-DNA is assembled with viral proteins. Released progeny display the fusion protein library.

3 Aim

To enable the specific application of nonnatural amino acids in peptide and protein engineering, a thorough characterization of their interaction profiles in native protein environments is required. The aim of the current study was to apply phage display to identify a suitable helical environment for peptide sequences containing either fluoroalkyl substituted amino acids or an alternating set of β - and γ -amino acids. Two previously established α -helical coiled-coil oligomers that readily accommodate such substitutions served as model systems. The predefined secondary and tertiary structure of coiled coils enables randomization of the direct interaction partners of an amino acid at a specified position and therefore the identification of favorable molecular recognition environments comprising canonical amino acids. The insights provided by these investigations served as the basis for a novel approach towards the rational design of a protease resistant peptide-based inhibitor of HIV's gp41 envelope protein subunit.

4 Concept and Previous Studies

4.1 Applied Screening System

A phage display screening system has previously been developed in the group of Prof. Kokschi, to allow for selection of preferred binding partners of nonnatural amino acids mediated by coiled-coil formation.^{177,178} Based on findings gained with this system several phage display libraries were generated in line with the present study to enable elucidation of favored protein environments for fluorinated as well as backbone extended amino acids. The concept of the applied phage-display assay is briefly outlined in the following.

4.1.1 Library Construction

The display of peptide libraries was accomplished via a pComb3 (*pComb3HSS*)¹⁷⁹ phagemid cloning vector in combination with bacteriophage M13 helper phages (see chapter 2.3.4). pComb3 harbors a gene conferring resistance to carbenicillin and features the nucleotide sequence that encodes for amino acids 230-406 (CT domain and glycine rich linker to N1 domain) of M13's pIII coat protein (*cf.* Figure 2.18).¹⁵² N-terminal peptide fusion was realized through oligonucleotide ligation onto the 5'-end of pIII via *Sfi*I restriction sites (Figure 4.1). The expression of this fusion protein is controlled by a weak promoter. This guarantees that potential toxic effects of the protein construct are minimized because mainly fully infectious phages, predominantly carrying only one pIII-fusion protein and four native pIII proteins that have been expressed from the helper phage genome, are generated.

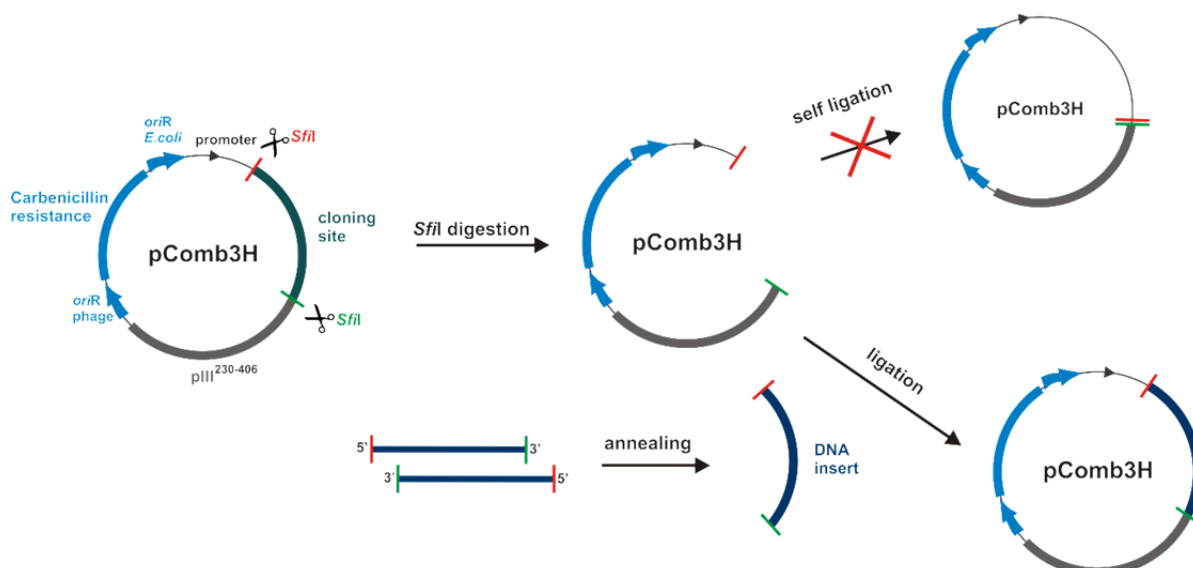


Figure 4.1: Schematic representation of applied cloning strategy

Chemically synthesized oligonucleotides that possess *Sfi*I sticky ends for ligation into the phagemid vector and encode for desired peptides with specifically randomized codons were purchased. Codon randomization was utilized during oligonucleotide synthesis via the NNK strategy, where N-positions can be occupied by all of the four nucleotides, whereas K

positions can only be filled with either guanine (G) or thymine (T). This strategy reduces the number of possible codons from 64 to 32, still encoding every canonical amino acid, while eliminating two of three STOP codons which would otherwise abort protein biosynthesis and thus lead to truncated nonfunctional protein products. The remaining amber STOP codon (TAG) is partially suppressed and translated as glutamine (Q) with an efficiency of 10 - 20 % in the suppressor *E.coli* ER2738 strain applied for library generation.¹⁸⁰ Randomized codons statistically encode for any of the 20 canonical amino acids. The DNA uptake limit of currently available *E. coli* ER2738 is 10^{10} . Thus, to guarantee a fully diverse peptide library, the number of randomized codons should not exceed six positions simultaneously.

In order to allow ligation of foreign DNA to the N-terminus of pIII, the phagemid is digested with the rare-cutting restriction enzyme *SfiI*.¹⁵² This enzyme recognizes the 8-bp sequence GGCCNNNN[^]NGGCC and cuts within the degenerate region of its interrupted palindromic recognition site. Use of unique 5' (GGCCAGG[^]CGGCC) and 3' (GGCCAGGC[^]CGGCC) *SfiI* sites allows for incorporation of the DNA library in the correct orientation and prevents self-ligation of the vector at the same time. Ligated phagemids were transformed into competent *E. coli* ER2738 bacteria by electroporation and transformed bacteria were selected from carbenicillin containing media.

4.1.2 Biopanning

The employed experimental procedure of affinity selection (denoted "biopanning") of phage displayed peptide sequences that bind to target peptides containing nonnatural amino acids is illustrated in Figure 4.2.

The target sequences are generated with a biotin tag by means of solid phase peptide synthesis applying standard Fmoc-chemistry.¹⁸¹ These target peptides (yellow) are immobilized on magnetic particles (green) that are covalently linked to streptavidin (Dyna-Beads®). With a dissociation constant (K_d) in the order of $\approx 10^{-14}$ mol/L, the binding of biotin to streptavidin is one of the strongest noncovalent interactions known in nature.¹⁸² Immobilized peptides are then incubated with the phage displayed peptide library (Figure 4.2a). Phages carrying a peptide, whose amino acid sequence is able to undergo coiled-coil formation with the target peptides bind to the magnetic particle via these noncovalent interactions. Subsequently, the particles are collected with the help of an external magnet and phages with no or low affinity fusion peptides are washed off (Figure 4.2b). By increasing the concentration of detergent in the washing solution (i. e. increasing stringency) with every new panning round, the selection pressure is successively increased and only very strong binders accumulated. Bound phages are eluted from the immobilized particles by proteolytic hydrolysis (Figure 4.2c). Here, the incubation time with the protease is short enough to only degrade the fusion peptide, leaving the phage intact and infectious. The obtained phage suspension is used to infect *E. coli* bacteria. Since infected

bacteria harbor phagemid DNA it can be selected from carbenicillin containing agar plates (Figure 4.2d). Particles without target peptide that are treated equivalently can serve as a negative control. Moreover, the phagemid DNA isolated from single colonies of infected bacteria can be used for DNA sequencing with suitable primers to reveal the peptide sequence of the phage displayed fusion protein which interacted with the target. Amplification of phagemid containing bacteria and subsequent infection with helper phages yields progeny phages for further panning rounds (Figure 4.2e).¹⁵²

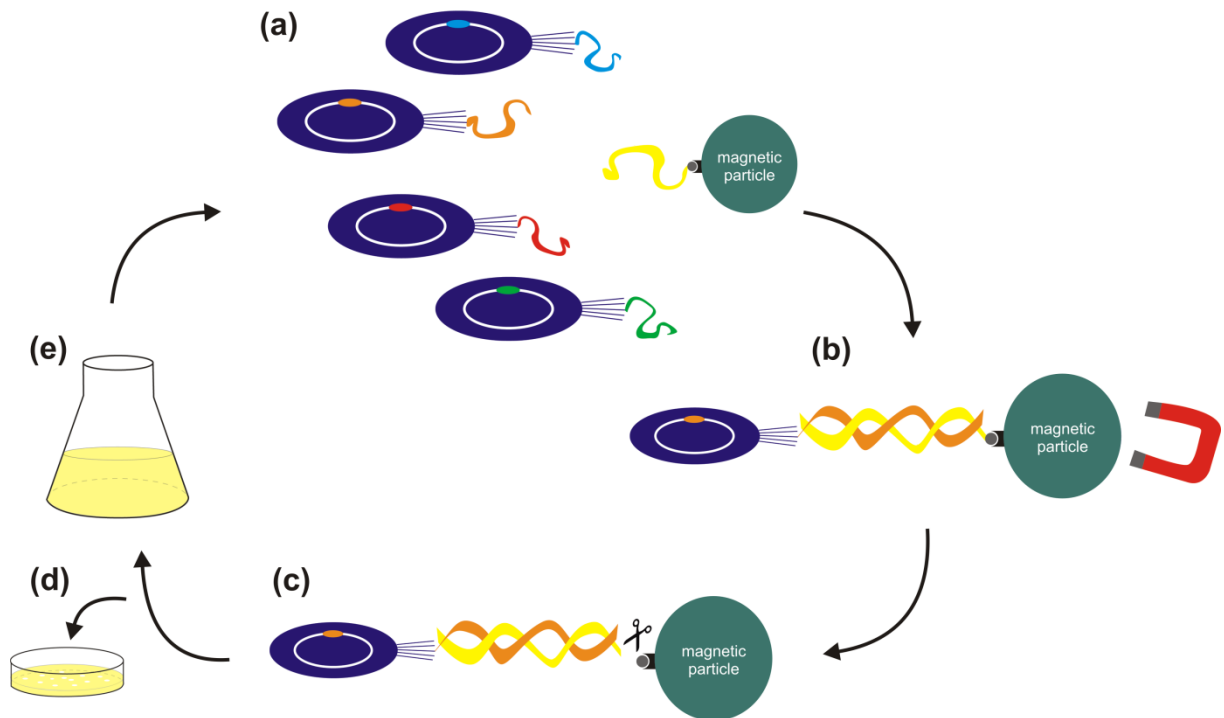


Figure 4.2: Selection of binding partners for a synthetic peptide out of a phage library via biopanning

4.2 Previous Phage-Display Experiments with VPE/VPK

In order to validate VPE/VPK as an appropriate system for coiled-coil formation mediated determination of preferred binding partners of a given amino acid via phage display, Vagt *et al.* initially randomized four positions of VPE that are interacting with a central hydrophobic residue of VPK in a phage displayed library.¹⁶² First, biopanning with this library was carried out against a VPK that contains glutamic acid at the respective substitution site according to above outlined procedures (see chapter 4.1). This experiment resulted in the selection of a lysine residue in a randomized position that is usually designated for hydrophobic amino acids. Structural investigations of the selected coiled-coil dimer by CD spectroscopy and MD simulations suggested that a buried salt bridge between the ammonium function of the lysine side chain and the carboxylate function of the glutamic acid side chain enables specific coiled-coil heterodimerization of the two peptides without disruption of the overall structure. This selection outcome successfully demonstrated the applicability of the phage display based screening system to select specific interaction partners of uncommon amino acids within the hydrophobic core of coiled coils.

Encouraged by these findings the Koksch group subsequently used the VPE/VPK screening system to investigate the interaction profiles of fluorinated amino acids within the hydrophobic environment of the coiled-coil interaction domain. Therefore, DfeGly, TfeGly or DfpGly was incorporated at position a_{16} of VPK.¹⁸³ Direct VPE interaction partners of this position, namely d'_{12} , g'_{15} , a'_{16} and d'_{19} , were randomized in a phage displayed library (Figure 4.3).

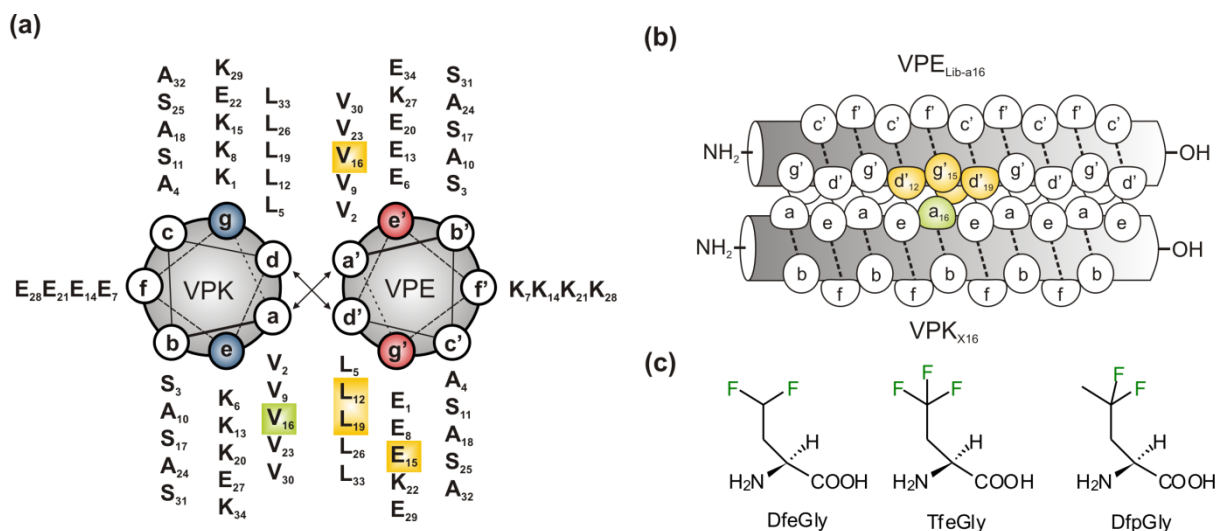


Figure 4.3: (a) helical wheel depiction of VPE/VPK and (b) side view of the heterodimer. The substitution side is shown in green and the four randomized positions in yellow. (c) fluorinated amino acids under investigation.

The amino acids selected for the variable positions followed in principle the pattern Leu₁₂Leu₁₅Ile₁₆Tyr₁₉ (VPE-L₁₅I₁₆Y₁₉) or Leu₁₂Tyr₁₅Ile₁₆Leu₁₉ (VPE-Y₁₅I₁₆) (Table 4.1). Leucine already occurred in the d -positions of VPE wild type and its selection in this position is in agreement with the general preference of d -positions for this amino acid in parallel coiled coil

dimers. The selection of isoleucine in position a'_{16} can also be explained by the fact that hydrophobic, β -branched amino acids are the most stabilizing amino acids in a -positions of parallel coiled-coil dimers. However, the selection of tyrosine in g'_{15} and d'_{19} was somewhat unexpected and was assumed to be due to possible cation- π -interactions between the aromatic side chain of tyrosine and the ammonium function of a lysine located in the opposite g -positions of VPK.

Table 4.1: Amino acids selected in four randomized positions of VPE after 5 rounds of panning against the different VPK- X_{16} analogues. Amino acids in these positions of the original VPE sequence are given in brackets.

| VPK variant | VPE- d'_{12} (Leu) | VPE- g'_{15} (Glu) | VPE- a'_{16} (Val) | VPE- d'_{19} (Leu) |
|--------------------------|-------------------------|-------------------------|-------------------------|-------------------------|
| VPK | Leu | Tyr/Leu | Ile | Tyr/Leu |
| VPK-DfeGly ₁₆ | Leu | Hydrophobic | Ile | Tyr/Leu |
| VPK-DfpGly ₁₆ | Leu | Tyr/Leu | Ile | Tyr/Leu |
| VPK-TfeGly ₁₆ | Leu/Phe | Tyr/Leu | Ile | Tyr/Leu |

To further characterize the VPE variants gained as binding partners of fluorinated VPK variants, these VPE peptides (VPE-L₁₅I₁₆Y₁₉, VPE-Y₁₅I₁₆, VPE-L₁₅I₁₆) were chemically synthesized and the thermal stability of all coiled-coil couples, which arose from the combination with the different VPK-variants, was determined. In comparison to the original VPK- X_{16} /VPE couple, a distinct increase of thermal stability could be observed for the VPE-peptides selected by phage display. The most stable combination of each VPK-variant was formed with VPE-L₁₅I₁₆. The extension of the hydrophobic interface, caused by the predominant selection of hydrophobic amino acids in position g'_{15} probably compensates the destabilizing effects caused by the fluorinated amino acids.²⁵

It is conceivable that this expansion stabilizes the coiled-coil structure and reduces the sensitivity of the system to detect changes in position a_{16} . To test this hypothesis and to check whether any differences in the selection outcome can be observed when the charged position remains unmodified, a second library for the selection against VPK variants containing fluorinated amino acids in a_{16} was constructed in line with the present thesis.¹⁸³

5 Results and Conclusions

5.1 Accommodating Fluorinated Amino Acids in Parallel Coiled-Coil Dimers

5.1.1 Substitutions in a Central α -Position

Initially, as a follow up to previous results gained in the Koksch group from biopanning against VPK's position a_{16} with a phage displayed VPE library that included a core flanking position (see chapter 4.2), a VPE library was constructed in which only d'_{12} , a'_{16} and d'_{19} were randomized (Figure 4.3).^{*} Moreover, in the course of these studies, the Koksch group gained access to the monofluorinated analogue of aminobutyric acid from Prof. Haufe (University of Münster), which was therefore included in the here presented studies.

The selected VPE-variants, resulting from screening against VPK and its above mentioned fluorinated analogues showed little variation, and mainly followed the pattern $\text{Leu}_{12}(\text{Ile/Leu})_{16}\text{Leu}_{19}$ (Figure 5.1). In contrast to the extended library with 4 randomized positions by Vagt *et al.*, here, exclusively leucine was observed in positions d'_{12} and d'_{19} , while formerly also aromatic amino acids had been selected in these sites (see chapter 4.2). These results suggest that the selection of tyrosine in hydrophobic core positions of the initial VPE-library was a consequence of extending the hydrophobic core by position g'_{15} , and was not specific for the investigated fluorinated amino acids.

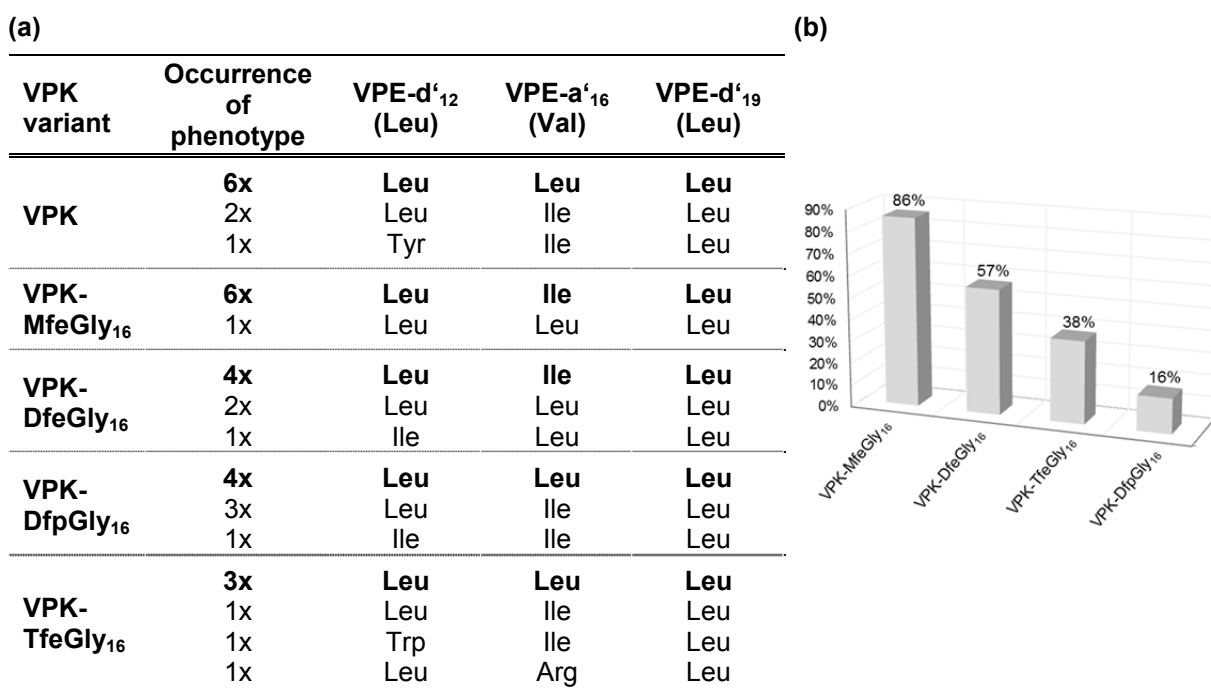


Figure 5.1: (a) Table of amino acids selected in randomized positions of VPE after 5 panning rounds against the different VPK- X_{16} analogues. Amino acids in these positions of the original VPE sequence are given in brackets. (b) Frequency of β -branched amino acids in VPE- a'_{23} post panning against the different VPK- X_{19} analogues.

^{*} Results of this section are partially summarized in: T. Vagt, E. Nyakatura, M. Salwiczek, C. Jäckel, B. Koksch, *Towards identifying preferred interaction partners of fluorinated amino acids within the hydrophobic environment of a dimeric coiled coil peptide*, *Org. Biomol. Chem.*, 2010, 8, 1382-1386.

As mentioned above, leucine is predominantly found in d -positions of naturally occurring coiled-coil assemblies,¹⁸⁴ and, thus, its emergence as the most conserved amino acid in position d'_{12} and d'_{19} is not surprising. Studies by Hodges *et al.* showed that β -branched amino acids (Val and Ile) in a -positions of parallel dimeric coiled coils exhibit the strongest contribution to stability.¹⁸⁵ The occurrence of β -branched amino acids in these packing arrangements has been attributed to the γ -methyl group, which projects from the β -carbon atom back into the center of the hydrophobic core.⁴³ Wild type VPE bears a valine in this position. The predominant selection of Leu and Ile in randomized a -positions might be attributed to their higher hydrophobicity in comparison to Val. Interestingly, upon increasing side-chain volume of the residues in VPK's substitution site a_{16} , the frequency of β -branched isoleucine in a'_{16} decreases (Figure 5.1b). It is conceivable that this is due to unfavourable packing of isoleucine, as the steric bulk increases with the degree of fluorination. DfpGly is exempt from this trend, which is presumably due to more pronounced constitutional differences, compared to the other amino acids.²⁵ Regarding hydrophobicity and size, solely TfeGly is comparable to Val (cf. Figure 2.6). Phage display selection led to the same interaction profile for the latter two amino acids (leucine in all three randomized positions).

Taken together, the interaction patterns obtained after biopanning of either of the two libraries against any of the fluorinated amino acids in a central a -position of the target peptide were found to be similar to that of the native hydrophobic residue valine. Earlier studies of the Kokschi group have shown that significant differences in hydrophobicity and side-chain volumes of the here investigated fluorinated building blocks can have an impact on the folding stability of coiled coils (see chapter 2.2.1).^{23,162} Yet, phage display mediated screening showed, when incorporated in position a_{16} , where the through fluorine polarized β -methylene groups point away from the hydrophobic core, all four investigated fluoroalkyl-substituted amino acids prefer a similar amino acid pattern. Despite their differences in hydrophobicity and size, they all prefer to interact with the aliphatic amino acids leucine and isoleucine. While both of these amino acids possess the same hydrophobicities and side-chain volumes, the position of the side-chain branch (β -methyl group in Ile and γ -methyl group in Leu) seems to control for optimal side-chain packing of the different (*S*)-2-aminobutyric acid analogues that possess increasing fluorine stoichiometry.

5.1.2 Substitutions in a Central *d*-Position

Since the packing patterns at the hydrophobic core of coiled-coil assemblies are imposed by the overall structure of the system, the amino acids in parallel assemblies at either *a*- or *d*-positions are embedded within a slightly different microenvironment (see chapter 2.1.1). If fluoroalkyl-substituted amino acids are incorporated at VPK's position a_{16} (*parallel packing*), the stability of the resulting VPK/VPE heterodimeric assembly is mainly dependent on their side-chain volume and hydrophobicity.^{25,97} Due to the different orientation of the local dipole, at position d_{19} the destabilizing impact of fluorine induced polarity prevails. Here, the polarized β -carbons project directly into the hydrophobic core (*perpendicular packing*). With regard to this, a second VPE library was engineered in line with the current study, in order to enable the identification of a suitable microenvironment for fluoromodified amino acids in VPK's position d_{19} .

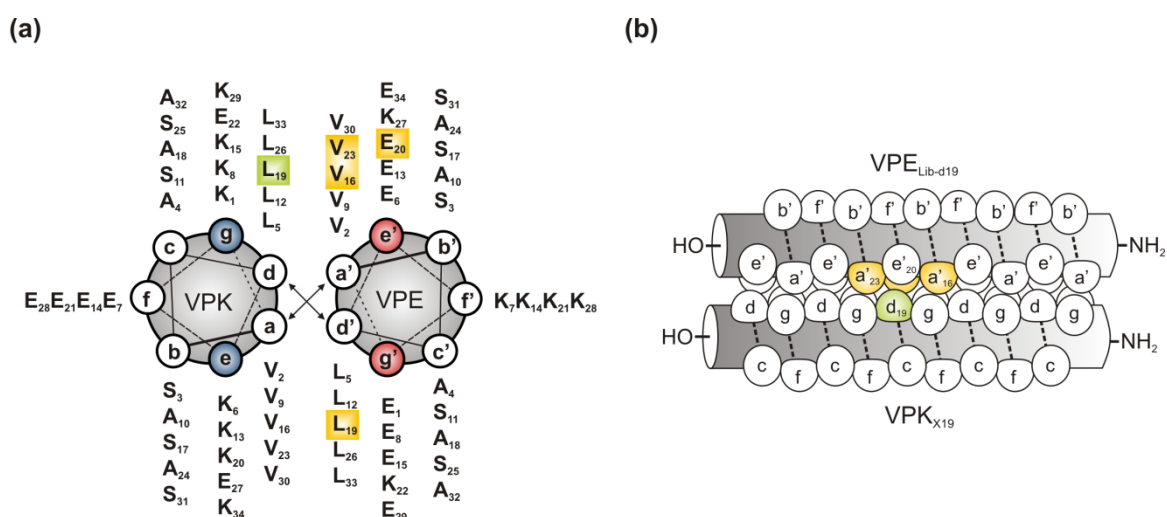


Figure 5.2: (a) helical wheel and (b) ribbon diagram of VPE/VPK. Substitution position is highlighted in green and randomized positions are highlighted in yellow.

Phage display library design was based on the randomization of the three central hydrophobic positions of VPE that are directly interacting with position d_{19} , namely a'_{16} , d'_{19} and a'_{23} (Figure 5.2). Because VPE is fused to the amino terminus of pIII, the different fluorinated VPK-variants were immobilized on streptavidin coated magnetic particles via an N-terminal biotin tag, in order to facilitate parallel helix alignment of the heteromeric assembly. A glycine-serine-glycine (GSG) sequence served as a spacer between biotin and VPK. Coiled-coil pairing selectivity was used to select best binding partners out of the pool of phage displayed VPE derivatives in five panning rounds according to procedures outlined above (see chapter 4.1).¹⁸⁶

Biopanning against VPK Wild Type

At first, wild type VPK, harboring leucine at the substitution site, was utilized as a target for selection with the newly generated library. Biopanning and subsequent sequencing of amplified phagemid vectors revealed a single consensus sequence (Table 5.1). Leucine is the most favored canonical amino acid regarding stability, when occupying *d*-positions of parallel α -helical coiled coils,^{79,184} and was used for the *d*-positions in the design of the original VPE.²⁵ Thus, its selection in *d'*₁₉ is conceivable. The selection of this amino acid in position *a*₁₆ might be attributed to its higher hydrophobicity in comparison to valine, which was used for *a*-positions during *de novo* design of VPK. Surprisingly, aromatic tyrosine was selected in VPE's position *a'*₂₃, even though only hydrophobic core positions have been randomized, and, thus, the extension of the hydrophobic core as observed with the library possessing a randomized core flanking position (see Table 4.1)¹⁸³ is impossible.

Table 5.1: Amino acids selected in randomized positions of VPE after 5 rounds of panning against VPK. Amino acids in these positions of the original VPE sequence are given in brackets.

| | Occurrence of phenotype | VPE- <i>a'</i> ₁₆ (Val) | VPE- <i>d'</i> ₁₉ (Leu) | VPE- <i>a'</i> ₂₃ (Val) |
|-----|-------------------------|---------------------------------------|---------------------------------------|---------------------------------------|
| VPK | 9x | Leu | Leu | Tyr |

To further investigate this selection outcome, the consensus sequence (VPE-L₁₆Y₂₃) was chemically synthesized and its interaction with VPK investigated in solution. CD spectra clearly indicate two distinct minima at 208 and 222 nm verifying the formation of helical assemblies (Figure 5.3a).

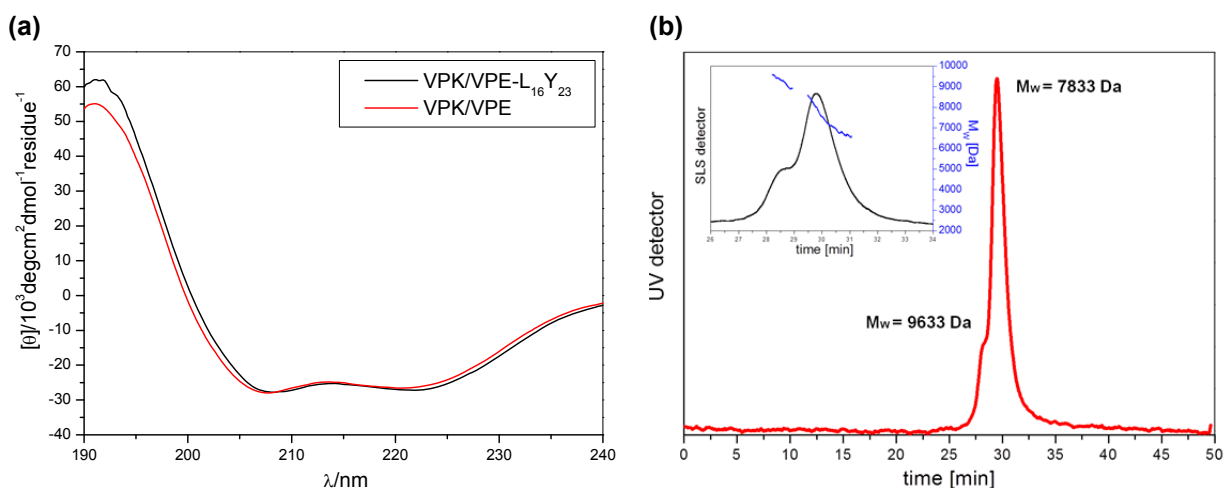


Figure 5.3: (a) CD spectra of VPK/VPE_{variant} heteromers depicted as the mean of 3 experiments. Peptide conc.: 20 μ M; 100 mM Phosphate buffer. (b) SEC/SLS chromatograms of VPK/VPE_{L₁₆L₂₃} depicted as the mean of 3 experiments. Peptide conc.: 35 μ M; eluent: 10mM PBS; flow rate: 0.3 mL/min; pH 7.4. Measurements were taken from 1:1 mixtures and repeated two times to confirm reproducibility and give standard deviations.

Applying size exclusion chromatography (SEC) in combination with static light scattering (SLS) showed that the dimeric oligomerization state of the VPE/VPK parent system is primarily retained (Figure 5.3b). However, a relatively small proportion of trimeric species is indicated by a shoulder of the size exclusion peak. The CD spectra of 1:1 mixtures of VPK and the selected VPE analogue as well as of VPK/VPE are nearly superimposable, indicating a similar folding behaviour. Seeking to elucidate the thermal stability of the selected assembly the CD signal at 222 nm which is distinctive for α -helices was monitored with rising temperature. Surprisingly, the obtained terminal stability ($T_M = 67^\circ\text{C}$, Figure 9.2; p. 113) is lower than the thermal stability of the parent VPE/VPK ($T_M = 72^\circ\text{C}$).^{25,183}

In order to understand this exceptional selectivity for VPE-L₁₆Y₂₃ (9 out of 9 clones), molecular dynamic studies with the heterodimeric VPK/VPE-L₁₆Y₂₃ in parallel strand orientation were carried out (Figure 5.4).⁹⁹

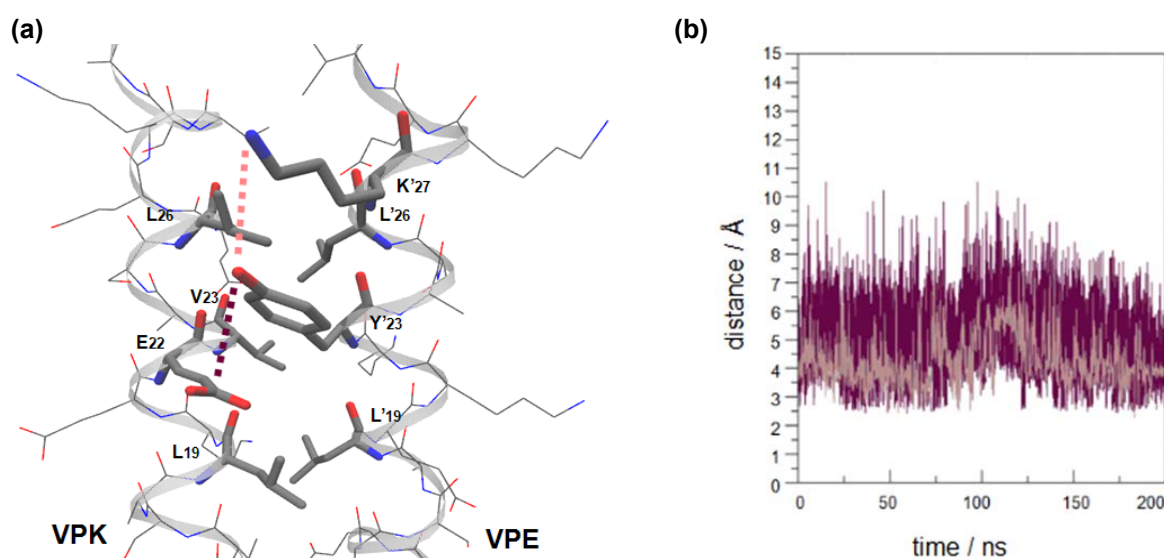


Figure 5.4:⁹⁹ (a) Snapshot of H-bond triade from MD simulations. (b) Distance between the hydroxyl group of Y'₂₃ and the side chain carboxyl group of E₂₂ (purple) and the secondary amine of K'₂₇ (pink).

Simulations revealed, that being situated at an *a*-position where the C_α-C_β bond vector points out of the hydrophobic core, the large hydrophobic surface of tyrosine's aromatic ring provides for ideal exclusion of water at the nonpolar helical interface. In general, the formation of hydrophobic interphases between two nonpolar side chains leads to a loss of side-chain conformational entropy.^{187,188} In regard to this, tyrosine might be particularly favorable since its side-chain rotamers are restricted. Moreover, simulations clearly indicate that tyrosine's hydroxyl group can serve as a hydrogen bond donor and acceptor at the same time (Figure 5.4). Its hydrogen can be involved in interhelical hydrogen bonding with the side chain carboxyl group of glutamic acid at VPK's position *g*₂₂, while the oxygen can form an intrahelical hydrogen bond with lysine's secondary amine which is located at VPE's position *e'*₂₇. During simulation tyrosine's side chain is always situated in between VPK-E₂₂ and VPE-K₂₇. Moreover, an H-bond distance plot indicates a certain flexibility of tyrosine's hydroxyl

group, since the individual distances of this group to the acidic functionality of E₂₂ and the side chain amine of K₂₇ vary between 2.5 Å and 8 Å throughout the entire course of the simulation. Taken together these results indicate that the tyrosine constitutes the center of an H-bond triad.

In order to further address the fundamental question of selectivity vs. stability raised by these results, theoretical investigations are currently conducted. Therefore, the energy barriers during the assembly process of VPK/VPE and VPK/VPE-L₁₆Y₂₃ are examined.⁹⁹

Biopanning Against VPK Containing Fluorinated Amino Acids at Position d₁₉

Leucine is both more hydrophobic and larger compared to any of the fluorinated amino acids investigated in the course of the present work (cf. chapter 2.2.1). Therefore, Abu, which constitutes the hydrocarbon analogue of these fluoromodified amino acids, was chosen to serve as a more adequate control in following studies.*

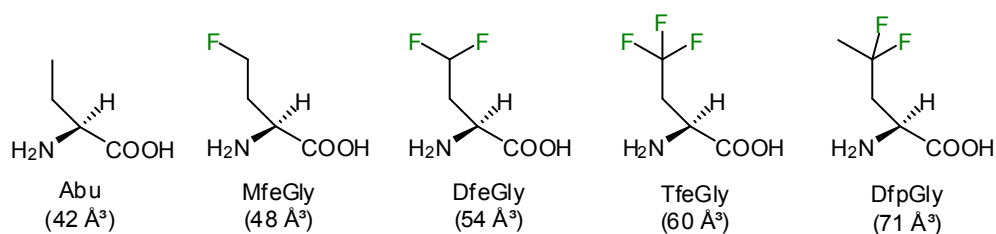


Figure 5.5: Structure and volume of (S)-2 aminobutyric acid (Abu), (S)-4-monofluoroethylglycine (MfeGly), (S)-4,4-difluoroethylglycine (DfeGly), (S)-4,4,4-trifluoroethylglycine (TfeGly), and (S)-4,4-difluoropropylglycine (DfpGly). vdW volumes of the amino acid side chains were calculated according to Zhao et al.¹⁸⁹

Phage display affinity selection with the VPE-phage library against VPK variants containing Abu, MfeGly, DfeGly, TfeGly or DfpGly in position d₁₉, revealed VPE analogues in which all three randomized positions were predominantly occupied by aliphatic amino acids (Figure 5.6a). Leucine emerged as the most conserved residue at position d₁₉. While this amino acid is likewise primarily selected in positions a₁₆' and a₂₃', here β-branched valine and more frequently isoleucine were also selected. However, when selected against VPK-TfeGly₁₉ all three randomized positions bear exclusively leucine. The selection of Leu and Ile in a-positions might be attributed to their higher hydrophobicity in comparison to Val, which was used in the original sequence of VPE for these positions. Albeit to a lesser extent, also the decrease in the frequency of β-branched isoleucine in a randomized a-position (a₂₃') upon increasing side-chain volume of the residues in VPK's substitution site is reproduced (Figure 5.6b). Consistent with the previous study, all randomized d-positions bear exclusively unbranched leucine when selected against TfeGly. Again, DfpGly is exempt from this trend, which might be attributed to more pronounced constitutional differences, compared to the other amino acids, having an impact on coiled-coil pairing specificity (*vide infra*).²⁵ Thus, the

* Results of this section are summarized in: E. K. Nyakatura, O. Reimann, T. Vagt, M. Salwiczek, B. Kokschi, *Accommodating fluorinated amino acids in a helical peptide environment*, *RSC Adv.*, 2013, 3(18), 6319-6322.

selected amino acid pattern exhibits a high degree of similarity to the selection pattern obtained post biopanning against VPK's position a_{16} (cf. chapter 5.1.1; Figure 5.1).

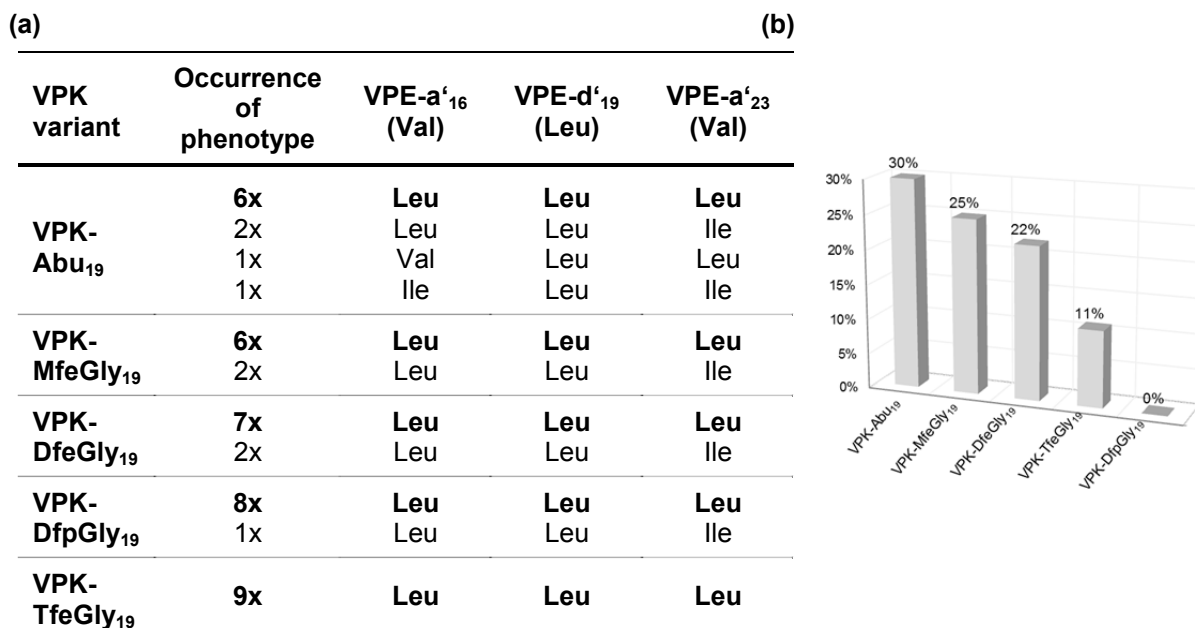


Figure 5.6: (a) table of amino acids selected in randomized positions of VPE after 5 rounds of panning against VPK- X_{19} . Amino acids in these positions of the original VPE sequence are given in brackets. (b) Frequency of β -branched amino acids in VPE- a'_{23} post panning against the different VPK- X_{19} analogues.

Because leucine was so frequently selected, in all three randomized positions, this selection pattern emphasizes that not only hydrophobicity, but also packing of interior hydrophobic residues affects coiled-coil stability.⁴² Leucine seems to provide the best shielding against water for fluorinated Abu analogues substituted in d -positions.

It is this combination of polarity, spatial properties and packing characteristics that determines the oligomerization state. Therefore, the consensus sequence of this phage display screen (VPE- $L_{16}L_{23}$) was chemically synthesized and the heteromeric oligomerization state of all VPK- X_{19} /VPE- $L_{16}L_{23}$ was determined by means of size exclusion chromatography (SEC) in combination with static light scattering (SLS). The results obtained indicate that the dimeric oligomerization state of the VPE/VPK parent system is largely retained (Table 5.2). However, VPK-DfeGly₁₉/VPE- $L_{16}L_{23}$, VPK-TfeGly₁₉/VPE- $L_{16}L_{23}$ as well as VPK-DfpGly₁₉/VPE- $L_{16}L_{23}$ also show a relatively small proportion of monomeric species, while a huge shoulder of the SEC peak of VPK-DfpGly₁₉/VPE- $L_{16}L_{23}$ represents an almost 1:1 mixture of dimers and trimers (see Figure 9.1; p.112).

Table 5.2: Molmasses and helical content of VPK- X_{19} /VPE- $L_{16}L_{23}$ heteromers (see Figure 9.1; p. 112). Standard deviations determined from 3 independent measurements.

| | Heterodimer mass [Da] (theoretical / experimental) | Helical content [%] |
|-------------------------------------------------|-------------------------------------------------------|---------------------|
| VPK-Abu ₁₉ /VPE- $L_{16}L_{23}$: | 7581 / 7471 ± 571 | 83 |
| VPK-MfeGly ₁₉ /VPE- $L_{16}L_{23}$: | 7599 / 7929 ± 279 | 63 |
| VPK-DfeGly ₁₉ /VPE- $L_{16}L_{23}$: | 7617 / 7884 ± 90 | 76 |
| VPK-TfeGly ₁₉ /VPE- $L_{16}L_{23}$: | 7635 / 7790 ± 114 | 62 |
| VPK-DfpGly ₁₉ /VPE- $L_{16}L_{23}$: | 7631 / 7979 ± 67 | 79 |

To further characterize the selection outcome, CD spectra were recorded and the thermal stabilities of all coiled-coil pairs, which arose from the combination of VPE- $L_{16}L_{23}$ with the different fluorinated VPK variants, were determined (Figure 5.7). Indeed, all 1:1 mixtures formed stable coiled coils with a helical content above 60 % and a increased thermal stability, when compared to 1:1 mixtures composed of the original VPE and the fluorinated VPK variants. The highest thermal stabilization was observed for VPK-DfeGly₁₉/VPE- $L_{16}L_{23}$ (+5.8 °C), followed by VPK-MfeGly₁₉/VPE- $L_{16}L_{23}$ (+5.2 °C).

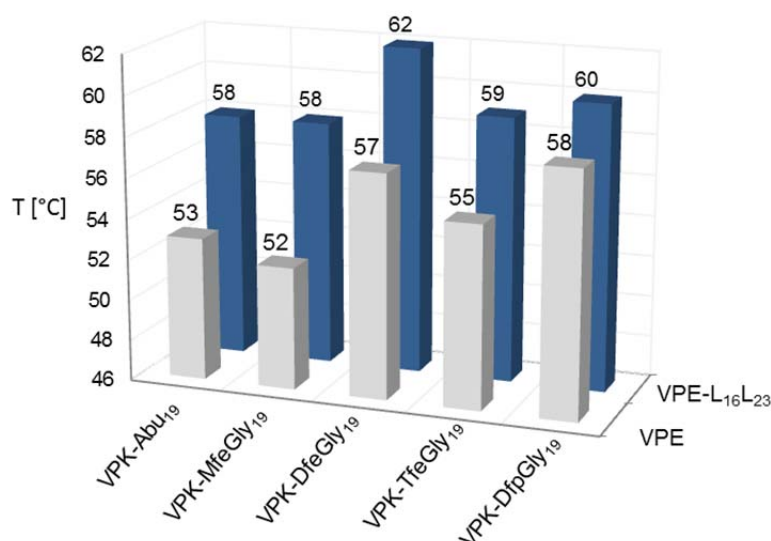


Figure 5.7: bar diagram showing melting points of VPK- X_{19} /VPE- $L_{16}L_{23}$. For comparison, melting points of VPK- X_{19} /VPE_{wild type} are shown in the foreground. Standard deviations from three independent measurements were ± 0.1 °C.

In summary, the interaction profiles of all four fluoroalkylated amino acids were found to be similar to that of Abu – despite their differences in hydrophobicity and size. Thus, these results complement the previous study, which focused on VPK's position a_{16} (cf. chapter 4.2, chapter 5.1.1, and Vagt *et al.*¹⁸³). Independent of the respective microenvironment, incorporating fluorinated aminobutyric acids into position d_{19} or position a_{16} of a parallel heterodimeric coiled coil leads to pairing characteristics that are similar to their hydrocarbon analogues. MfeGly, DfeGly, TfeGly, and DfpGly, in principle, prefer a hydrophobic environment within protein structures. Considering that in position d_{19} the through fluorine substitution polarized β -carbon group points directly into the hydrophobic core, and that

polarity differences of Abu analogues with increasing fluorine content have a substantial impact on the biophysical properties of peptides,²⁵ these findings are somewhat surprising. On the other hand, it is conceivable that the applied system directs the outcome of this experiment, and that in its structural context the alphabet of canonical amino acids might not be sufficiently diverse to report on the comparably subtle changes induced by the incorporation fluorine atoms.

Also, the evaluation of high resolution X-ray structures of other peptides and proteins containing fluorinated hydrophobic amino acids have shown that, although they are somewhat larger, fluorinated residues only minimally perturb the overall structure when introduced in low stoichiometries, probably due to their close shape complementarity to the hydrocarbon side chains they replace.^{17,133,190,191}

If correct packing is achieved, a criterion most likely to be fulfilled when the amino acid identity is retained (e.g., replacing hexafluoroleucine for leucine, etc.), the unique intrinsic physiochemical properties of fluorinated amino acids might be exploited to generate fluorinated peptides and proteins with novel properties that can be applied for clinical and industrial purposes.

5.2 Stabilizing a Coiled Coil that Contains a Set of Alternating β - and γ -Amino Acids

Protein-protein interactions in which at least one partner contributes an α -helix to the protein interface are attractive candidates for the design of foldamer-based inhibitors, since a variety of helices comprising foldameric sequences can be predictably generated. However, despite the growing number of such systems, the precise mimicry of side chain topology and recognition properties of natural α -helical sequences still remains challenging. Efforts of the Kokschi group to design a heterogeneous coiled-coil foldamer have led to the generation of the tetrameric B3 β 2 γ /Acid-pp, which possesses a pentad of alternating β - and γ -amino acids (*cf.* chapter 2.2.2).⁸⁴ While this system tolerates the heterogeneous modification, a loss in thermal stability is observed when compared to its parental system Base-pp/Acid-pp.

In line with the present thesis a phage display screening system was generated, in order to facilitate the empirical search for high-affinity binders to B3 β 2 γ . The peptide library construction was based on the Acid-pp (*wt*) sequence, and included the randomization of the central heptad key positions (a'_{15} , d'_{18} , e'_{19} , g'_{21}), which are directly interacting with the $\beta\gamma$ -segment of B3 β 2 γ .

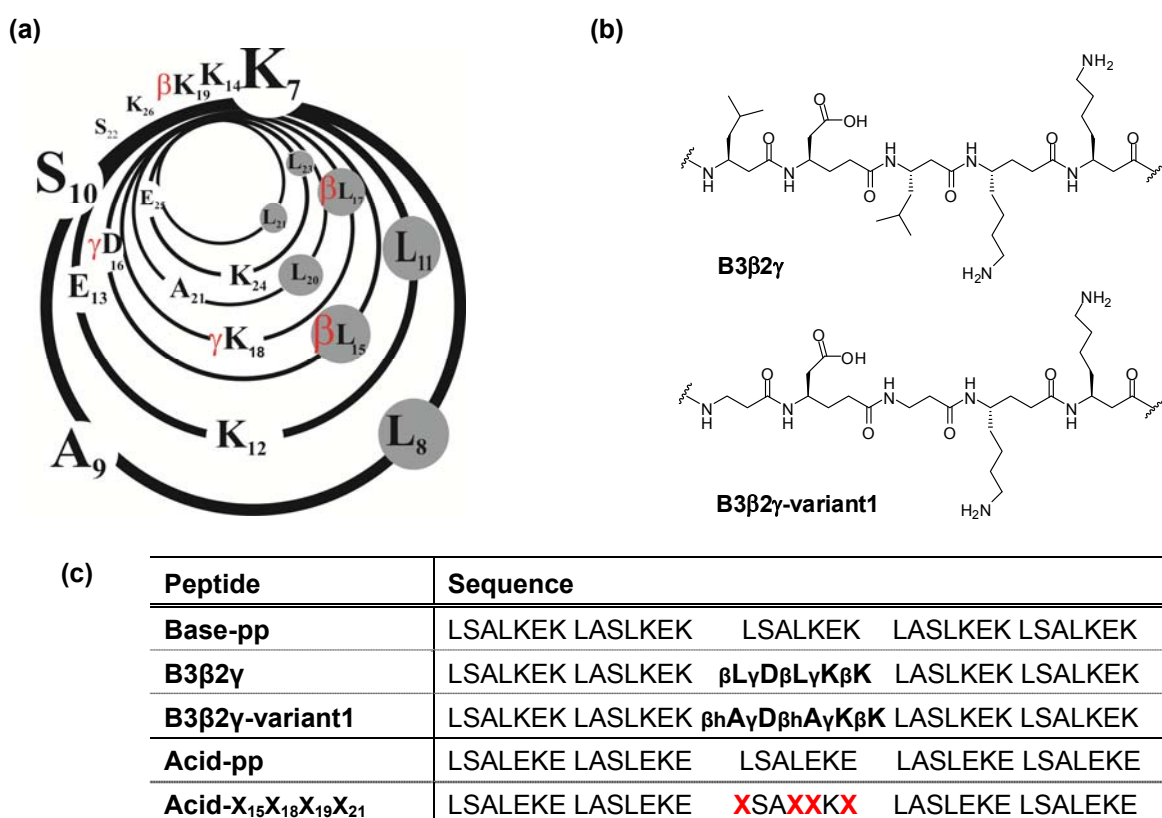


Figure 5.8: (a) Helical representation of B3 β 2 γ 's 21 central residues.⁸³ Amino acids at the hydrophobic core are in grey. (b) Chemical structure of the pentad of alternating β - + γ -amino acids of here investigated foldamer sequences. (c) Amino acid sequences.

Parallel Helix Alignment

In a first experimental set up, the chimera was N-terminally labelled with biotin to provide for loading on streptavidin coated magnetic particles and the Acid-pp library was displayed on the surface of filamentous bacteriophage M13 as an amino terminal fusion of pIII according to procedures outlined in chapter 4.1.^{192,*} Again, the GSG sequence served as a spacer between biotin and B3 β 2 γ . For the evaluation of the selection outcome, a second foldameric sequence, B3 β 2 γ -variant1, was generated accordingly. In this control chimera the two β^3 -homoleucine residues of B3 β 2 γ were substituted with β -alanine, thereby removing two key side chains for coiled-coil interactions.⁸⁴ Both chimeras were screened against the phage displayed Acid-pp library in six panning rounds.

The screening against B3 β 2 γ -variant1 resulted in a low phage number comparable to that of unspecifically bound phages in the negative control (see Figure 9.4, p. 114). Moreover, no consensus sequence but rather a completely random distribution of amino acids occurred in all randomized positions, indicating that there was no specific recognition of the β -alanine backbone. The selection against B3 β 2 γ , on the other hand, led to significant colony enrichment and revealed two most prevalent peptide sequences (denoted Acid-CFLE and Acid-ICEF) of high homology (Table 5.3).

Table 5.3: Amino acids selected in randomized positions of Acid-pp after the 5th and 6th panning round against N-terminally biotin labeled B3 β 2 γ . Amino acids in these positions of the original Acid-pp sequence are given in brackets.

| | Occurrence of phenotype | Acid-pp a' ₁₅ (Leu) | Acid-pp d' ₁₈ (Leu) | Acid-pp e' ₁₉ (Glu) | Acid-pp g' ₂₁ (Glu) |
|-------------------------------|-------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| 5 th panning round | 2x | Ile | Cys | Glu | Phe |
| | 1x | Cys | Phe | Leu | Asp |
| 6 th panning round | 3x | Cys | Phe | Leu | Glu |
| | 2x | Cys | Phe | Leu | Asp |
| | 1x | Val | Leu | Leu | Asp |

Both sequences comprise a cysteine residue, either at position a'₁₅ or d'₁₈, with a Phe next to it in position d'₁₈ or g'₂₁, respectively. In addition, an aliphatic residue (Leu or Ile) occupies one of the four randomized positions. Unlike to their parent peptide Acid-pp, only one Glu residue is found in core flanking positions, instead of two. It is either located at g'₂₁ (Acid-CFLE) or at position e'₁₉ (Acid-ICEF). Similar to previous studies where aromatic residues were found at positions designated for charged amino acids,¹⁸³ the selection of an additional hydrophobic residue at position e'₁₉ or g'₂₁ indicates an extension of the hydrophobic core. The bulky hydrophobic side chains of Ile and Phe are presumably selected to provide ideal shielding from water in the binding groove of the foldamer.

*Results of this section are summarized in: E. K. Nyakatura, ‡ R. R. Araghi, ‡ J. Mortier, S. Wiecezorek, C. Baldauf, G. Wolber, B. Kokschi, *An unusual interstrand H-bond stabilizes the hetero-assembly of helical $\alpha\beta$ -chimeras with natural peptides*, *ACS Chem. Biol.*, 2013, in revision. (‡equal contributors)

The most prevalent Acid-pp variants were chemically synthesized and their interaction with B3 β 2 γ assessed in solution. Applying static light scattering (SLS), it was verified that the tetrameric oligomerization state of the parent system is retained when selected Acid-pp variants assemble with B3 β 2 γ (see Table 9.6, p. 115). The 1:1 mixtures of B3 β 2 γ /Acid-ICEF and B3 β 2 γ /Acid-CFLE formed coiled coils with an increased helical content when compared to B3 β 2 γ /Acid-pp, and investigations on the thermal stabilities of these assemblies leads to significantly higher thermal stabilities in comparison to the B3 β 2 γ /Acid-pp bundle ($T_m=61^\circ\text{C}$) (Figure 5.9). This revealed that the combination of a thiol side chain with a bulky residue seems to match the specific packing requirements of the $\beta\gamma$ -pattern. While B3 β 2 γ /Acid-CFLE has a T_m value of 70°C , the equimolar mixture of B3 β 2 γ /Acid-ICEF starts melting at 74°C , and thus its thermal stability closely resembles that of its native parental system Acid-pp/Base-pp ($T_m > 70^\circ\text{C}$).¹⁹³

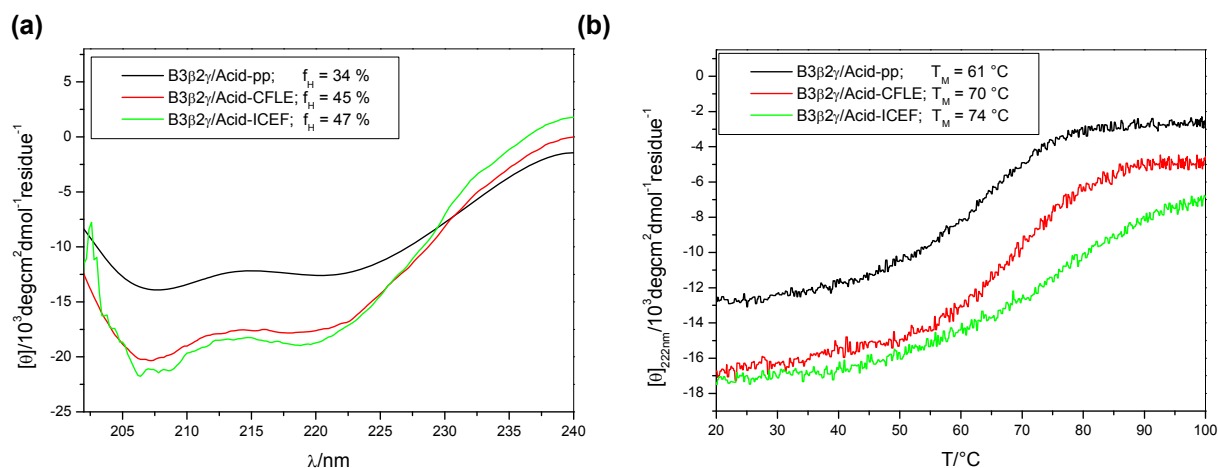


Figure 5.9: (a) CD spectra of equimolar mixtures of B3 β 2 γ /Acid-pp and its variants. Helical content, f_H , given in figure legend. (b) Thermal denaturation spectra of equimolar mixtures of B3 β 2 γ /Acid-pp and its variants. Melting points given in figure legend.¹⁹³ Standard deviations from three independent measurements were $\pm 0.1^\circ\text{C}$. The total peptide concentration was $20 \mu\text{M}$. The spectra were recorded in phosphate buffer 50 mM , containing 0.25 M GndHCl at $\text{pH } 7.4$.

To determine the relative orientation of the helices in the parent B3 β 2 γ /Acid-pp and the selected B3 β 2 γ /Acid-ICEF assemblies, a FRET assay using *o*-aminobenzoic acid (Abz) as the fluorescence donor and 3-nitrotyrosine (Y(NO₂)) as the acceptor was applied.^{25,83,194} Fluorescence quenching by resonance energy transfer from Abz to Y(NO₂) only occurs when the donor and the acceptor are in proximity. In the absence of any orientation-selective element,⁸⁴ the fluorescence spectra of both, N-terminally as well as C-terminally Abz-labeled Acid-pp (Acid-NAbz and Acid-CAbz, respectively), show a progressive decrease in fluorescence intensity at increasing concentrations of N-terminally YNO₂-labeled B3 β 2 γ (B3 β 2 γ -NY(NO₂)) (Figure 5.10). Control experiments in the presence of denaturant (GndHCl) demonstrated that the observed quenching is the result of specific folding rather than self-quenching. Taken together, these results verify the unspecific orientation of the helices towards each other (Figure 9.7, p. 118).

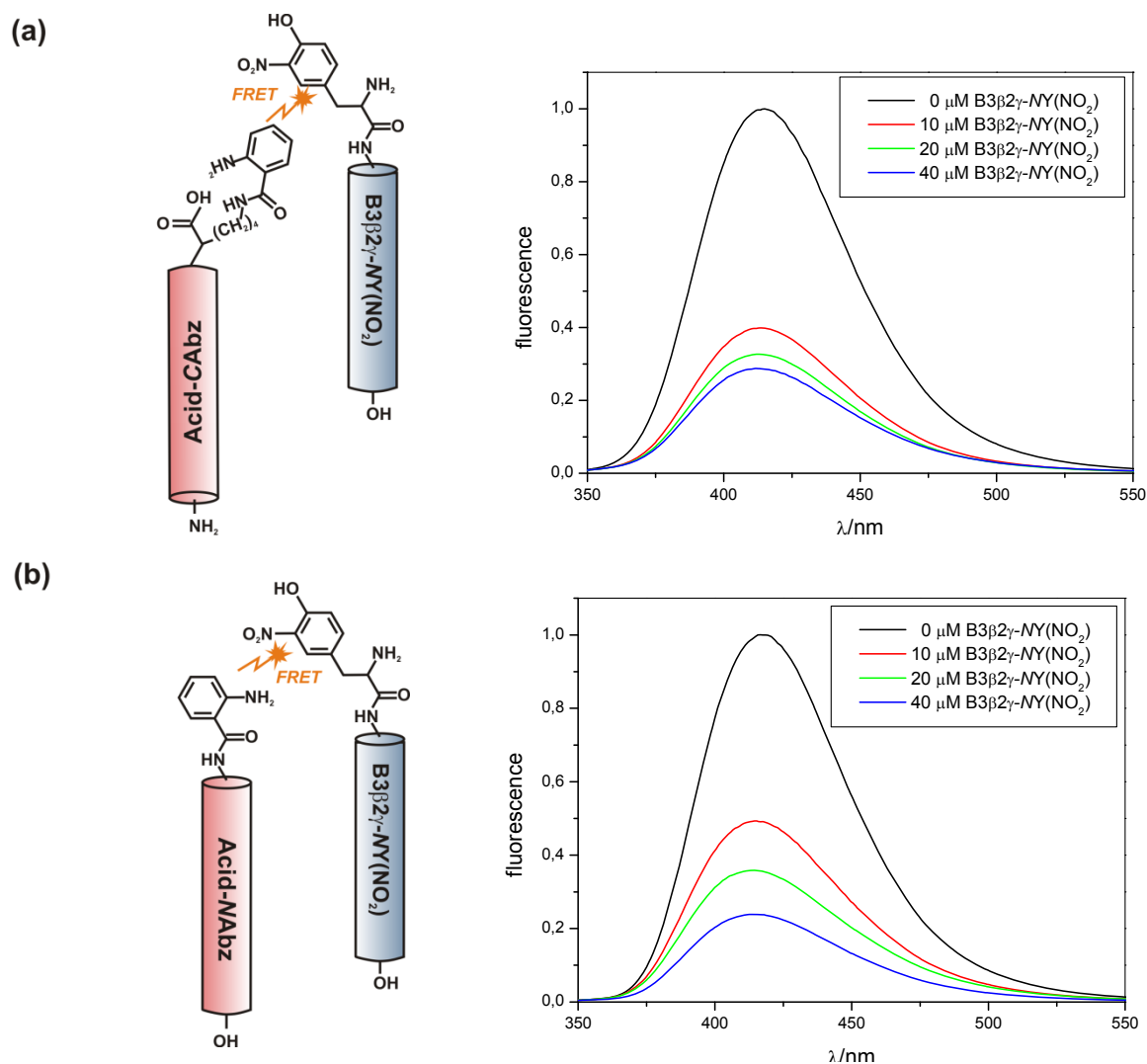


Figure 5.10: Fluorescence spectra of Acid-pp variants at pH 7.4 (10 mM phosphate buffer) in the presence of different concentrations of B3β2γ-NY(NO₂) as depicted in respective figure legend. (a) 20 μM Acid-pp-CAbz; (b) 20 μM Acid-pp-NAbz. The spectra were normalized.

In contrast to these results, FRET experiments in which the fluorescence donor Abz was present at the C-terminus of Acid-ICEF show much weaker quenching than similar experiments with its N-terminally labeled analogue and thus suggest that Acid-ICEF and B3β2γ preferentially form parallel heterooligomers (Figure 5.11). The experimental set up of phage display might have directed the selection of high-affinity binders for the chimera in parallel helix orientation. During panning B3β2γ was N-terminally biotinylated and the Acid-pp library was fused to the N terminus of the pIII coat protein. Thus, a parallel coiled-coil formation on the magnetic particle is feasible, whereas an antiparallel orientation is likely to be sterically hindered.

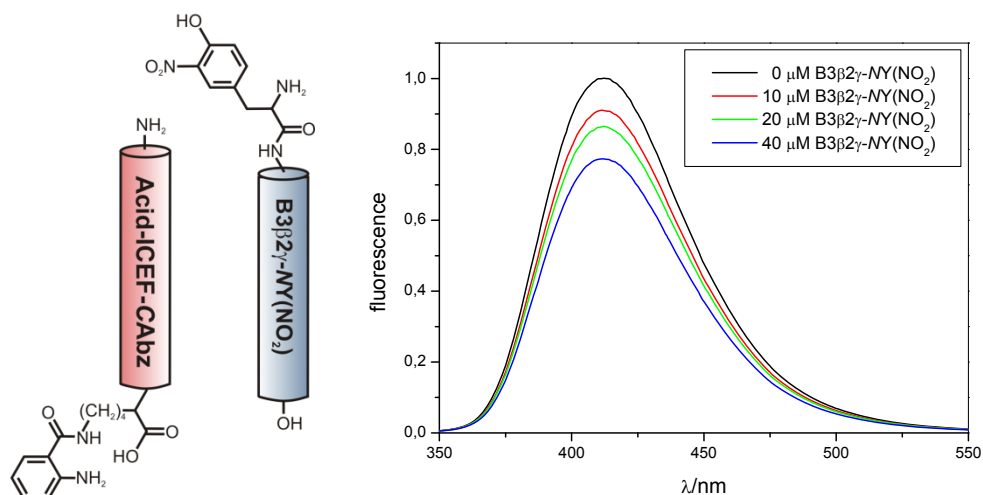


Figure 5.11: Fluorescence spectra of 20 μ M Acid-ICEF-CAbz at pH 7.4 (10 mM phosphate buffer), in the presence of different concentrations of B3 β 2 γ -NY(NO₂) as depicted in figure legend. The spectra were normalized.

Molecular dynamics with the B3 β 2 γ /Acid-ICEF system further confirmed experimentally gained results.⁹⁹ When simulating the tetrameric system in an antiparallel strand orientation no significant alterations could be observed (data not shown). If simulated in a parallel strand orientation, however, three major structural changes occurred: (i) The centers of mass distance between the helices decreased drastically. (ii) The intramolecular H-bond network of the B3 β 2 γ helices rearranged such that the carbonyls of β Leu₁₇ are set free. (iii) The side chain orientation of Acid-ICEF's Cys switches from an intra-molecular interaction with the backbone carbonyl of Glu₁₄ to an inter-molecular interaction with the liberated carbonyl of β Leu₁₇ (Figure 5.12). Moreover, a root mean square deviation (RMSD) analysis indicated two *plateau's* corresponding to two main states adopted by the complex, and thus further verifies the switch from intra- to interhelical H-bonding (Figure 9.8, p. 118). These observations suggest that a strong intermolecular interaction occurs when Acid-ICEF faces the β - and γ -residues. Moreover they further confirm the crucial role that Cys has on the stability of the fold as it might enable an intermolecular H-bond which stabilizes the quaternary structure.

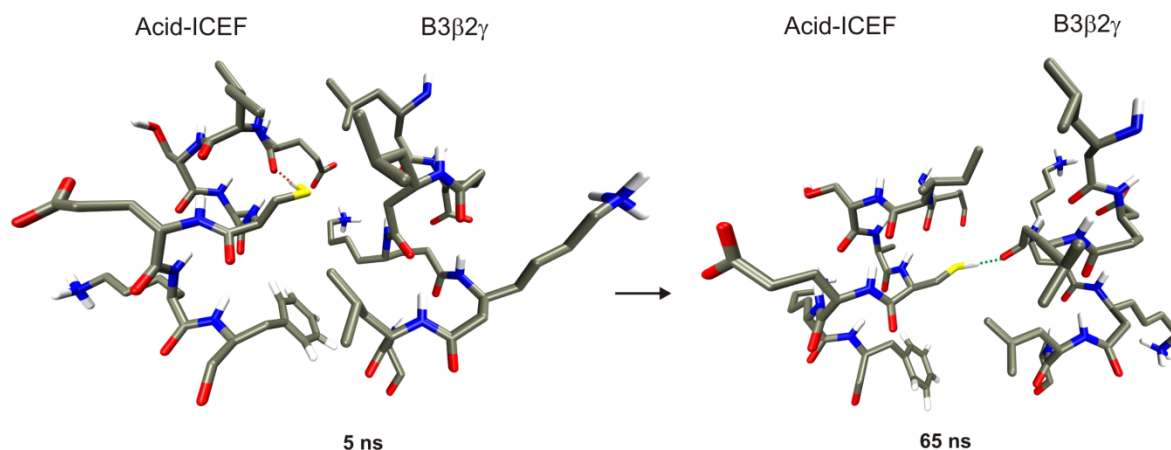


Figure 5.12: Snapshots of MD simulation showing amino acids 14-21 and 15-20 of one Acid-ICEF and one B3 β 2 γ strand, respectively. Red dotted line depicts intrahelical SH Cys18 - CO Glu14 H-bond (2.3 Å). Green dotted line depicts interhelical SH Cys18 - CO β L17 H-bond (2.2 Å).

Hydrogen bond formations involving Cys have been shown to contribute to structural stability of proteins in a large extent.^{195,196} To rule out cysteine's oxidation and thus potential disulfide bond formation, the melting curves as well as size exclusion chromatograms of 1:1 B3 β 2 γ /Acid-ICEF mixtures prior and post incubation with reducing agent DTT were compared (Figure 5.13).^{83,193} The resemblance of the curves before and after DTT treatment gave no indication for oxidation.

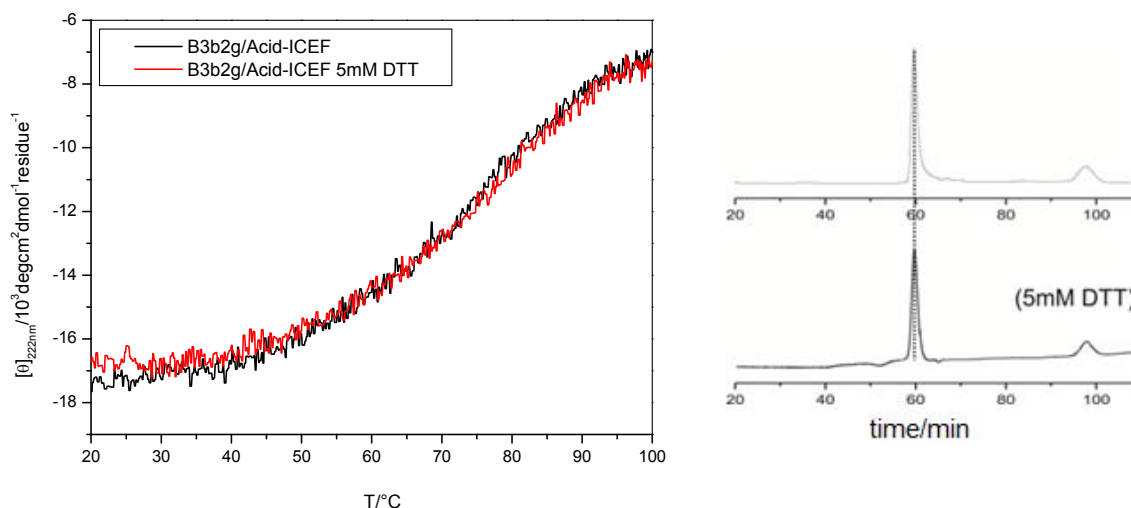


Figure 5.13:¹⁹³ (a) Thermal denaturation spectra of the equimolar mixture of B3 β 2 γ /Acid-ICEF. The spectra are shown before and after treatment with DTT as indicated in figure legend. The total peptide concentration was 20 μM . The spectra were recorded in phosphate buffer 50 mM, containing 0.25 M GndHCl at pH 7.4. (b) SEC chromatograms of 1:1 mixtures of B3 β 2 γ /Acid-ICEF before and after treatment with DTT as indicated in figure legend. The peptides were analysed in 100 mM sodium phosphate pH 7.4. Gly-anthranilic acid was used as the internal standard reference (100 min).

Considering the key function of Cys, the mutation of this residue was expected to affect the stability of the entire motif. To test this hypothesis, two control variants in which the Cys of Acid-ICEF is substituted by either aminobutyric acid (Acid-IAbuEF) or Ser (Acid-ISEF) were generated in line with the present study. While 1:1 mixtures of either of these variants with B3 β 2 γ exhibited considerably less stability in solution than the selected analogue, the variant containing hydrophobic Abu is slightly more stable than the variant containing the polar Ser (Figure 5.14), indicating that the polarity of Ser is stability deterrent. Moreover, no stable intermolecular interaction involving serine's hydroxyl group could be observed in MD simulations with Acid-ISEF. As opposed to Cys residues that are involved in disulfide bonds, free Cys residues appear to have a rather hydrophobic character and are only slightly polar.^{197,198} These results suggest that cysteine's side chain is just polar enough to undergo an interhelical H-bond to a backbone carbonyl and at the same time hydrophobic enough to prevent disruption of the hydrophobic core. Overall, its physicochemical properties (i.e. volume, polarity and side chain reactivity) seem to provide for ideal core packing involving $\alpha\beta\gamma$ -foldamers.

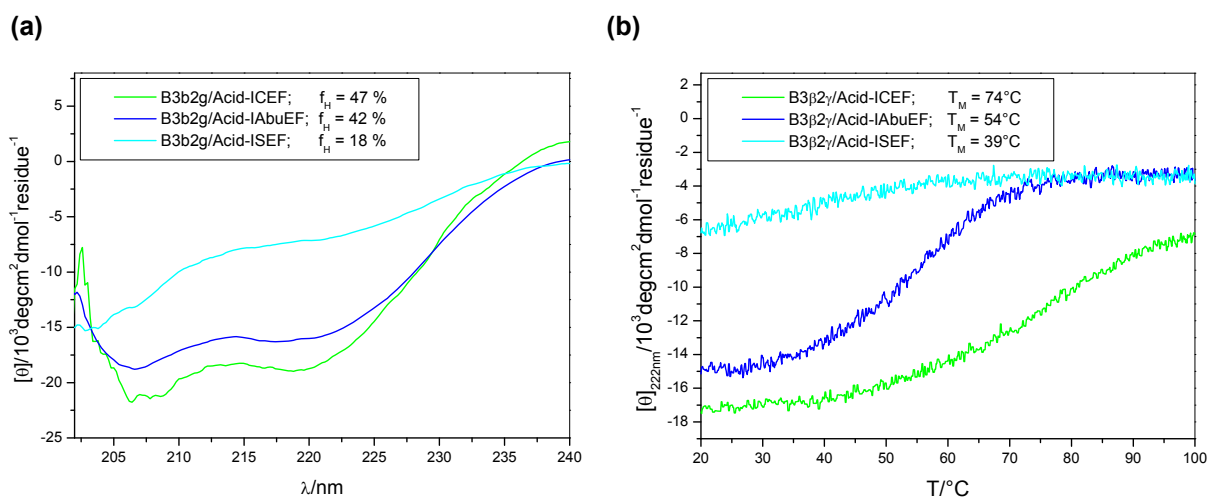


Figure 5.14: (a) CD spectra of equimolar mixtures of B3β2γ/Acid-IXEF as depicted in figure legend. Helical content, f_H , given in figure legend. (b) Thermal denaturation spectra of equimolar mixtures of B3β2γ/Acid-IXEF as depicted in figure legend. Melting points given in figure legend. Standard deviations from three independent measurements were ± 0.1 °C. The total peptide concentration was 20 μ M. The spectra were recorded in phosphate buffer 50 mM, containing 0.25 M GndHCl at pH 7.4.

To conclude, affinity binding studies of B3β2γ/Acid-pp in parallel helix orientation were successfully facilitated. Two of the selected peptides bound the chimeric sequence with higher thermal stability in comparison to its α -parent peptide, while a similar oligomerization state is maintained. The presence of the Cys side chain in combination with a bulky aromatic residue provides for an ideal geometric arrangement of hydrophobic core atoms. The buried Cys residue can significantly influence the core-packing of chimeric coiled coils through the formation of an interhelical H-bond with a non H-bonded backbone carbonyl of the $\alpha\beta\gamma$ -chimera, which was shown to account for an outstanding stability.

5.2.1 Antiparallel Helix Alignment

The in 5.2.1 described study allows for the conclusion that the setup of the undertaken phage display experiment directed the selection of high-affinity binders for the chimera in parallel helix orientation. It was shown that the N-terminally immobilisation of B3β2γ probably led to the selection in parallel helix alignment with the Acid-pp library. To facilitate antiparallel helix orientation on the magnetic particle during biopanning, B3β2γ and B3β2γ-variant1, were synthesized with the biotin linked to the side chain amine of a Lys residue that was introduced at the carboxy terminus. GSG served as a spacer between Lys(bio) and B3β2γ.

Phage display of this second experimental series was carried out in six panning rounds with the phage displayed Acid-pp library of the previous study and according to therein established procedures. Similar to the first experiment, the screen against B3β2γ-variant1 resulted in a low colony number, comparable to that of unspecifically bound phages in the negative control. Screening against B3β2γ, however, resulted in substantial colony enrichment after the fourth and fifth panning round. In agreement with the previous study, these results underline that the β Leu side chains in position a_{15} and c_{17} of B3β2γ are

necessary for the formation of a stable coiled coil, and thus, for the selection of specific binding partners of the $\beta\gamma$ -segment in the undertaken phage display experiments.

Table 5.4: Amino acids selected in randomized positions of Acid-pp after the 5th and 6th panning round against C-terminally biotin labeled B3 β 2 γ . Amino acids in these positions of the original Acid-pp sequence are given in brackets.

| | Occurrence of phenotype | Acid-pp a' ₁₅ (Leu) | Acid-pp d' ₁₈ (Leu) | Acid-pp e' ₁₉ (Glu) | Acid-pp g' ₂₁ (Glu) |
|-------------------------------------|-------------------------|--------------------------------|--------------------------------|--------------------------------|--------------------------------|
| 5th Panning round | 3x | His | Cys | Ala | Asn |
| | 2x | Met | Thr | Glu | Arg |
| | 1x | Leu | Leu | Leu | Leu |
| 6th Panning round | 3x | Leu | Leu | Leu | Leu |
| | 2x | Met | Thr | Glu | Arg |
| | 1x | His | Cys | Ala | Asn |
| | 1x | Leu | Phe | Tyr | Leu |

Selected Acid-pp peptides show a larger variance than the sequences selected post panning in parallel helix orientation. While in case of a parallel helix orientation, two sequences that harbour amino acids of high similarity in randomized positions were found, in this second experimental set up four sequences evolved from sequencing isolated phagemids after the 5th and 6th panning round (Table 5.4). Nonetheless, in the absence of charged residues in any of the randomized positions, two of the selected Acid-pp variants partially reproduce the extension of the hydrophobic core, as observed in previous studies. One of these sequences bears exclusively leucine in all four randomized positions (denoted Acid-LLLL). Similarly, a second peptide (Acid-LFYL) contains primarily hydrophobic amino acids in randomized positions; leucine in a'₁₅ as well as g'₂₁, and phenylalanine in d'₁₈. However, here, the fourth randomized position, e'₁₉, is occupied by polar Tyr, which is potentially involved in H-bonds. Being situated at a hydrophobic core flanking position, intrahelical H-bonding with carboxyl groups of adjacent glutamic acids in *i*+7 or *i*-7 positions are possible. Also an H-bond formation with a backbone carbonyl or Lysine residues of the interacting $\beta\gamma$ -segment of B3 β 2 γ is feasible.

The remaining two sequences harbour charged amino acids in randomized positions. Acid-MTER possesses two oppositely charged amino acids; position e'₁₉ is filled with negatively charged glutamic acid, and g'₂₁ is occupied by positively charged arginine. While glutamic acid equates to the original Acid-pp sequence and is thus most likely engaged in interhelical H-bonds with lysine residues in B3 β 2 γ , arginine might be involved in intrahelical H-bonds with adjacent glutamic acid residues in *i*+7 or *i*-7 distances. Surprisingly, only one of the two randomized core positions harbors a hydrophobic residue (methionine), the other one is filled with polar threonine.

In Acid-HCAN a hydrophobic core position (a'₁₅) is occupied by a charged residue (His). However, in analogy to the selection in parallel helix orientation, cysteine is found in the remaining hydrophobic core position d'₁₈. The occurrence of another polar residue in g'₂₁ in

combination with the rather small hydrophobic alanine in the remaining two randomized positions indicates a stabilization of the assembly via polar and ionic interactions rather than extended hydrophobic interphases.

To study their interaction with B3 β 2 γ , all selected Acid-pp variants were chemically synthesized. Static light scattering revealed that the tetrameric oligomerization state of the parent system is retained when either of the two hydrophobic core extended Acid-pp variants (Acid-LLLL or Acid-LFYL) assembles with B3 β 2 γ (Figure 9.10, p. 121). Interestingly, Acid-MTER forms primarily dimers with B3 β 2 γ , while a small shoulder in its size exclusion peak indicates the presence of monomeric species. In a heterogeneous mixture of B3 β 2 γ and Acid-HCAN the monomeric proportion even prevails and only a small dimeric fraction is observed.

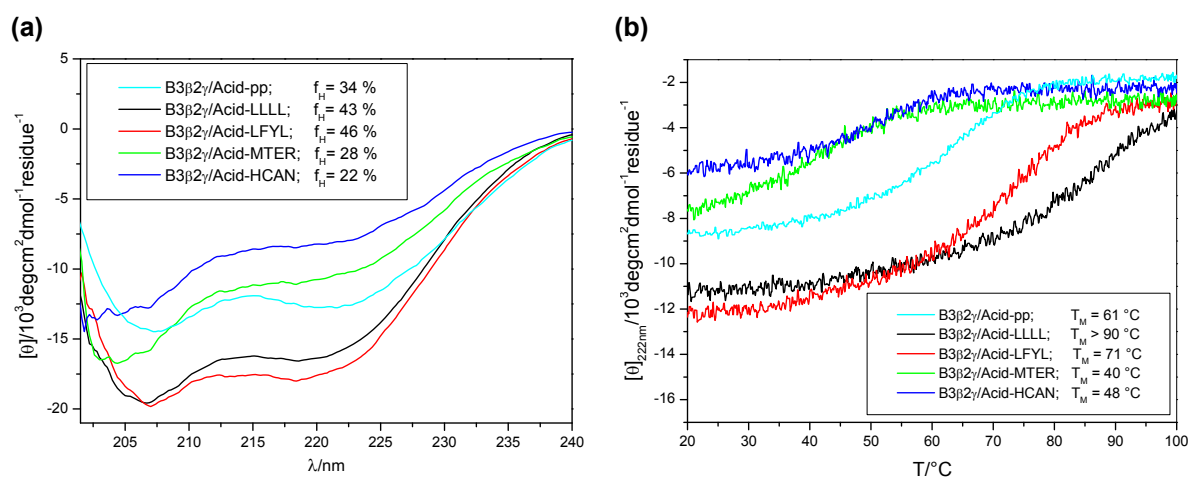


Figure 5.15: (a) CD spectra of equimolar mixtures of B3 β 2 γ /Acid-pp and its variants. (b) Thermal denaturation spectra of equimolar mixtures of B3 β 2 γ /Acid-pp and its variants. Melting points given in figure legend. Standard deviations from three independent measurements were $\pm 0.1^\circ\text{C}$. The total peptide concentration was $20 \mu\text{M}$. The spectra were recorded in 50 mM phosphate buffer, containing 0.25 M GndHCl at $\text{pH } 7.4$.

CD studies showed that, while the 1:1 mixtures of B3 β 2 γ /Acid-LLLL and B3 β 2 γ /Acid-LFYL formed coiled coils with an increased helical content when compared to B3 β 2 γ /Acid-pp, the two selected variants harboring charged amino acids in randomized positions are less helical than B3 β 2 γ /Acid-pp (Figure 5.15). Moreover, investigations on the thermal stabilities of these assemblies revealed fundamental differences among the selected variants. Significantly higher thermal stabilities in comparison to the B3 β 2 γ /Acid-pp bundle ($T_M = 61^\circ\text{C}$), are observed for B3 β 2 γ /Acid-LLLL and B3 β 2 γ /Acid-LFYL. The melting point of B3 β 2 γ /Acid-LFYL is in the range of the thermal stability of Base-pp/Acid-pp, whereas the equimolar mixture of B3 β 2 γ /Acid-LLLL starts melting above 90°C . The latter thus exhibits a substantially improved thermal stability when compared to the native parental system.

Taken together, these results reveal that the extension of the hydrophobic core can significantly contribute to the stability of the tetrameric assembly, while additional polar or ionic interactions might facilitate the formation of coiled-coil bundles that contain alternating β - and γ -amino acids. In light of unsuccessful previous attempts to *de novo* design dimeric

foldamer systems,^{83,134} the reduced thermal stabilities of B3 β 2 γ /Acid-MTER and B3 β 2 γ /Acid-HCAN, compared to B3 β 2 γ /Acid-pp, are not surprising. Additional noncovalent interactions, and the enlarged interface found in higher oligomerization states, such as trimers or tetramers, provide stabilizing effects which counteract the perturbation caused by the unnatural building blocks. The results presented here indicate that the generation of dimeric heteroassemblies involving chimeric sequences is feasible if additional charged amino acids are introduced at the helical interface.

Further studies are required to thoroughly characterize the selection outcomes of the phage display experiments. FRET will be applied to assess whether this second experimental set up directed the selection of high-affinity binders for the chimera under the regime of antiparallel helix orientation. Moreover, MD simulations are currently being conducted to investigate whether the selected side chains could engage in inter- and/or intrahelical interactions.⁹⁹ If such interactions are observed, mutational analyses will be carried out to determine their individual contribution to the overall stability of the selected bundles.

5.3 Towards Protease Stable Fluorinated HIV Entry Inhibitors

The entry of HIV's genome and proteins into the target cell's cytoplasm is orchestrated by its envelope protein subunit gp41, which induces fusion of the viral envelope with the cell membrane. The mechanism of cell fusion proceeds via a conformational rearrangement of gp41 involving association of three C-heptad repeat (CHR) helices along the conserved hydrophobic grooves of a central trimeric N-heptad repeat (NHR) coiled coil in an antiparallel fashion (*cf.* chapter 2.1.2; Figure 5.16).^{56,57} Experimental evidence suggests that a fairly long-lived intermediate exists (pre-hairpin state) in which the CHR and NHR are exposed and available as drug targets.^{57,61,199}

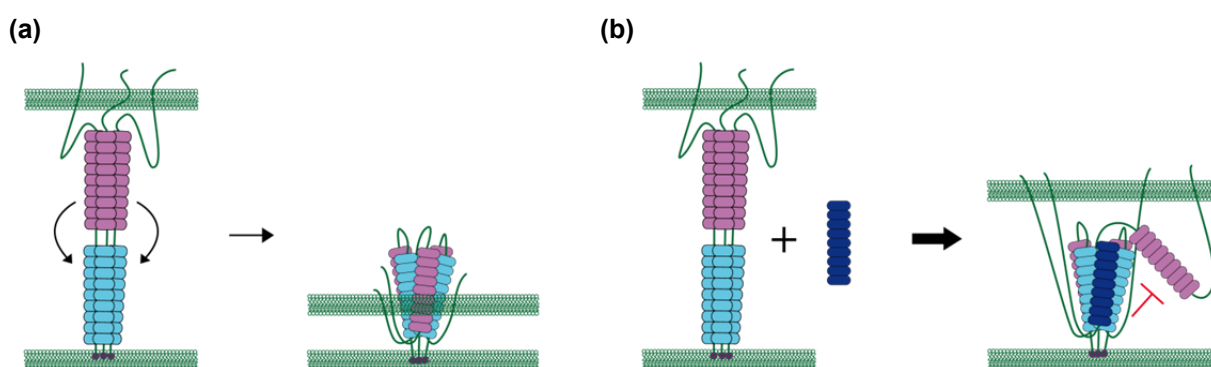


Figure 5.16: Schematic representation of (a) HIV-1 viral fusion via six-helix bundle formation and (b) its disruption through peptide inhibition (navy blue).

5.3.1 Scientific Context of Design

In the early 1900s, a number of synthetic peptides derived from the NHR and CHR regions of gp41 were discovered to have potent anti-HIV activity. Two of the CHR derived peptides C34 and T20 (the latter is also known as enfuvirtide) inhibit HIV infection at low nanomolar levels.²⁰⁰ T20 is a 36 amino acid peptide which corresponds to residues 637-673 of gp41 (Figure 5.17). While T20 is clearly able to interact with the pre-hairpin state of gp41, the host cell membrane also plays a putative role in its inhibitory mode of action.^{64,201} C34 is a 34 amino acid peptide consisting of residues 628-661 that, in contrast to T20, include the eight amino acid pocket binding domain (628-635) which is thought to be essential for association with the NHR. Three highly conserved hydrophobic residues (Trp⁶²⁸, Trp⁶³¹ and Ile⁶³⁵) of this motif penetrate into a deep hydrophobic pocket formed by residues 565-579 of the NHR, resulting in a fairly strong interaction.²⁰²

Resistance to T20, the only peptide based fusion inhibitor currently in clinical use, is governed by mutational changes in a stretch of amino acids in and adjacent to the ⁵⁴⁷GIV⁵⁴⁹ motif of the NHR.²⁰³ Moreover, the pharmacologic utility of this lengthy peptide, which is purely composed of canonical α -amino acids, is hindered by a loss of bioactive structure and rapid proteolysis.⁶⁸ Since this limited bioavailability attenuates the efficacy of the drug, T20 has to be applied intravenously.

While the C-peptides are α -helical when bound to the gp41 NHR, they are unstructured when isolated in solution.²⁰⁴ Therefore, the C-peptide binding event exacts a large energetic penalty due to the loss of conformational entropy, going from various unstructured configurations to a fixed helical conformation. Studies by Matusoka *et al.*, have shown that the introduction of several oppositely charged glutamic acid/lysine pairs in i and $i+4$ positions, at the solvent accessible site of C34, increases its α -helicity and water solubility, while also stabilizing the CHR/NHR bundle.^{205,206}



Figure 5.17: Sequence of NHR and CHR regions of gp41 depicted as a hairpin. Heptad repeat shown in black. Pocket forming domain (red) and pocket binding domain (PDB; blue) crucial for 6-HB formation. Region N-terminal adjacent to PDB (green) stabilizes gp41 core structure.⁵⁴⁷GIV⁵⁴⁹ (purple) is responsible for resistance in T20 and C34. FP corresponds to fusion peptide. The regions corresponding to two NHR derived peptides, N36 and T21, are shown above the sequence. The regions corresponding to two CHR derived peptides, C34 and T20 are shown below the sequence.

Multiple mutation and binding studies have pointed to a critical role for the residues W628, W631, I635, and I642 of gp41, as well as the formation of a salt bridge between D632 and K574.²⁰⁷⁻²¹² These residues contribute substantially to the stability of the six helix bundle, and thus the inhibitory potency of CHR derived peptides. Moreover, He *et al.* identified the ⁶²¹QIWNMT⁶²⁷ motif located upstream of the CHR, immediately adjacent to the pocket binding domain (PBD), as being critical for NHR and CHR interhelical interactions, and that CHR derived peptides containing this motif possess nanomolar anti-HIV activity.^{200,213} Contemporaneous investigations by Gochin *et al.* with a set of CHR derived peptides that differed in α -helical content and charge, revealed that a negative net charge (-2 or -3) combined with nonspecific hydrophobic interactions and high α -helical content correlates with the inhibitory potency of investigated peptides.²⁰⁷ Examining the upstream region of the gp41 CHR including residue E620 the authors found that specific salt bridge and hydrogen bond networks may reside in this region and thus their results support He *et al.*'s findings that an extended interaction site upstream of the PBD may be necessary to achieve higher inhibitory potency. Moreover, snap shots derived from two independent homology modeling experiments found in literature, revealed that R579 of NHR can form an intermolecular salt bridge with E620.^{207,214}

Since the hydrophobic pocket is highly conserved among strains,⁵⁶ it is considered to be an ideal target for the development of novel HIV fusion inhibitors. The high degree of

conservation likely renders this region less prone to mutations that confer drug resistance than the upstream region of the NHR. Taking these recent findings into account, CHR derived HIV fusion inhibitors were designed in the context of the current thesis with the purpose of targeting the hydrophobic pocket as well as its C-terminal region. Therefore, two approaches were applied:

- (i) Introducing one or several covalent backbone crosslinks in C-peptides of different length to increase their α -helical content, and thus investigate a potential shortening of the required sequence length for sufficient interaction with the NHR.
- (ii) Site specific substitution of an aliphatic residue within the pocket binding domain with fluorinated amino acids of different side-chain volume and polarity to elucidate their optimal packing and binding specificity at the hydrophobic interface.

5.3.2 Short Constrained C-Peptides

To stabilize the helical conformation of the C-peptide in the unbound state and thus to counteract the energetic penalty during binding, covalent backbone crosslinks were introduced into C-peptide analogues of differing length in the course of the current study (Figure 5.18). For this, the from Verdine *et al.* engineered amino acid (*S*)- α -methyl, α -pentenylglycine (referred to as S5)^{142,144} was utilized. If positioned at the *i* and *i*+4 positions, S5 side chains are able to undergo ring-closing metathesis (RCM, *cf.* chapter 2.2.3). This crosslink should not only stabilize the helical peptide conformation, but should also enhance proteolytic stability.

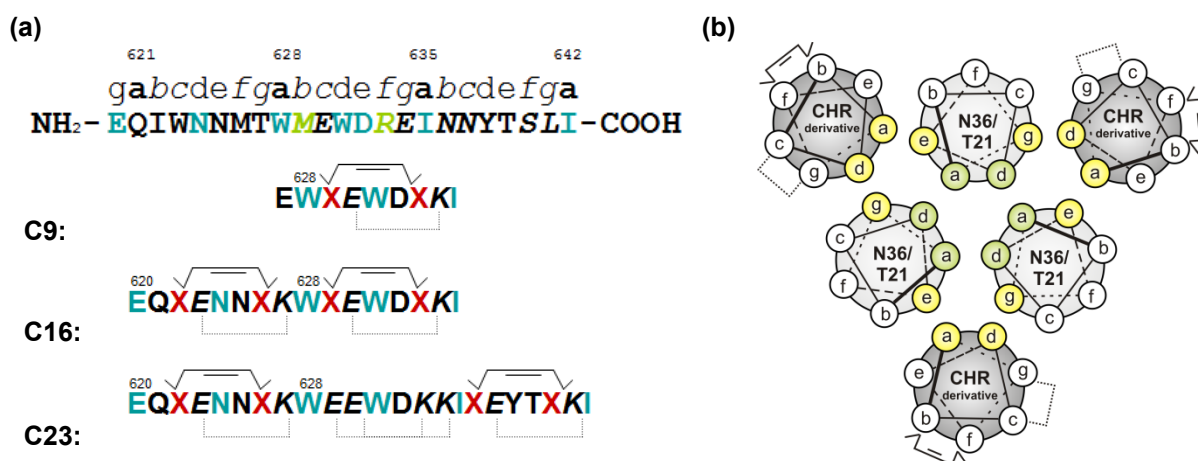


Figure 5.18 (a) Sequences of synthesized peptides as well as the CHR sequence they correspond to (top). Residues interacting with NHR highlighted in turquoise. S5 residues indicated as red aces. E/K ion pairs are in italic. (b) Helical wheel depiction of their proposed interaction with NHR derived peptides.

The shortest of the peptides engineered here, C9, corresponds to the 8 amino acid sequence of the pocket binding domain (628-635). This sequence was extended by an N-terminal glutamic acid residue, to achieve an overall netcharge of -2. Amino acids that have been identified to be essential during NHR binding remained unmodified. S5 was introduced in *i*, *i*+4 distance at the solvent exposed positions *b* and *f*. Since the backbone crosslink of these

two residues introduces additional hydrophobicity, glutamic acid/lysine pairs were incorporated at the remaining solvent exposed positions (*c* and *g*), in analogy to studies by Matusoka *et al.* (see chapter 5.3.1).^{205,206}

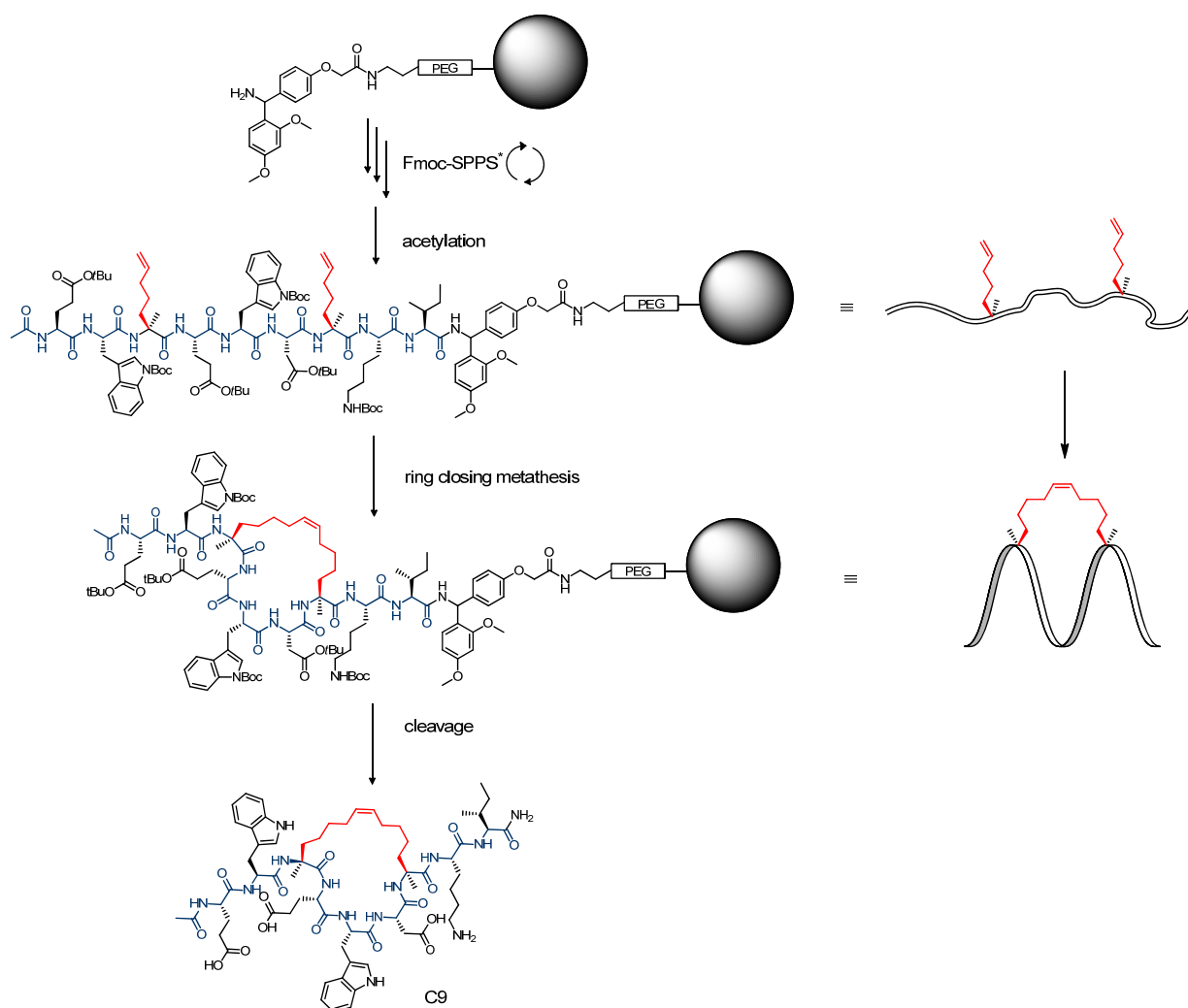


Figure 5.19: Synthesis scheme for backbone crosslinked peptide C9. *see chapter 7.1.1, p. 83-85 for experimental details.

A second peptide, C16, extends C9 N-terminally by 7 amino acid residues, corresponding to the end residue E620, to enable an intermolecular salt bridge with R579 of the NHR. Here, two pairs of S5 as well as two glutamic acid/lysine pairs were introduced, to further enhance α -helicity. The overall net charge of this peptide is also -2. Finally, a third peptide, C23, was generated that extends the PBD N-terminally as well as C-terminally by 7 amino acids. One pair of S5 residues and one E/K pair were introduced at each terminus, while the central region harbors a glutamic acid/lysine pair in *b/f* as well as *c/g* positions. Figure 5.19 shows the synthetic strategy with C9 as an example.

To examine the impact of the covalent backbone crosslink on α -helicity, CD spectroscopic analysis was performed. While CD spectra reveal that C9 and C16 are unstructured in their open form, the CD spectrum of open C23 indicates slight α -helical character (Figure 5.20a).

However, post RCM, all three peptides possess increased α -helicities with similar α -helical contents (Figure 5.20b).

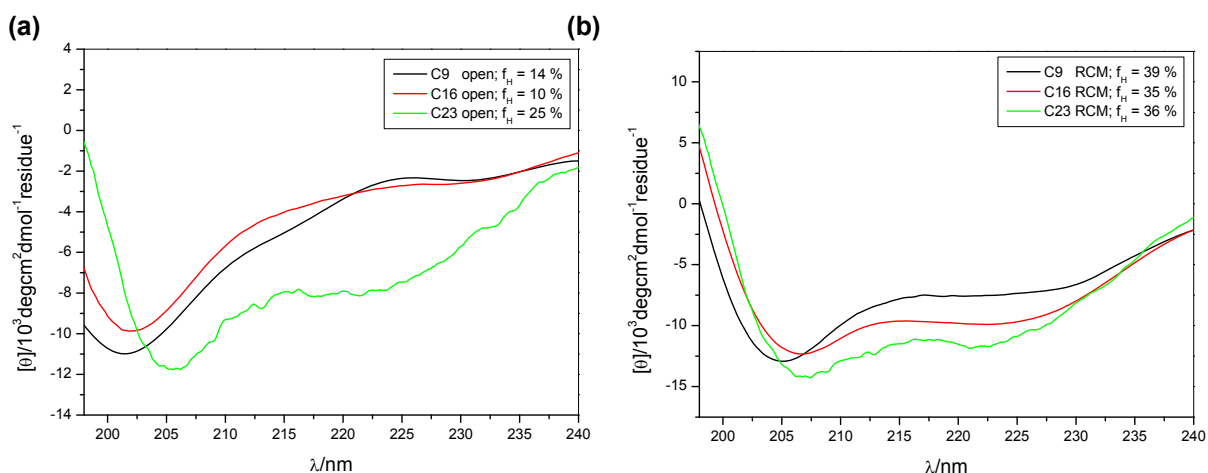


Figure 5.20: CD spectra of CHR derived peptides in their open (a) and closed (b) form. $10\mu\text{M}$ peptide concentration; 50mM Phosphate buffer, 150mM NaCl.

As mentioned above, the amino acids that have been shown to interact with the NHR are well distributed over the full length of C23, and, thus, a large interfacial surface area is involved in the binding of CHR derived peptides to the complimentary binding groove. With regard to this, the longest of the C-peptides investigated here is likely to be more potent in interacting with NHR derived peptides than its shorter analogues. To determine whether the engineered peptide sequences possess any binding affinity to the NHR, initially a 1:1 mixture of C23_{RCM} and N36 was incubated for 30 min at 37°C , and subsequently spectroscopically analyzed according to procedures outlined in literature.²¹³ In this experiment, the corresponding peptide sequence, solely consisting of the natural occurring amino acids served as the unmodified reference (C23_{Ref}). C34 was implemented as a positive control. The six-helix bundle formation of C34/N36 is demonstrated by the two distinct minima characteristic for α -helical structures. On the contrary, the CD spectra of 1:1 mixtures of C23_{RCM} and C23_{Ref} indicate that these peptides do not bind to the N-peptide of gp41.

N36 corresponds to the amino acid sequence that directly interacts with C34 (see Figure 5.17). When compared to C34, the sequence of C23 is N-terminally shifted, which would result in an overlap of noninteracting single strand peptide on both termini of a theoretical C23/N36 complex. Therefore, the interaction with a second N-peptide sequence, T21, comprising a larger theoretical interfacial contact area with the C23 sequence, was assessed in a second experiment to allow for improved alignment of the helices (Figure 5.21b). Again, no six-helix bundle formation could be detected by CD spectroscopic analysis under the applied conditions.

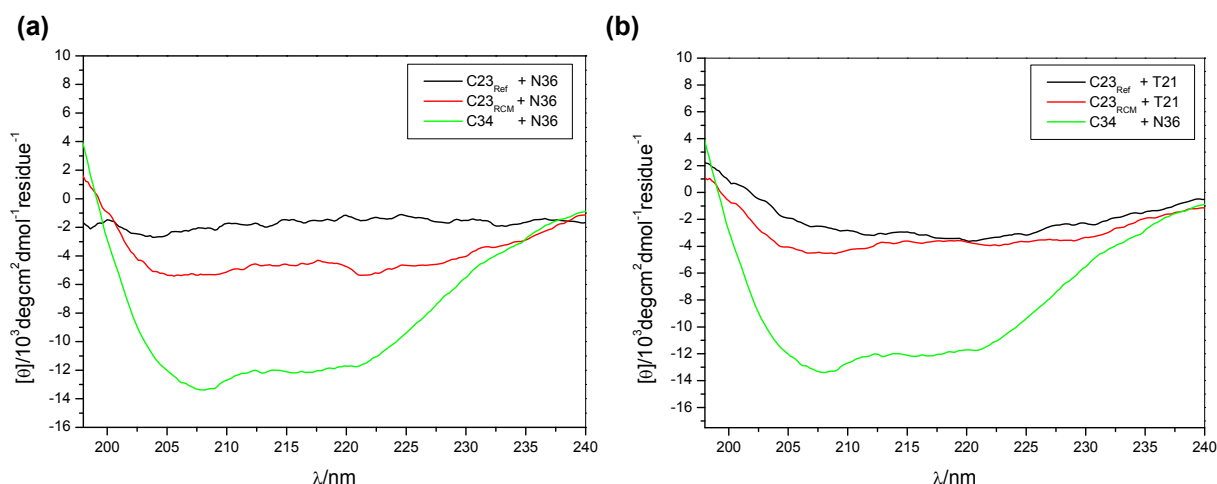


Figure 5.21: CD spectra of $C23_{RCM}$ and its reference in equimolar mixture with N36. (b) CD spectra of $C23_{RCM}$ and its reference in equimolar mixture with T21. 10 μ M peptide concentration; 50mM Phosphate buffer, 150 mM NaCl.

Even though the α -helicity was clearly improved with $C23_{RCM}$ in comparison to $C23_{Ref}$, apparently neither of the two peptides is able to undergo binding to the NHR. Nearly simultaneous to the generation of this results, data published by Gellman *et al.* indicates that the affinity between CHR and the complimentary N-helices depends upon interactions distributed across the extended protein interface.⁶⁹ While their findings highlight a very tight binding within the hydrophobic pocket, several residues downstream of the PBD contribute significantly to NHR binding that are not comprised in the sequence of $C23_{Ref}$. These findings contradict the prior generally accepted hypothesis that the binding affinity is dominated by interactions of a small cluster of side chains within the C-terminal region of the NHR. In agreement with these findings, here outlined results indicate that a CHR peptide that comprises only one heptad of the downstream region of the PBD does not provide for sufficient binding affinity, even if the stability providing ⁶²¹QIWNNMT⁶²⁷ motif is included. Therefore, two additional CHR derived peptide sequences were generated, which were extended by an additional amino terminal heptad repeat (denoted C33 and C31; Table 5.5), and their ability to interact with the NHR derived sequence T21 assessed by CD spectroscopy.

Table 5.5: Sequences of C-terminally extended CHR derived peptides

| Peptide | Sequence |
|------------|---------------------------------------------------------------------------------|
| C33 | 620 652 EQIWNNMTWMEWDREINNYTSLIHSLIEESQNQ |
| C31 | 620 650 EQIWNNMTWMEWDREINNYTSLIHSLIEESQ |

CD spectra of 1:1 mixtures of these newly generated C-peptides and T21 clearly suggest the formation of hetero-assemblies (Figure 5.22). Moreover, the helical content of both mixtures as well as their thermal stabilities are substantially increased in comparison to C34/N36 (T_M literature value for C34/N36 bundle at identical conditions = 64 °C).²¹³ Hence, these

results support previous findings pointing at a stability providing character of the region C-terminally adjacent to the PBD.

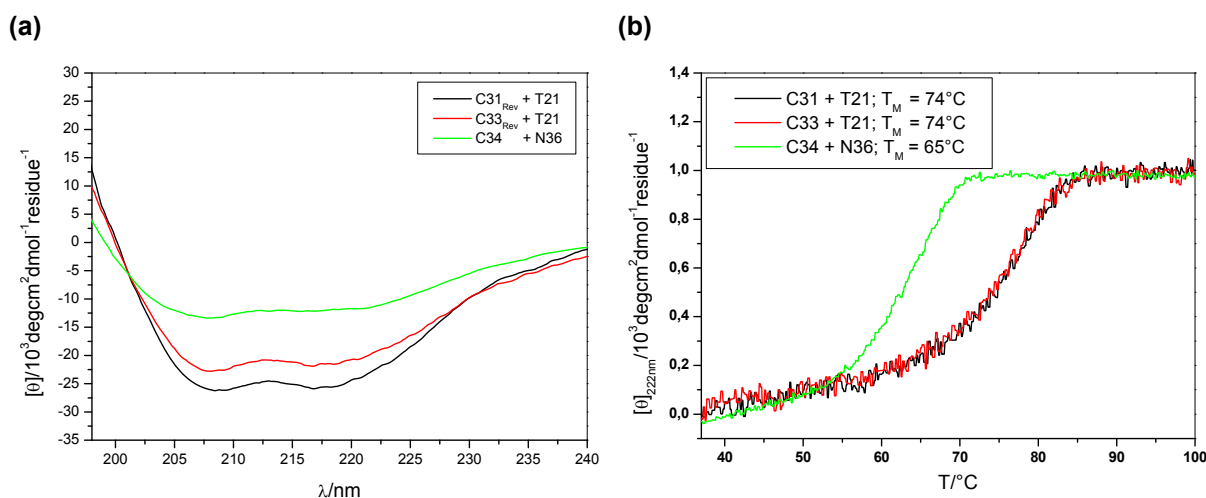


Figure 5.22: (a) CD spectra of newly generated peptides in equimolar mixtures with T21. (b) Thermal denaturation spectra of equimolar mixtures. $10\mu\text{M}$ peptide concentration; 50mM Phosphate buffer, 150mM NaCl.

To further verify heterooligomerization, an N-PAGE experiment according to procedures introduced by He *et al.* was applied.²¹³ Again, the C34/N36 bundle served as a positive control. Consistent with He *et al.*'s observations, the N-peptides N36 and T21 show no band in the native gel. This effect can be attributed to their net positive charges which result in their migration up and off the gel. All investigated C-peptides (negative net charge) show specific bands. The bands of the 1:1 mixtures of N- and C-peptides migrate significantly slower than the peptides alone. Hence, native PAGE confirms the six helix bundle formation (Figure 5.23).

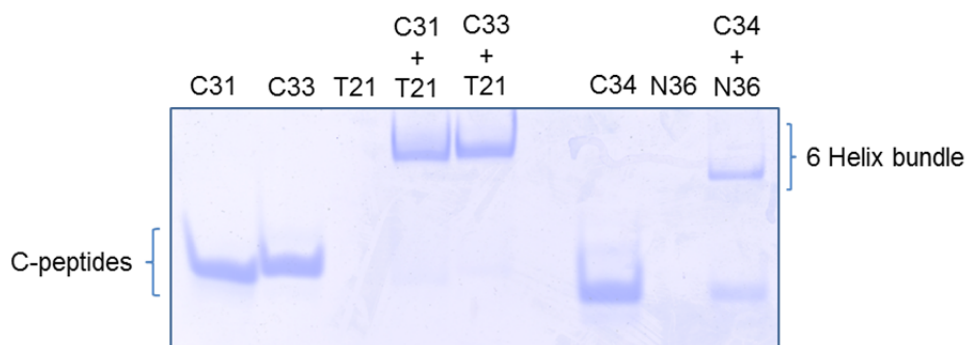


Figure 5.23: Native PAGE (18% Tris glycine) of C31, C33, T21, and 1:1 mixtures of C-peptides with T21. C34, N36 and C34/N36 served as a positive control. $60\mu\text{M}$ peptide concentration.

In summary, gained results suggest that herein generated intrahelically crosslinked C-peptide analogues do not appreciably bind to NHR derived target sequences, even though the combination of covalent backbone linkages and ion pairs substantially enhanced their α -helical content. Because the intermolecular binding interface of the gp41 fusion complex is large, shallow, and hydrophobic, peptides with a length of 30 residues and longer are generally significantly more potent than short peptides and small molecules.^{68,69} However,

several short peptides with stabilized α -helical conformations and moderate inhibitory efficiency have been reported.^{151,204,215} Hence, the inability of here investigated peptides to bind to the NHR cannot exclusively be attributed to their sequence length. All crystal structures available at the time of peptide design lack the residues N-terminal to the PBD. A recently obtained high resolution crystal structure comprising the majority of this motif revealed that this motif does not maintain the α -helical conformation.²⁰² Moreover, results of Gellman *et al.* indicate the broad distribution of energetically important contacts across the extended CHR/NHR interface located C-terminal to the PBD.⁶⁹ The longest of the peptide sequences that were initially investigated in the current study comprised only one heptad unit upstream of the PBD. According to Gellman *et al.*'s studies, NHR-binding is also facilitated via residues in *e* and *g*-positions, which were used as modification positions during rational design of the CHR analogues described here. Combined, recent findings explain why the initially generated C-peptide sequences fail to bind to the NHR derived peptides.

Because the length of the shorter analogue of the two additionally generated C-peptides (C31) is sufficient for stable six-helix bundle formation with T21, it was chosen as a starting point for further modification. To circumvent the substitution of key *e* and *g*-positions, bis-(4-pentenyl)glycine could be introduced in suitable *f*-positions together with S5 amino acids in adjacent *c*- and *b*-positions in future intrahelical crosslinking approaches (Figure 5.24). Dihydroxylation of the double bonds post RCM²¹⁶ would render the covalent backbone crosslinks less hydrophobic, resulting in a polar face at the solvent exposed positions of the helix.

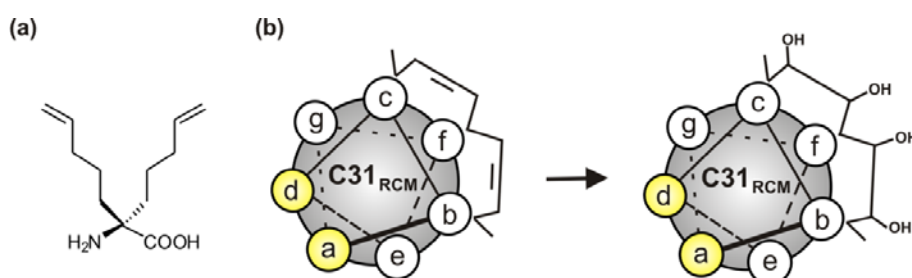


Figure 5.24: (a) structure of bis-(4-pentenyl)glycine, (b) dihydroxylation of introduced double bonds post RCM.

5.3.3 Incorporation of Fluorinated Amino Acids

Studies outlined in chapter 5.1 show that fluorinated analogs of hydrophobic amino acids are readily accommodated in coiled-coil assemblies. Previous hydrophobicity studies of the Kokschi group have shown that fluoroalkyl side chains possess two seemingly contrary physicochemical properties, hydrophobicity and polarity.^{96,97} Complimenting these findings, Goa *et al.* showed that tetrafluorinated phenylalanine isomers possess both enhanced hydrophobicity and a substantially greater partial positive charge on the remaining hydrogen atom (*cf.* chapter 2.2.1).^{105,107} When accurately positioned, these fluorinated building blocks

can tremendously stabilize CH- π interactions at protein cores. Investigating the crystal structure obtained from Kim *et al.*⁵⁶ reveals that the hydrogen atoms of the γ -methyl group of the highly conserved CHR residue I635 might be involved in CH- π interactions with W631 within the hydrophobic pocket (Figure 5.25).

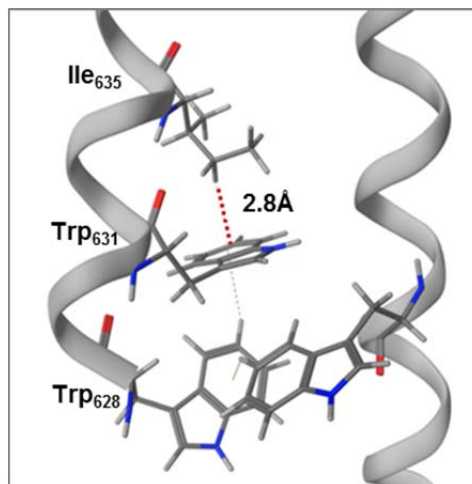


Figure 5.25: segment of the crystal structure of C34/N36, (PDB code: 1AIK.) Figure created using pyMOL v0.99.

In the context of this thesis, a line of investigation was proposed, in which three fluorinated amino acids with increasing side-chain volume and fluorine content (namely difluoroethylglycine (DfeGly), trifluoroisoleucine (5^3 -F₃Ile), and hexafluoroisoleucine ($5^3,5^3$ -F₆Leu)) will substitute for the relevant Ile in C31 (I16) (Figure 5.26).

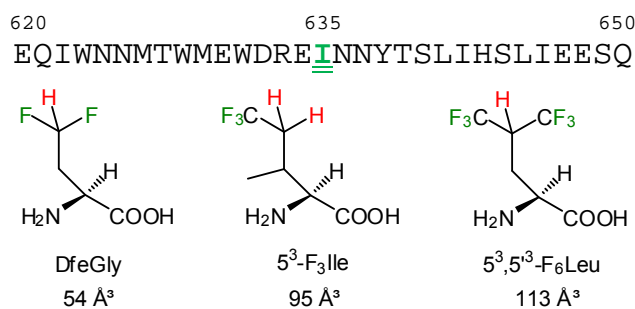


Figure 5.26: Top: sequence of C31. Substitution site highlighted in green. Bottom: Structures and side-chain volumes of fluorinated amino acids. vdW volumes of the amino acid side chains were calculated according to Zhao *et al.*¹⁸⁹ The through fluorine atoms polarized hydrogen atoms are highlighted in red.

While the overall hydrophobicity of these analogues should be increased in comparison to their hydrocarbon analogues, the close proximity of the fluorine atoms should simultaneously enhance the partial charges of relevant H-atoms. Thus, with the fluorinated peptides the influence of CH- π interactions might be probed in the context of size and electronic alterations. The fluorinated building blocks hexafluoroisoleucine and trifluoroisoleucine have been synthesized and kindly provided by the Czekelius group.

Since an understanding of side-chain hydrophobicity and α -helix propensity is of crucial importance for interpretation of the upcoming study, these two properties were assessed in preliminary studies.*

Hydrophobicity

First, the relationship between side-chain volume and hydrophobicity of 5^3 -F₃Ile and $5^3,5^{3'}$ -F₆Leu were investigated, by plotting their side-chain van der Waals volumes versus their retention times in an RP-HPLC experiment.²¹⁷ Here, the nonpolar phase of the reversed-phase column serves as a mimic for a hydrophobic interaction.⁹⁸ DfeGly was previously investigated in similar studies of the Kokschi group and results thereof,^{96,97} were kindly provided to the author of this thesis. Initial studies of the group included the canonical aliphatic amino acids and 2-aminobutyric acid. In line with current studies, 2-aminoheptanoic acid (Aha) was used to extend the initial set of hydrocarbon residues. The van der Waals volumes of the amino acid side chains were calculated according to Zhao *et al.*¹⁸⁹ Aha correlates very well with its smaller nonfluorinated analogues. Their retention time increases nonlinearly with increasing side-chain volume (Figure 5.27). As expected, the enlargement of the aliphatic side chain results in an increase of hydrophobicity.

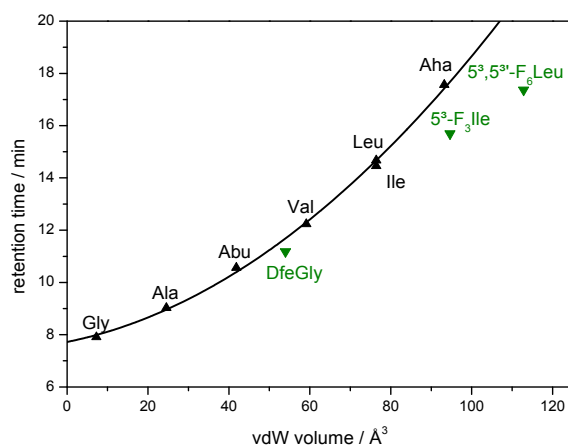


Figure 5.27: Retention times of Fmoc amino acids plotted against the van der Waals volume of their side chains. Nonfluorinated amino acids are depicted as black triangles; their correlation is shown as a black line. Fluorinated amino acids are represented by green triangles.

In accordance with earlier studies, the retention times of the fluorinated amino acids 5^3 -F₃Ile, and $5^3,5^{3'}$ -F₆Leu do not fit into the correlation between side-chain volume and retention time.⁹⁶ As previously observed, DfeGly is larger and more hydrophobic than its nonfluorinated analogue Abu. Although similar in size, 5^3 -F₃Ile is less hydrophobic than Aha. $5^3,5^{3'}$ -F₆Leu is similar to Aha in hydrophobicity, while exhibiting a much larger volume. In a free energy perturbation study, the hydration energy of $5^3,5^{3'}$ -F₆Leu was shown to be 1.1

* Results of this section are summarized in: H. Erdbrink‡, E. K. Nyakatura‡, S. Huhmann, U. I. M. Gerling, D. Lenz, B. Kokschi, C. Czekelius, *Synthesis of enantiomerically pure (2S,3S)-5,5,5-trifluoroisoleucine and (2R,3S)-5,5,5-trifluoro-allo-isoleucine*, *Beilstein J. Org. Chem.*, 2013, 9, 2009-2014. (‡equal contributors)

kcal/mol higher than that of leucine.¹¹⁶ In combination, herein and previously gained results suggest that there are two factors that determine the overall hydrophobicity of the fluorinated amino acids.⁹⁶ On the one hand, substitutions of hydrogen by fluorine increase the solvent accessible surface area and thus lead to an increase in hydration energy. On the other hand, the C-F bond is more polarized than the C-H bond, and electrostatic interactions of the fluorinated group with the solvent are energetically more favored. As a consequence, fluoroalkyl side chains possess two seemingly contrary physicochemical properties, hydrophobicity and polarity, and the combination of both renders fluorinated amino acids to be less hydrophobic than their surface area alone would suggest.

***α*-Helix Propensity**

In general, fluorination of amino acids leads to a dramatic decrease in helix propensity.^{111,112,218} In the course of the current study, the α -helix propensity of 5³-F₃Ile was investigated according to methods established by Cheng *et al.*, who showed that when an amino acid of interest is incorporated into an α -helical polyalanine model peptide (KX), its α -helix propensity can be calculated from circular dichroism (CD) spectroscopy.^{111,112} Therefore, 5³-F₃Ile was converted into its Fmoc protected form and subsequently used in solid-phase synthesis of K-5³-F₃Ile by applying standard Fmoc-based chemistry.^{181,217} The α -helix propensity [ω] was calculated from CD data (Table 5.6).

Table 5.6: Ellipticity [Θ] at 222 nm was taken from normalized CD data. Fraction helix [f_{helix}] and helix propensities [ω] were calculated from [$\Theta_{222\text{nm}}$] applying a modified Lifson–Roig theory^{219,220}

| Peptide | [$\Theta_{222\text{ nm}}$] | f_{helix} | ω |
|-----------------------------------------------------------------------|------------------------------|--------------------|-------------|
| K-Ile | -13813 ± 156 | 0.40 ± 0.01 | 0.52 ± 0.05 |
| K-5³-F₃Ile | -10776 ± 216 | 0.31 ± 0.01 | 0.26 ± 0.03 |
| K-Leu | -17400 ± 200 | 0.50 ± 0.01 | 1.06 ± 0.12 |
| K-5³,5^{3'}-F₆Leu ¹¹² | - 3602 ± 130 | 0.10 ± 0.01 | 0 |
| K-Abu ¹¹² | -18100 ± 200 | 0.52 ± 0.01 | 1.22 ± 0.14 |
| K-DfeGly ²²¹ | -13969 ± 569 | 0.40 ± 0.02 | 0.50 ± 0.06 |

Although the α -helix propensity of 5³-F₃Ile is half of that for Ile, the α -helix propensity of this amino acid is partially retained. Similarly, the α -helix propensity of DfeGly, is significantly reduced when compared to its hydrocarbon analogue Abu. The extensive fluorination of Leu, 5³,5^{3'}-F₆Leu, even leads to a complete loss of α -helix propensity. The dramatic decrease in helix propensity upon fluorination was previously attributed to a possible burial of fluorocarbon side chains in the unfolded state of the model peptide.¹¹² The exposure of this side chain in the helical state would lead to destructive helix formation energetics, due to the enhanced hydrophobic nature of the fluorocarbon side chains. However, it remains to be elucidated to what extent helix propensity affects protein stability at buried positions, like the hydrophobic pocket of gp41, since the reduced helix propensity may in part be attributed to

unfavorable solvent interactions at exposed positions of the applied monomeric model peptide.

In conclusion, DfeGly, 5³-F₃Ile, and 5³,5^{3'}-F₆Leu possess distinctly different hydrophobic and spatial properties that might enable the investigation of the influence of CH- π interactions on the structural stability of proteins. Thus, the partial charges for the through fluorine polarized hydrogen atoms (Figure 5.26) will be calculated and compared to their nonfluorinated analogues to elucidate to what extent the respective polarity is increased by the introduction of fluorine. Moreover, the three fluorinated C31 analogues containing DfeGly, 5³-F₃Ile, and 5³,5^{3'}-F₆Leu at the substitution site proposed here, I16, are currently being generated.²¹⁷ To determine the parameters underlying their assembly with T21 (i.e. binding affinity, enthalpy, entropy, as well as kinetics) isothermal titration calorimetry (ITC) and surface plasmon resonance (SPR) experiments will be carried out in future studies.

6 Summary and Outlook

The detailed and systematic characterization of the noncovalent interactions of nonnatural amino acids in native protein environments provides valuable insights for their specific application in peptide and protein engineering. The aim of this study was to apply phage display to identify a suitable helical environment for peptide sequences that contain either fluoroalkyl substituted amino acids, or an alternating set of β - and γ -amino acids. Subsequently, the observations made in the course of these investigations, coupled with studies published by other research groups in the field of HIV infection, were applied to design a candidate protease resistant fluorinated peptide-based HIV entry inhibitor that targets the envelope protein subunit gp41. The following three sections individually summarize the various studies conducted in the framework of this thesis.

Accommodating fluorinated amino acids in helical peptide environments

To identify preferred interaction partners of fluorinated aminobutyric acid analogues (MfeGly, DfeGly, TfeGly, and DfpGly), a previously established phage display screening system was utilized. In this system, the identification of specific binding partners is made possible by using coiled-coil pairing selectivity of the parallel heterodimeric VPK/VPE model assembly as a readout. In initial studies a central *a*-position in one strand of this system (VPK) served as a substitution site. The three directly interacting hydrophobic core positions of the complementary VPE strand were randomized in a phage displayed library. When incorporated in an *a*-position the through fluorine substitution polarized β -methylene groups point away from the hydrophobic core. Whereas, when incorporated in a central *d*-position, the polarized β -carbons project directly into the hydrophobic core. Based on this insight, a second VPE library was engineered, in order to enable the specific identification of suitable microenvironments for fluoromodified amino acids in either of both hydrophobic core positions individually.

Results show that, independent of the respective hydrophobic microenvironment, and despite all differences in hydrophobicity and size, the incorporation of these fluorinated aminobutyric acids into central *a*- or *d*-positions of a parallel heterodimeric coiled coil leads to similar pairing characteristics. All fluorinated amino acids studied here, in principle, prefer to interact with aliphatic amino acids. Considering that previous studies showed that fluorine induced polarity has a stronger effect at *d*-positions, while at *a*-positions side-chain volume is the prevailing factor that determines stability, these findings are somewhat surprising. On the other hand, it is conceivable that the applied system directs the outcome of this experiment, and that, within the hydrophobic core, aliphatic amino acids will be preferably selected regardless of the precise nature of the physicochemical properties of the fluorine analogues. A complementary approach would be the incorporation of these nonnatural building blocks into the charged interaction domain of a coiled coil, which would provide information about

specific interaction profiles in a solvent exposed environment. The parallel heterodimerization of the VPK/VPE model system would be suitable for such studies, since position dependent investigations (*e* vs *g*-position) are feasible.

The results highlight that coiled coils readily accommodate fluorinated amino acids as nonnatural building blocks in their hydrophobic cores, suggesting that such substitutions could be well tolerated in the hydrophobic cores of numerous biologically interesting domains. However, the selection outcome also indicates that the incorporation of fluorinated amino acids into natural protein folds requires subtle rearrangements on the sequence level (Ile vs. Leu) to optimize the packing of their immediate environment. Thus, improved packing, and likely shielding from solvent waters, leads to the observed stability enhancements.

Stabilizing a coiled coil that contains a set of alternating β - and γ -amino acids

The second part of this study was a broad survey for peptides that specifically match an $\alpha\beta\gamma$ -chimera in a coiled-coil assembly. Here, the tetrameric acid-pp/B3 β 2 γ coiled-coil system, which harbors a pentad of alternating β - and γ -amino acids served as a model. It was previously shown that the substitution of a central heptad with the pentad of backbone extended amino acids within the parental Acid-pp/Base-pp system is accompanied by a substantial loss in thermal stability.⁸⁴

Aiming to stabilize the chimeric system, an Acid-pp phage displayed library was generated in line with the current study, in which the four central heptad positions that directly interact with the $\beta\gamma$ -segment of B3 β 2 γ were randomized. In a first screen B3 β 2 γ was immobilized on streptavidin coated magnetic particles via an N-terminally attached biotin, while the Acid-pp library was fused to the N-terminus of a bacteriophage coat protein. In this experimental phage display setup, two highly similar Acid-pp variants comprising a cysteine and a phenylalanine were selected. Both peptides were shown to assemble with the chimeric sequence with substantially increased thermal stabilities than the original Acid-pp parent peptide, while a similar oligomerization state is maintained. Mutational analysis in combination with MD simulations revealed that the cysteine residue significantly improves the core-packing of the chimeric coiled coil through the formation of an interhelical H-bond with a non H-bonded backbone carbonyl of the $\alpha\beta\gamma$ -chimera. Moreover, it was shown that the N-terminal immobilization of B3 β 2 γ probably led to the selection of parallel coiled coils with the Acid-pp library.

Therefore, a second selection with a C-terminally immobilized B3 β 2 γ target was carried out. Here, four sequences evolved from phage display experiments. In two of these peptides the hydrophobic core was extended to positions that are usually designated for charged residues. It was shown that these peptides assemble with B3 β 2 γ under retention of the parental tetrameric oligomerization state, while the thermal stabilities of the resulting coiled

coils are comparable or even higher. The remaining two sequences harbor charged amino acids in positions that were originally designated for hydrophobic residues. Both tend to form dimers with the chimeric system. Here, the thermal stabilities were shown to be substantially lower than that of the tetrameric acid-pp/B3 β 2 γ chimera. However, these results indicate that the generation of dimeric heteroassemblies involving $\alpha\beta\gamma$ -chimeras is feasible if charged amino acids are introduced at the helical interface, an endeavor that had previously been unsuccessful.

These studies may pave the way for the generation of foldameric sequences, completely composed of $\beta\gamma$ -sequences ($\beta\gamma$ -hybrid peptides), that are able to form coiled-coil assemblies with α -peptides. For example, a $\beta\gamma$ -hybrid could be generated purely composed of a repeat of B3 β 2 γ 's stretch of alternating β and γ -amino acids (pentad repeat). As the complimentary strand, an α -peptidic heptad repeat comprising the ICEF recognition motif that was selected here could be employed. It is conceivable that this leads to an ideal geometric arrangement of hydrophobic core atoms, which in turn could enable the assembly of the $\beta\gamma$ -sequence with an α -peptide.

Towards protease stable fluorinated HIV entry inhibitors

The third part of this work focused on the core structure of HIV-1 gp41, which constitutes a stable six helix bundle folded by its trimeric N- and C-terminal heptad repeats (NHR and CHR). Inhibiting bundle formation with NHR or CHR derivatives blocks viral infection of host cells. One approach of the current investigations was to overcome the entropic penalty during the binding of unstructured C-peptide analogues. Applying rational design, the helical structure of three CHR derived peptides (9, 16, and 23 residues in length) that comprise the sequence of the pocket binding domain (PBD) as well as its N-terminal region has been stabilized by means of covalent backbone linkage and ion pairs. Even though the adequate placing of these modifications substantially enhanced α -helical content, CD experiments suggest that the investigated peptides do not appreciably bind to NHR derived target sequences. Two additional peptides solely composed of the wild type sequence, denoted C31 and C33, comprising 8 or 10 additional C-terminal residues, respectively, were generated. Both peptides were confirmed to undergo six helix bundle formation, while thermal stability results indicate an equally strong interaction with a NHR peptide. Hence, C31 was chosen to serve as a starting point for systematic substitution studies.

Being located in the middle of C31, the Ile residue of the highly conserved Trp-Trp-Ile motif which constitutes the PBD serves as the substitution site for fluorinated amino acids in ongoing studies. DfeGly, 5³-F₃Ile, and 5³,5^{3'}-F₆Leu will be incorporated to probe its involvement in a CH- π interaction with one of the Trp residues of the PBD. Preliminary assays were carried out to investigate the hydrophobicity and α -helix propensity of 5³-F₃Ile and 5³,5^{3'}-F₆Leu. Gained results verify that the herein investigated fluoroalkyl side chains

possess a unique combination of hydrophobicity and polarity. While the substitution with fluorine locally increases the solvent accessible surface area and thus hydrophobicity, the great electronegativity of fluorine draws electron density away from the adjacent hydrogen atoms. The α -helix propensity of 5^3 -F₃Ile is only half of that of Ile, while $5^3,5^{3'}$ -F₆Leu was previously shown to lack any propensity to form helices.¹¹² The dramatic decrease in α -helix propensity might be attributed to the burial of fluorocarbon side chains in the unfolded state of the monomeric model peptide in which they were investigated. The exposure of this highly hydrophobic side chain in the helical state is likely to be enthalpically unfavorable.

In future studies, these relatively conservative substitutions will provide information about the unique physicochemical properties of fluorine modified side chains. These studies might also elucidate to what extent the reduced α -helix propensities of fluorinated amino acids affect protein stability at a position within gp41's hydrophobic pocket.

7 Experimental Procedures and Analytical Methods

7.1 Peptide Synthesis and Characterization

The development of peptide based drugs and screening systems for the thorough investigation of physicochemical properties of unnatural peptide building blocks requires reliable methods for fast and efficient peptide synthesis.

7.1.1 Solid Phase Peptide Synthesis

The synthesis of long peptide sequences in solution is generally limited by a loss of yield during each coupling step. This problem has been overcome by the development of a method for solid phase peptide synthesis (SPPS) by R. B. Merrifield.²²² The principle of this method is rather simple. A growing peptide sequence is immobilized on solid, insoluble polymeric resins by coupling their C-terminal end via cleavable linkers to reactive sites of the solid support. Unlike ribosomal protein synthesis, SPPS proceeds in a C-terminal to N-terminal fashion. The N-terminus as well as eventual side chain functionalities of the incoming amino acid are protected while coupling proceeds. The peptide remains immobilized throughout the entire synthesis allowing for fast and efficient removal of liquid reagents and by-products in several washing steps.

When applying the orthogonal protecting group Fmoc/^tBu-strategy the α -amino group is reversibly blocked by Fmoc (9-Fluorenylmethoxycarbonyl), which is eliminated in the alkaline media of piperidine. *Tert*-Butyl-type groups are used for the semipermanent protection of the side chains of trifunctional amino acids (e.g. asparagine and glutamate as *tert*-butyl ester; serine, tyrosine, and threonine as *tert*-butyl ethers and lysine with Boc). These protecting groups are cleaved under rather mild conditions by trifluoroacetic acid, and the peptide can be discharged from the linker simultaneously. Generally, an increased positive partial charge of the carbonyl carbon atom is generated via the activation of the C-terminus prior to amino acid coupling. This leads to a greater reactivity towards the nucleophilic attack of the α -amino group and thus increases coupling efficiency. The description "orthogonal" implies that the acid labile side chain protecting groups are stable under conditions of Fmoc cleavage. This strategy has been applied to the synthesis of the peptides described in the current thesis (Table 7.1).

The following peptides were synthesized by former and current group members, Dr. Toni Vagt, Dr. Mario Salwiczek, and MSc Ulla Gerling, and were kindly provided to the author of this thesis:

| | | | | |
|------------------------------------|------------------------------------|------------------------------------|------------------------------------|-----------------|
| VPK-DfeGly₁₆ | Bio-VPK-TfeGly₁₆ | VPK-DfeGly₁₉ | Bio-VPK-TfeGly₁₉ | Bio-VPK |
| Bio-VPK-DfeGly₁₆ | VPK-DfpGly₁₆ | Bio-VPK-DfeGly₁₉ | VPK-DfpGly₁₉ | K-DfeGly |
| VPK-TfeGly₁₆ | Bio-VPK-DfpGly₁₆ | VPK-TfeGly₁₉ | Bio-VPK-DfpGly₁₉ | K-Ile |

T21 was purchased from Peptide Specialty Laboratories GmbH.

Table 7.1: Sequences of peptides generated in the course of this study:

| Peptide/ Chimera | Sequence |
|--------------------------------------------------|---------------------------------------------------------------------------------------------------------------------|
| VPE | Abz-EVSALEKEVASLEKEVSALEKKVASLKKEVSALE-OH |
| VPE-L ₁₆ L ₂₃ ¹ | Abz-EVSALEKEVASLEKE L SALEKK L ASLKKEVSALE-OH |
| VPK | Abz-KVSALKEKVASLKEKVSALKEEVASLEEKVSALK-OH |
| VPK-MfeGly ₁₆ | Abz-KVSALKEKVASLKEK MfeGly SALKEEVASLEEKVSALK-OH |
| Bio-VPK-MfeGly ₁₆ | Biotin-GSGKVSALKEKVASLKEK MfeGly SALKEEVASLEEKVSALK-OH |
| Bio-VPK-Abu ₁₉ | Biotin-GSGKVSALKEKVASLKEKVS A AbuKEEVASLEEKVSALK-OH |
| VPK-Abu ₁₉ | Abz-KVSALKEKVASLKEKVS A AbuKEEVASLEEKVSALK-OH |
| VPK-MfeGly ₁₉ | Abz-KVSALKEKVASLKEKVS A MfeGly KEEVASLEEKVSALK-OH |
| Bio-VPK-MfeGly ₁₉ | Biotin-GSGKVSALKEKVASLKEKVS A MfeGly KEEVASLEEKVSALK-OH |
| Acid-pp | Abz-LSALEKELASLEKELSALEKELASLEKELSALEKE-OH |
| Acid-pp-CAbz | NH ₂ -LSALEKELASLEKELSALEKELASLEKELSALE K (Abz)E-OH |
| Acid-pp-CFLE ² | Abz-LSALEKELASLEKE CSAFLKE LASLEKELSALEKE-OH |
| Acid-pp-ICEF ² | Abz-LSALEKELASLEKE ISA CEKFL ASLEKELSALEKE-OH |
| Acid-pp-ICEF-CAbz | H ₂ N-LSALEKELASLEKE ISA CEKFL ASLEKELSALE K (Abz)E-OH |
| Acid-pp-IAbuEF | Abz-LSALEKELASLEKE ISA AbuEKFL ASLEKELSALEKE-OH |
| Acid-pp-ISEF | Abz-LSALEKELASLEKE ISA SEKFL ASLEKELSALEKE-OH |
| Acid-pp-LLLL | Abz-LSALEKELASLEKE LSALLKLL ASLEKELSALEKE-OH |
| Acid-pp-LFYL | Abz-LSALEKELASLEKE LSAFYKLL ASLEKELSALEKE-OH |
| Acid-pp-MTER | Abz-LSALEKELASLEKE MSATEKRL ASLEKELSALEKE-OH |
| Acid-pp-HCAN | Abz-LSALEKELASLEKE HSACAKNL ASLEKELSALEKE-OH |
| Base-pp | Abz-LSALKEKLASLKEKLSALKEKLASLKEKLSALKEK-OH |
| B3β2γ | Abz-LSALKEKLASLKEK βLyDβLyKβK LASLKEKLSALKEK-OH |
| B3β2γ-GSG-NBio ² | Biotin-GSGLSALKEKLASLKEK βLyDβLyKβK LASLKEKLSALKEK-OH |
| B3β2γ Variant1-GSG-NBio ² | Biotin-GSGLSALKEKLASLKEK βhAyDβhAyKβK LASLKEKLSALKEK-OH |
| B3β2γ-NY(NO ₂) | H ₂ N- Y(NO₂) -LSALKEKLASLKEK βLyDβLyKβK LASLKEKLSALKEK-OH |
| B3β2γ-GSG-CBio ³ | H ₂ N-LSALKEKLASLKEK βLyDβLyKβK LASLKEKLSALKEK GSGK (Bio)-OH |
| B3β2γ Variant1-GSG-CBio ³ | H ₂ N-LSALKEKLASLKEK βhAyDβhAyKβK LASLKEKLSALKEK GSGK (Bio)-OH |
| N36 | Ac-SGIVQQQNNLLRAIEAQQHLLQLTVWGIKQLQARIL-NH ₂ |
| C34 | Ac-WMEWDREINNYTSLIHSLIEESQNQQEKNEQELL-NH ₂ |
| C33 | Ac-EQIWNNMTWMEWDREINNYTSLIHSLIEESQNQ-NH ₂ |
| C31 | Ac-EQIWNNMTWMEWDREINNYTSLIHSLIEESQ-NH ₂ |
| C23 | Ac-EQIWNNMTWMEWDREINNYTSLI-NH ₂ |
| C9 _{St} | Ac-EW S₅ EW D S₅KI -NH ₂ |
| C16 _{St} | Ac-EQ S₅ENNS₅KWS₅EW S₅KI -NH ₂ |
| C23 _{St} | Ac-EQ S₅ENNS₅KWEEWD KKIS₅EYT S₅KI -NH ₂ |
| K- 5-F ₃ Ile ⁴ | Ac-YGGKAAAAKA 5-F₃Ile AAAAAAK-NH ₂ |

¹ Synthesized by Verena Jeschke under the supervision of the author of thesis.² Synthesized by Sebastian Wiczorek under the supervision of the author of this thesis.³ Synthesized by Vanessa Radtke under the supervision of the author of this thesis.⁴ Synthesized by Susanne Huhmann under the supervision of the author of this thesis.

Chemicals and Solvents

The following chemicals were used as purchased: Boc-**Abz**-OH (Bachem), acetonitrile (**ACN**, HPLC gradient grade, Acros), D(+)-**biotin**, 98% (Acros Organics), 1,8-Diazabicyclo[5.4.0]undec-7-ene (**DBU**, Merck), dichloromethane (**DCM**, Fischer Scientific), 1,2-dichloroethane (**DCE**, extra dry, 99.8%, Fischer Scientific), diethylether (analytical grade, Fischer Scientific), *N,N*-diisopropyl-carbodiimide (**DIC**, 99%, Acros Organics), *N,N*-diisopropylethylamine (**DIPEA**, 98%, Acros Organics), dimethylformamide (**DMF**, p.a., Acros Organics), 1,2-ethanedithiol (**EDT**, 95 %, Acros Organics), guanidine hydrochloride (**GndHCl**, 99.5%, Acros Organics), 1-hydroxy-7-azabenzotriazole (**HOAt**, Carbolution), 1-hydroxy-benzotriazole (**HOBt**, Fa. Gerhardt), *N*-methyl-2-pyrrolidinone (**NMP**, Acros Organics), phenol (99+%, Acros Organics), piperidine (99% extra pure, Acros Organics), sodium dihydrogenphosphate dihydrate (Fluka), sodium perchlorate (Acros Organics), 2-(1H-benzotriazol-1-yl)-1,1,3,3-tetramethyluronium tetrafluoroborate (**TBTU**, Fa. Gerhardt), trifluoro acetic acid (**TFA**, p.a. 100%, Roth), trifluoro acetic acid (**TFA**, Uvasol, Merck), thioanisole (99%, Acros Organics), triisopropylsilane (**TIS** 99%, Acros Organics), trimethylsilyl bromide (**TMSBr**, Fluka). Water was purified on a MilliPore device (MilliQ-Advantage A10 Millipore), and acetic anhydride (99%, Acros) was distilled prior to use. All remaining solvents, chemicals, and reagents are listed in the relevant section in which they were used.

Amino acids

Fmoc-protected building blocks of canonical α -amino acid (i.e.: Fmoc-Asp(OtBu)-OH, Fmoc-Gln(Trt)-OH, Fmoc-Gly-OH, Fmoc-Ile-OH, Fmoc-L-Ala-OH·H₂O, Fmoc-L-Arg(Pbf)-OH, Fmoc-L-Asn(Trt)-OH, Fmoc-L-Glu(OtBu)-OH, Fmoc-L-His(Trt)-OH, Fmoc-L-Leu-OH, Fmoc-L-Met-OH, Fmoc-L-Tyr(tBu)-OH, Fmoc-Lys(Boc)-OH, Fmoc-Ser(tBu)-OH, Fmoc-Thr(tBu)-OH, Fmoc-Trp(Boc)-OH) were purchased from Orpegen Pharma. Fmoc-Nle-OH was purchased from Novabiochem. Fmoc-protected (*S*)- β 3-Homolysine, (*S*)- β 3-Homoleucine, and (*S*)- β 3-Homoalanine were purchased from Fluka, and (*R*)- γ 4-Homoaspartic acid and (*S*)- γ 4-Homolysine from RareChemicals. Fmoc-3-Nitrotyrosine was purchased from Bachem. (2*S*)-amino heptaonic acid was purchased from Sigma Aldrich. (*S*)-*N*-Fmoc- α -pentenylalanine (S5-Fmoc) was purchased from Nagase's Unnatural Amino Acids. Since this building block was received as a hard to handle oily substance, it was divided into 0.1 mmol aliquots after dilution to a volume of 10 mL with DCM and stored at -20°C.

The fluorinated amino acid building block (*S*)-4-monofluoroethylglycine (MfeGly) was synthesized in the group of Prof. Haufe, Fmoc protected by Ulla Gerling in the group of Prof. Kokschi, and kindly provided to the author of this thesis. The fluorinated amino acid building blocks (2*S*,3*S*)-5,5,5-trifluoroisoleucine (5³-F₃-Ile) and (2*S*)-5,5,5,5',5',5'-hexafluoroisoleucine (5³,5'³-F₆-Leu) were synthesized in the group of Prof. Czekelius (H. Erdbrink and C. Czekelius, respectively), and kindly provided to the author of this thesis. Fmoc protection

of (2S,3S)-5,5,5-trifluoroisoleucine, (2S)-5,5,5',5',5'-hexafluoroisoleucine, and (2S)-amino heptanoic acid was carried out by Susanne Huhmann in the course of the experimental work on her MSc thesis under the laboratory supervision of the author of the current study.²¹⁷

Automated Synthesis

The synthesis of all model peptides as well as the α -peptide fragments of B3 β 2 γ was performed via the Fmoc standard protocol given in Table 7.2 on a fully automatic peptide synthesizer *SyroXP* manufactured by Multi-SynTech GmbH. Fmoc-Glu(OtBu)-NovaSyn[®]-TGA-resin (0.21 mmol/g loading) and Fmoc-Lys(Boc)-NovaSyn[®]-TGA-resin (0.20 mmol/g loading) were purchased from Merck. Synthesis was performed at a 0.05 mmol scale in disposable syringes with polytetrafluoroethylene (PTFE) frits.

Before coupling a new amino acid to the resin/peptide, the temporary Fmoc-protecting group was removed with an equimolar mixture of the bases piperidine (PIP) and DBU. Activation of the new amino acid's C-terminal carboxyl group was carried out with TBTU and HOBt in a base catalyzed reaction using DIPEA. The coupling mixture contained 0.23 M NaClO₄ to prevent on-resin aggregation.

Table 7.2: Coupling procedure in Fmoc standard protocol applied in *SyroXP* synthesizer.

| Process | Reagent | Time |
|-------------------|------------------------------------------------------------------------------------------|------------------|
| Fmoc-Deprotection | 2 mL 2 % PIP / 2 % DBU in DMF 2 mL 2 % PIP / 2 % DBU in DMF | 10 min 10 min |
| Washing | 2.5 mL DMF | 6 × 1 min |
| Coupling | 4 eq. Fmoc-Amino acid-OH 4 eq. HOBt in DMF 4 eq. TBTU in DMF 4 eq. DIPEA in NMP | 30 min |
| Washing | 2.5 mL DMF | 6 × 1 min |
| Coupling | 4 eq. Fmoc-Amino acid-OH 4 eq. HOBt in DMF 4 eq. TBTU in DMF 4 eq. DIPEA in NMP | 30 min |

The synthesis of C34 and N36 was performed via the protocol given in Table 7.3 on the fully automatic peptide synthesizer *Activo-P11* manufactured by *ActivoTec*. NovaSyn[®]-TGR-resin (0.21 mmol/g loading) was purchased from Merck and loaded with the corresponding C-terminal amino acid by means of DIC/HOAt activation (1:1 in DMF) in a 5-fold excess prior to synthesis. Quantitative loading was verified according to procedures described by Gude *et al.*²²³ Synthesis was performed at a 0.05 mmol scale in disposable syringes with polytetrafluoroethylene (PTFE) frits.

Table 7.3: Coupling procedure in Fmoc standard protocol applied in Activo-P11 synthesizer.

| Process | Reagent | Time |
|-------------------|----------------------------------------------------------------------|-------------------------|
| Fmoc-Deprotection | 2 mL 2 % PIP / 2 % DBU in DMF 2 mL 2 % PIP / 2 % DBU in DMF | 2 × 5 min 2 × 10 min |
| Washing | 2.5 mL DMF | 5 × 1 min |
| Coupling | 10 eq. Fmoc-Amino acid-OH 10 eq. HOBt in DMF 10 eq. DIC in DMF | 60 min |
| Washing | 2.5 mL DMF | 5 × 1 min |
| Coupling | 5 eq. Fmoc-Amino acid-OH 5 eq. HOAt in DMF 5 eq. DIC in DMF | 60 min |

Manual Synthesis

- The synthesis of C33, C31, C23, C9_{St}, C16_{St}, and C23_{St} was carried out manually in a double coupling procedure. NovaSyn[®]-TGR-resin (0.21 mmol/g loading) was purchased from Merck and loaded with the according C-terminal amino acid by means of DIC/HOAt activation (1:1 in DMF) in a 5-fold excess prior to the synthesis. Quantitative loading was verified according to procedures described by Gude *et al.*²²³ Synthesis was performed at a 0.05 mmol scale in disposable syringes with polytetrafluoroethylene (PTFE) frits. Canonical amino acid building blocks were activated by means of DIC/HOAt 1:1 in DMF. The molar excess of amino acid and coupling reagents was 10-fold during the first and 5-fold during the second coupling. Fmoc-S5 amino acid building blocks were also activated by means of DIC/HOAt 1:1 in DMF. Here, single coupling was carried out, and the molar excess was reduced to 2 equivalents while the reaction time was extended to 12 hours. Complete coupling was verified with a negative Kaiser test.²²⁴ The subsequent amino acid was coupled manually in a double coupling process and activated by means of DIC/HOAt 1:1 in DMF. The molar excess of amino acid and coupling reagents was increased to 10 fold for each coupling step, while the first coupling time was 6h and the second 12h.
- Fluorinated amino acids as well as the subsequent amino acids were incorporated into respective model peptides manually in a double coupling process and activated by means of DIC/HOAt 1:1 in DMF. The molar excess of amino acid and coupling reagents was reduced for fluorine-containing residues to 1.5-fold for the first and 0.8-fold for the second coupling. These couplings were performed with varying reaction times (4-8h) until completion was indicated by a negative Kaiser test.²²⁴
- The manual coupling of β - and γ -amino acids within $\alpha\beta\gamma$ -chimeric peptides was carried out in a single coupling step by means of 1:1 DIC/HOAt activation in DMF. The molar excess of corresponding Fmoc- β - or γ -amino acid-OH was 1.5 fold and

the reaction time was 6h. Complete coupling was verified with a negative Kaiser test.²²⁴ Only if necessary, a second coupling with 0.5 equivalents of the backbone extended amino acid was performed overnight.

- The Abz-label for concentration determination via UV-spectroscopy was coupled manually by means of 1:1 DIC/HOBt activation in DMF in a double coupling procedure with 5 eq. excess each (2×1 h). The Biotin tag for streptavidin binding was coupled in a double coupling procedure with 4 eq. excess of D(+)-Biotin and 1:1 DIC/HOAt activation in NMP. The reaction time was 1h for each coupling step. Complete coupling was verified with a negative Kaiser test.²²⁴
- Also the GSG-linker between peptides synthesized by automated SPPS and the biotin tag was coupled manually according to double coupling procedure by means of 1:1 DIC/HOBt activation in DMF. The molar excess of amino acid and coupling reagents was 10-fold for each coupling step and the reaction time was 2×1 h.

Acetylation

To avoid failures in the sequence, potentially free amino groups were capped by acetylation with 10% acetic anhydride and 10% DIPEA in DMF (3×10min) subsequent to resin loading, subsequent to coupling of a nonnatural building block and the residue immediately downstream. Free N-termini of all gp41 C- and N-peptide variants were acetylated in the same manner post final Fmoc deprotection.

Ring-Closing Metathesis

Ring-closing metathesis was performed with resin-bound, N-terminally acetylated, and side chain protected peptides using 20 mol% of Grubbs I catalyst. Therefore, the previously dried resin was swollen in extra dry DCE for 15 min while being degassed with argon. The adequate amount of catalyst was dissolved in extra dry DCE and degassed for 15 min prior to its addition to the peptide. The reaction was carried out in disposable syringes with PTFE frits under avoidance of oxygen (argon filled balloon attached to syringe). After a first reaction of 6 h at room temperature, the resin was treated with fresh catalyst for additional 12 h. The reaction was monitored by analytical HPLC and ESI-ToF-Mass Spectrometry after cleavage of the peptides from a resin aliquot (*vide infra*).

Cleavage from Solid Support

To cleave the synthesized peptide from the support, the resins were treated between 2 and 4.5 hours at room temperature with a general cleavage cocktail of 89 % TFA, 10 % TIS and 1 % H₂O (v/v/v). These strong acidic conditions also remove permanent side-chain protecting groups (Boc, tBu, Trt, Pbf). Subsequently, TFA was removed under argon gas and the peptides were precipitated in ice-cold diethyl ether. The crude peptide material was collected via centrifugation and dried by exposure to air. Peptides carrying methionine or the biotin tag

for phage display experiments were cleaved by adding 89 % TFA, 5 % TIS, 5 % EDT and 1 % H₂O (v/v/v/v) to prevent oxidation of thiol groups. Reduction of yet occurred S-oxides was carried out with TMSBr according to procedures described by Beck *et al.*²²⁵ All peptides related to gp41 were cleaved by adding 81.5 % TFA, 1 % TIS, 5 % H₂O, 2.5 % EDT, 5 % thioanisole, and 5 % phenol (v/v/v/v/v).

7.1.2 Purification, Identification, and Characterization

Chromatography is a commonly applied technique used for the separation of mixtures. Here, mixtures are dissolved in a fluid called the mobile phase, which carries it through a structure holding another material referred to as stationary phase. Subtle differences in the partition coefficient of compounds lead to differential retention on the stationary phase and thus cause their separation. In the here described studies, analytical as well as preparative reversed-phase high-performance liquid chromatography (RP-HPLC) was applied for purity determination and purification purposes, respectively. Columns filled with silica resins that possess covalently bonded alkyl chains (C8 and C18) served as the hydrophobic stationary phase, while a polar (aqueous) mobile phase was employed. Peptides were eluted from the column by decreasing the polarity of the mobile phase with acetonitrile. The exact mass of eluted compounds was determined with electrospray ionization time-of-flight mass spectrometry (ESI-ToF-MS) which separates ionized compounds according to their mass (m) - to - charge (z) ratios (m/z). The concentration of a given peptide solution was determined with UV absorbance spectroscopy, as the absorbance is a linear function of the molar concentration.

Analytical HPLC

The purity of intermediate and final products was verified by analytical RP-HPLC. Analytical HPLC was performed on either a LaChrom Elite HPLC (Hitachi Europe Ltd. Berkshire, SL6 8YA, GB), with two L-2130 HTA Pumps, L-2200 Autosampler, L-2455 Diode Array Detector, L-2485 Fluorescence Detector, or a LaChrom-HPLC-system containing an Interface L-7000, two pumps L-7100, diode array detector L-7450, and an autosampler L-7200. A LUNATM C8(2) column (10 µm particle size, 250 × 4.60 mm inner diameter, phenomenex®) or a LUNATM C18 column (5 µm particle size, 250 × 4.60 mm inner diameter, phenomenex®) was used. Solvent A was H₂O containing 0.1 % TFA, solvent B ACN (Acetonitrile) containing 0.1 % TFA. The flow rate was 1 mL/min, and the absorbance between 200 nm and 450 nm was recorded. To enable efficient separation, linear gradients and columns were optimized for each sample individually.

Preparative HPLC

Post determination of an appropriate linear solvent gradient and column for efficient separation by analytical HPLC, purifications of crude peptides were carried out with a RP-

HPLC system from Knauer GmbH (D-14167, Berlin). This system included a *Smartline Manager 5000* with interface module, two *Smartline Pump 1000*, a *UV Detector 2500*, a 6-port-3-channel injection valve with a 5 ml sample loop. Separation was either carried out on a *LUNA™ C8(2)* column (10 µm particle size, 250 mm × 21.2 mm inner diameter, phenomenex®) or on a *Gemini-NX C18 column* (10 µm particle size, 250 mm × 21.2 mm inner diameter, phenomenex®). A flow rate of 20 mL/min was used, and the absorbance was recorded at 320 nm for Abz containing peptides, at 280 nm for peptides containing aromatic amino acid residues, and at 230 nm for the remaining peptides (peptide bond absorbance).

ESI-ToF-Mass Spectrometry

Intermediate and final products were identified by molecular mass measurement (Table 7.4) using a 6210 Time-of-flight LC/MS with ESI-injector from Agilent Technologies, Inc.. Therefore, the compounds were dissolved in a mixture of acetonitrile and water containing 0.1% TFA. Mass spectra were analyzed using *Agilent MassHunter* Data Analysis software.

Table 7.4: Identification of peptides and chimeras by ESI-TOF mass spectrometry. (Unless stated otherwise, all peptides bear an N terminal Abz label for precise concentration determination.)

| Peptide/ Chimera | Charge | Molecular Weight Calculated [g/mol] | Molecular Weight Measured [g/mol] |
|--------------------------------------------------|----------------------|-------------------------------------|-----------------------------------|
| VPE | [M+4H] ⁴⁺ | 948.27 | 948.27 |
| VPE-L ₁₆ L ₂₃ ¹ | [M+4H] ⁴⁺ | 955.27 | 955.79 |
| VPK | [M+4H] ⁴⁺ | 947.80 | 947.80 |
| VPK-MfeGly ₁₆ | [M+4H] ⁴⁺ | 948.30 | 948.42 |
| Bio-VPK-MfeGly ₁₆ | [M+4H] ⁴⁺ | 978.53 | 978.51 |
| Bio-VPK-Abu ₁₉ | [M+4H] ⁴⁺ | 1017.55 | 1018.32 |
| VPK-Abu ₁₉ | [M+4H] ⁴⁺ | 940.79 | 940.79 |
| VPK-MfeGly ₁₉ | [M+4H] ⁴⁺ | 945.29 | 945.35 |
| Bio-VPK-MfeGly ₁₉ | [M+4H] ⁴⁺ | 1022.31 | 1022.32 |
| Acid-pp | [M+4H] ⁴⁺ | 998.29 | 998.29 |
| Acid-pp-CAbz | [M+4H] ⁴⁺ | 998.29 | 998.29 |
| Acid-pp-CFLE ² | [M+4H] ⁴⁺ | 1000.28 | 1000.27 |
| Acid-pp-ICEF ² | [M+4H] ⁴⁺ | 1000.28 | 1000.77 |
| Acid-pp-ICEF-CAbz | [M+4H] ⁴⁺ | 1000.28 | 1000.28 |
| Acid-pp-IAbuEF | [M+4H] ⁴⁺ | 995.77 | 995.78 |
| Acid-pp-ISEF | [M+4H] ⁴⁺ | 996.28 | 996.28 |
| Acid-pp-LLLL | [M+4H] ⁴⁺ | 990.30 | 990.71 |
| Acid-pp-LFYL | [M+4H] ⁴⁺ | 1011.29 | 1011.80 |
| Acid-pp-MTER | [M+4H] ⁴⁺ | 1006.53 | 1006.94 |
| Acid-pp-HCAN | [M+4H] ⁴⁺ | 983.56 | 983.91 |
| Base-pp | [M+4H] ⁴⁺ | 997.10 | 997.11 |
| B3β2γ | [M+4H] ⁴⁺ | 978.59 | 979.11 |
| B3β2γ-GSG-NBio ² | [M+4H] ⁴⁺ | 1055.67 | 1055.64 |
| B3β2γ Variant1-GSG-NBio ² | [M+4H] ⁴⁺ | 1027.64 | 1027.62 |
| NH ₂ -Y(NO ₂)-G-OH | [M-1H] ¹⁻ | 282.04 | 282.07 |
| B3β2γ-NY(NO ₂) | [M+4H] ⁴⁺ | 1000.83 | 1000.87 |
| B3β2γ-GSG-CBio ³ | [M+4H] ⁴⁺ | 1091.37 | 1091.92 |

¹ Synthesized by Verena Jeschke under the supervision of the author of thesis.

² Synthesized by Sebastian Wieczorek under the supervision of the author of this thesis.

³ Synthesized by Vanessa Radtke under the supervision of the author of this thesis.

| | | | |
|------------------------------------------------------|----------------------|---------|---------|
| B3β2γ Variant1-GSG-CBio³ | [M+4H] ⁴⁺ | 1059.34 | 1059.89 |
| N36 | [M+4H] ⁴⁺ | 1041.35 | 1041.85 |
| C34 | [M+3H] ³⁺ | 1430.01 | 1430.03 |
| C33 | [M+3H] ³⁺ | 1398.31 | 1398.33 |
| C31 | [M+3H] ³⁺ | 1317.61 | 1317.63 |
| C23 | [M+2H] ²⁺ | 1514.19 | 1514.19 |
| C9_{St} | [M+2H] ²⁺ | 662.85 | 662.85 |
| C16_{St} | [M+2H] ²⁺ | 1081.06 | 1080.56 |
| C23_{St} | [M+2H] ²⁺ | 1017.86 | 1017.87 |
| K-5³-F₃-Ile¹ | [M+2H] ²⁺ | 872.52 | 872.03 |

Determination of Peptide Concentration

In the absence of any aromatic amino acid in the peptide sequence, concentrations of Abz or biotin harboring peptides were estimated by UV spectroscopy on a Cary 50 UV/Vis spectrometer (Varian) using the absorption of *o*-aminobenzoic acid ($\lambda_{\max} = 320$ nm at pH 7.4) or 3-nitrotyrosine ($\lambda_{\max} = 420$ nm at pH 7.4). Therefore, calibration curves were recorded using different concentrations of H₂N-Abz-Gly-COOH × HCl (Bachem) and H₂N-Tyr(NO₂)-Gly-COOH, respectively, in the buffer used for CD spectroscopy containing 6M guanidine hydrochloride.

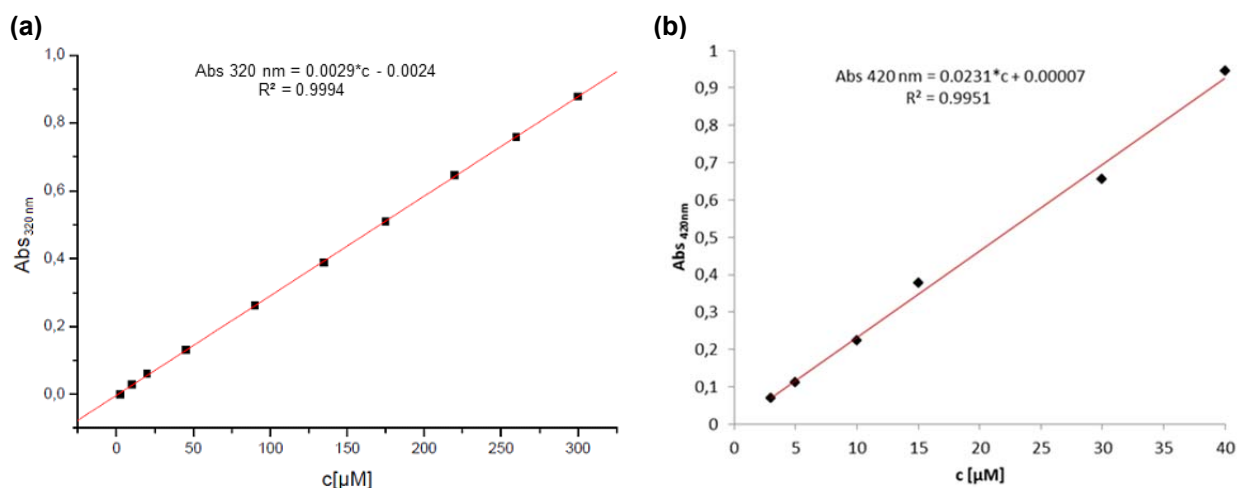


Figure 7.1: Calibration curves for the determination of concentrations of (a) Abz containing and (b) Nitrotyrosine containing peptides recorded at 20°C in the respective buffer

Concentrations of peptides containing none of the above mentioned UV labels but UV active amino acid residues were determined by UV spectroscopy in 6M guanidine hydrochloride. From the absorbance at 280 nm, the concentration of each stock solution was calculated according to the following equation:²²⁶

$$\epsilon(280 \text{ nm})[M-1\text{cm}^{-1}] = [(\#\text{Trp})(5500) + (\#\text{Tyr})(1490) + (\#\text{Cys})(125)]$$

For all UV measurements disposable Plastibrand® PMMA cuvettes (purchased from VWR) with 1 cm path lengths were used. The spectra were recorded at wavelengths from 500 nm

¹ Synthesized by Susanne Huhmann under the supervision of the author of this thesis.

to 300 nm with baseline correction through a blank sample containing only buffer without peptide.

Hydrophobicity Studies

N^α-Fmoc protected amino acids were dissolved in deionized water containing 40% ACN and 0.1% TFA. Their retention times were determined on a C18 column (Capcell PAK C18, 5 μm). A linear gradient from 40 to 70% ACN over 20 min was applied at room temperature. All experiments were performed in triplicate. The van der Waals volumes of the side chains were estimated from the β-carbon outward by summation of the atomic increments and bond contributions.²²⁷

7.2 Structural Analysis

7.2.1 Circular Dichroism Spectroscopy

Circular dichroism (CD) is an excellent tool for secondary structure determination of biomacromolecules.²²⁸ It allows not only for rapid determination of protein structure under a variety of experimental conditions, but also detailed monitoring of conformational transitions of the studied system. Both of these features were extensively explored for studies described in this thesis.

CD describes the phenomenon that occurs when two oppositely circularly polarized light beams are absorbed by an optically active molecule.²²⁹ Differential absorption of the left- and right circularly polarized components of plane polarized light results in an overall elliptical beam shape (Figure 7.2a). When the electric vectors of the two circular components point in the same direction, the major axis of such an ellipse corresponds to the sum of the magnitudes of these vectors. If the electric vectors point in opposite directions, the difference in their magnitudes gives the minor axis of the ellipse. CD data is commonly discussed in terms of ellipticity, θ , which is defined as the angle between the major and minor axes.

Proteins and peptides are composed of enantiomeric amino acids (except for glycine) which are linked via chromophoric peptide bonds. These two distinct features render them most suitable targets for CD spectroscopic analysis.²³⁰

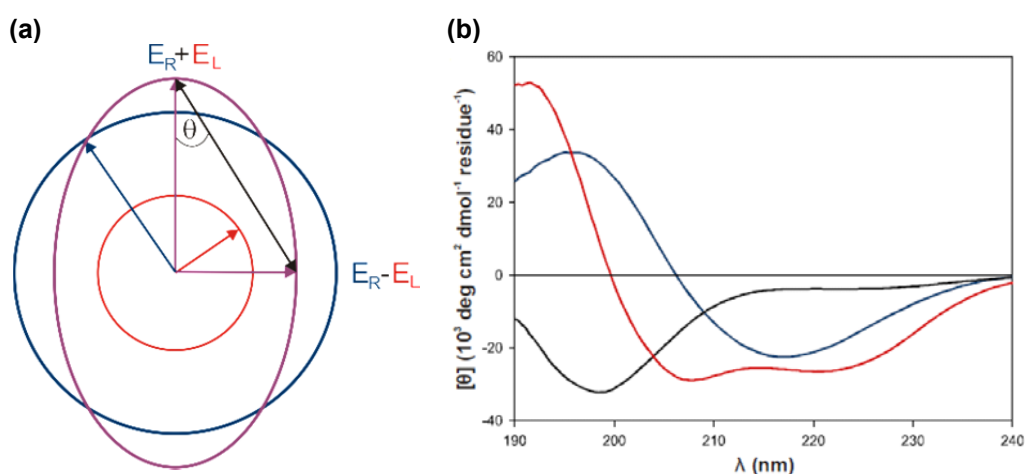


Figure 7.2: (a) elliptical polarized light (purple) is composed of unequal contributions of right (blue) and left (red) circularly polarized light. (b) Characteristic CD-spectra of a β -sheet (blue), an α -helix (red), and a random coil (black).

Peptide bonds display two characteristic absorption bands in the far-UV region (240 nm and below) of electromagnetic spectra: a weak but broad $n \rightarrow \pi^*$ transition around 220 nm and a more intense $\pi \rightarrow \pi^*$ transition around 190 nm. Every secondary structure found in proteins gives rise to a characteristic CD signature in far-UV (Figure 7.2b). The characteristics of β -sheet spectra are a broad negative band in the 215-220 nm region, resulting from $n \rightarrow \pi^*$ transition, and a positive band in the 190-200 nm region ($\pi \rightarrow \pi^*$ transition). Isolated α -helical structures are depicted in a spectrum with two specific minima at 222 nm ($n \rightarrow \pi^*$ transition) and 208 nm ($\pi \rightarrow \pi^*$), and a positive band at 193 nm ($\pi \rightarrow \pi^*$). Primarily unfolded structures

(random coil) are usually recognized by one minimum at approximately 200 nm ($n \rightarrow \pi^*$ transition).

CD Measurements

If not stated otherwise, CD-spectra were obtained in the far-UV range (190–240 nm) using 0.1 cm Quartz Suprasil cuvettes (Hellma) equipped with a stopper as recorded on a *Jasco J-715 spectropolarimeter* (*Jasco PTC-348WI peltier thermostat*). The nitrogen flow rate was set to 3 L/min. Individual concentration and buffer conditions are given with respective data. All spectra were background corrected by buffer spectrum subtraction. Moreover, all spectra were measured in triplicate to verify reproducibility and standard deviations. Acquired data was analyzed using different versions of Microsoft Excell (Microsoft) and Microcal Origin 6.0 (OriginLab). Ellipticity was normalized to concentration (c [mol/L]), number, n , of respective residues (including the Abz or biotin label) and path length ($l = 1$ cm) using Equation 7.1. Thus CD-data are given in terms of molar ellipticity, $[\theta]$, with the dimension [$10^3 \text{ deg cm}^2 \text{ d mol}^{-1} \text{ residue}^{-1}$].

$$[\theta] = \frac{\theta}{10000 \cdot c \cdot l \cdot n} \quad (7.1)$$

Helical Content Determination

Fractional helicity within a given protein structure was obtained from experimentally molar ellipticities at 222 nm, $[\theta]_{222}$. The basis on which this helical content is calculated was provided by *Chen et al.*, who demonstrated a chain length dependence of the $n\pi^*$ transition intensity according to the following equation:²³¹

$$[\theta]_{222nm} = [\theta]_{222nm}^{\infty} \cdot \left(1 - \frac{k}{n}\right) \quad (7.2)$$

$[\theta]_{222nm}$ is the normalized signal intensity at 222nm, $[\theta]_{222nm}^{\infty}$, the signal intensity of an infinite helix at this wavelength, and n the number of backbone amides. For a nearly 100% α -helical peptide such as $(\text{Lys})_n$ the signal intensity is ca. $-40,000 \text{ deg cm}^2 \text{ d mol}^{-1}$.²³² The values for k vary in the literature and depend on the applied method and polypeptide. Taking the four C-terminal carbonyls that cannot undergo hydrogen bonding ('helix fraying') into account, investigations by *Gans et al* with C- and N-terminally unprotected polypeptides revealed a value of 4.6 for k . The authors thus introduced the following equation for the determination of the α -helical fraction f_H of a given peptide.²³²

$$f_{helix} = \frac{[\theta]_{222nm}}{-40000 \cdot \left(\frac{4.6}{n}\right)} \quad (7.3)$$

Capping the ends of a polypeptide chain can reduce helix fraying and thus also the k value can be reduced to 2.5.^{71,233,234}

Thermal Denaturation Experiments

The transition from native fold to random coil and its dependence on temperature or denaturant concentration can be monitored by plotting CD signals as a function of the respective denaturing condition. This thermal stability can be either recorded by measuring the entire spectrum of a sample at different temperatures, or, alternatively, a single wavelength that represents a certain feature of a given secondary structure can be measured continuously with rising temperature. As such, the signal at 222 nm is most diagnostic for α -helical structures (cf Figure 7.2) and is commonly used to monitor their denaturation. The full reversibility of the unfolding (“melting”) process can be verified by measuring the same wavelength while decreasing the temperature back to the starting point or by starting a whole new denaturation after the sample is already cooled down. All spectra were measured in triplicate to verify reproducibility and standard deviations. If the results are reproducible, the sample is not aggregating or precipitating after unfolding and, thus, the melting temperature is directly related to conformational and thermodynamic stability. In a two-state transition from folded oligomeric coiled coil to unfolded monomeric species, the fraction of unfolded species can be expressed in terms of the molar ellipticity $[\theta]$. The temperature dependence of $[\theta]$ can be described as the sum of the temperature dependence of the ellipticities of the pure monomer and oligomer that are related to “fraction unfolded” as follows:^{25,235}

$$[\theta]_T^M = f_u [\theta]^M \quad (7.4)$$

$$[\theta]_T^D = [\theta]^D - f_u [\theta]^D \quad (7.5)$$

where $[\theta]^M$ and $[\theta]^D$ are temperature dependencies of the fully monomeric and oligomeric peptides, respectively. Thus f_u can be written as:

$$f_u = \frac{[\theta] - [\theta]^D}{[\theta]^M - [\theta]^D} \quad (7.6)$$

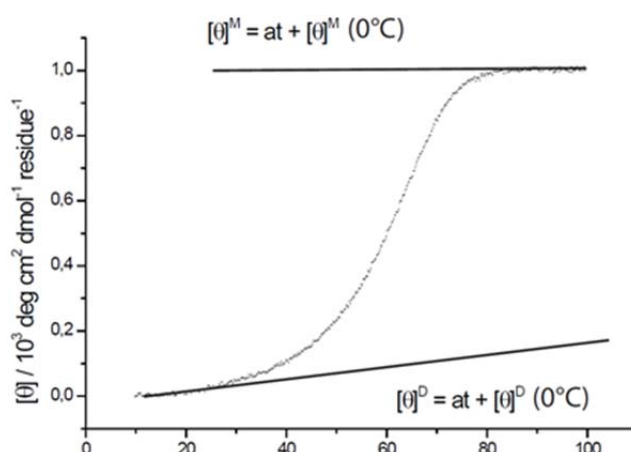


Figure 7.3: example for a melting curve obtained from thermal denaturation. Linear baselines for the oligomeric state and the monomeric state are shown, where a is the slope and t the temperature in $^{\circ}\text{C}$. $[\theta]^M$ and $[\theta]^D$ are the temperature dependencies of the fully monomeric and dimeric peptides. $[\theta]^M$ and $[\theta]^D$ represent the hypothetical ellipticity values for the unfolded and the folded peptides at 0°C . The turning point of the curve represents the melting point of the peptide assembly.

α -Helix Propensity Studies

Alanine based model peptides that were used for the determination of helix propensities of fluorinated amino acids investigated here were dissolved in 1 M NaCl, 1 mM sodium phosphate, 1 mM sodium citrate and 1 mM sodium borate buffer (pH 7.0) according to studies by Cheng *et al.*¹¹² CD measurements were performed at peptide concentrations of 30, 50, and 80 μ M at pH 7.0 and 0 °C. Data were collected from 250 to 200 nm at 0.2 nm intervals, 2 nm bandwidth, and 2 s response time, and spectra were background corrected by buffer spectrum subtraction. Each reported CD value represents the mean of at least three independent measurements. The fractional helical content of peptide (f_{helix}) was calculated from the mean residue molar ellipticity at 222 nm and the number of backbone amides ($n = 19$) using the equation $f_{helix} = [\Theta]_{222}/(40000(1-2.5/n))$, to enable comparison with values obtained from Cheng *et al.*¹¹² The helix propensity of the fluorinated amino acid at the guest position of the model peptide was calculated from the f_{helix} based on a modified Lifson-Roig theory.^{219,220}

7.2.2 Fluorescence Spectroscopy (FRET experiments)

Fluorescence resonance energy transfer (FRET) is a nonradiative process whereby an excited state donor (usually a fluorophore) transfers energy to a proximal ground state acceptor through long-range dipole–dipole interactions.^{236,237} The acceptor must absorb energy at the emission wavelength(s) of the donor, but does not necessarily have fluoresce itself (i.e. dark quenching). The rate of energy transfer depends on several factors, including the extent of spectral overlap, the relative orientation of the transition dipoles, and, most importantly, the distance between the donor and acceptor molecules.²³⁷

FRET from subunits labeled with a donor group to subunits labeled with an acceptor group was used for analysis of the relative orientation of helices in the coiled-coil bundle assemblies containing backbone extended amino acids studied here. 4-aminobenzoic acid (Abz; $\lambda_{ex} = 320$ nm) served as the donor molecule and 3-nitrotyrosine (Y(NO₂); $\lambda_{abs} = 420$ nm) served as the acceptor. A decrease in fluorescence intensity indicates that the acceptor and donor molecules have moved closer to one another. Because fluorescence should be recovered upon chemically induced unfolding, different amounts of GndHCl were added, as indicated in the corresponding figure captions.

Fluorescence spectra were recorded on a luminescence spectrometer LS 50B (Perkin Elmer) equipped with a Julabo temperature controller F12 (Julabo GmbH) using a 1 cm Quartz Suprasil cuvette (Hellma) at 20 °C. Three scans from 350 to 550 nm were performed, averaged, and the spectra were normalized to the respective maximum fluorescence. Acquired spectra were analyzed using FLWinLab 2.0 (Perkin Elmer) and Microcal Origin 6.0 (OriginLab). Individual concentration and buffer conditions are given with the respective data.

7.2.3 Oligomerization State Determination

Static Light Scattering

Size exclusion chromatography (SEC), coupled with static light scattering (SLS), refractive index (RI), and ultraviolet (UV) detection provides a universal approach for the determination of the molar mass and oligomerization state of peptides and proteins in solution (Figure 7.4).²³⁸ SEC, also referred to as gel filtration, solely serves as a fractional step separating oligomeric assemblies based on their size. The UV detector monitors absorbance at a given wavelength, the RI detector monitors changes in refractive index, and the LS detector records the access of scattered light. If the overall concentration of the sample is known, the molar mass M_W of physically separated oligomers can be calculated using the following equation:

$$M_W = k \frac{(LS)(UV)}{A(RI)^2} \quad (7.7)$$

Where M_W is the molar mass, k is the calibration constant, LS is the response of the LS detector, A is the extinction coefficient, and RI and UV are the responses from the refractive index and UV detectors, respectively.

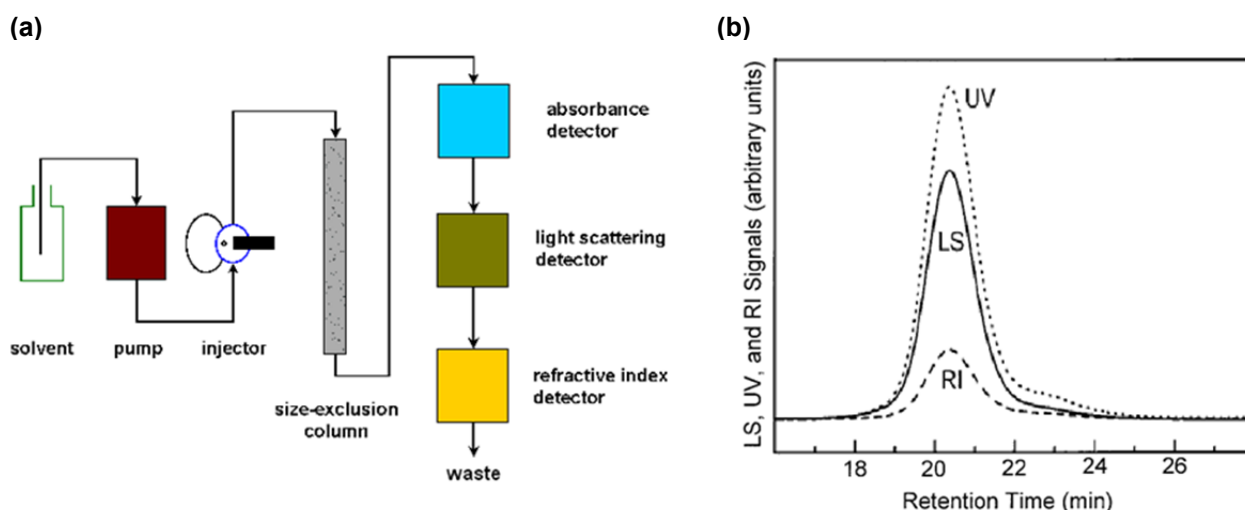


Figure 7.4: (a) experimental set up of SEC/SLS. (b) example for a chromatogram as obtained in SEC/SLS experiments

Data were collected on a Dawn Heleos 8 light scattering instrument (Wyatt Technology) coupled with an analytical SEC system (workstation: La Chrom, VWR; Pump L-2130, UV Detector L-2400; column WTC-015S5; 5 μm , 150 \AA , 7.8 x 300 mm, Wyatt Technology) at $\lambda = 220$ nm. Measurements were taken at pH 7.4 at a total peptide concentration of 60 μM and repeated two times to confirm reproducibility and give standard deviations. All measurements were performed in PBS at room temperature with a flow rate of 0.3 mL/min. Data were analyzed using the ASTRA software version 5.3.4.20 (Wyatt Technology) and Microcal Origin 6.0 (OriginLab).

N-PAGE

Polyacrylamide gel electrophoresis separates molecules according to their size. Since in native polyacrylamide gel electrophoresis (N-PAGE) gels are run under nondenaturing conditions, complexes remain associated and folded, this method can be applied to determine six helix bundle formation between the N- and C-peptides of gp41.^{213,239}

N-peptide (N36 or T21) was mixed with a C-peptide at a final concentration of 60 μ M and incubated at 37 °C for 30 min. The mixture was loaded onto 18% Tris-glycine gels at 30 μ L/per well. 6x DNA loading dye (New England Biolabs) served as loading buffer. Gel electrophoresis was carried out with 140 V constant voltage at room temperature for 2 h utilizing a *SDS-Whatman Multigel G44* electrophoresis chamber (Bometra) and a *Elektrophoresis Power Pac 300* power supply (Bio-Rad). The gel was stained with Coomassie brilliant blue and destained overnight.

7.3 Phage Display

Phage display has proven to be one of the most powerful selection techniques in the field of protein engineering.¹⁵² In the course of the current thesis it has been extensively used for the selection of preferred binding partners of nonnatural amino acids at α -helical interfaces.

Chemicals and Solvents

Agar (Carl Roth GmbH & Co. KG), agarose (Carl Roth GmbH & Co. KG), carbenicillin (Sigma-Aldrich), dNTP-Mix 10 mM (New England Biolabs), ethanol 96% (Carl Roth GmbH & Co. KG), guanidine hydrochlorid BioChemika, 99.0% (Sigma-Aldrich), milk powder - blotting grade (Carl Roth GmbH & Co. KG), MOPS PUFFERAN® (Carl Roth), polyethylene glycol (PEG) 8000 (Sigma-Aldrich), peptone made of casein – tryptically digested (Carl Roth), potassium chloride, 99.5% (Carl Roth), potassium-hydrogenphosphate, 99.0%, (Carl Roth), Rotilabo®-syringe filters, 0,22 μ m, sterile (Carl Roth), sodium acetate, 99.0%, waterfree (Carl Roth), sodium-azide Bio-Chemika Ultra, 99.5% (Sigma-Aldrich), sodium chloride, 99.5% (Carl Roth), tris-hydrochloride - buffer quality (AppliChem GmbH), Tween 20, for molecular biology (Sigma-Aldrich), yeast extract BioChemika (AppliChem GmbH).

Enzymes

Sfi I, recombinant (New England Biolabs GmbH), T4 DNA ligase (New England Biolabs GmbH), Taq DNA-Polymerase (New England Biolabs GmbH), trypsin from bovine pancreas, TPCK-treated, ≥ 10000 BAEE (Sigma Aldrich).

Bacteria, Helper Phages, Plasmids

E. coli K12 ER2738 (New England Biolabs), VCSM13 helper phages (Stratagene), pComb3H phagemid vector (GenBank database accession number: AF268280, Barbas Laboratory, TSRI).

7.3.1 Buffers and Media

Agar Plates

Lysogeny Broth (LB) agar petri dishes were prepared according to standard molecular biological practice. Carbenicillin was used at a final concentration of 100 μ g/mL.

Carbenicillin Stock Solution

Carbenicillin was dissolved in deionized water (100 mg/mL), filter-sterilized (0.22 μ m), aliquoted (1 mL), and stored at -20° C.

Guanidine Hydrochloride PBS (1 M)

Guanidine hydrochloride (25 mmol) was dissolved in 25 mL PBS-buffer and stored at RT.

Kanamycin Stock Solution

Kanamycin was dissolved in deionized water (50 mg/mL), filter-sterilized (0.22 μ m), aliquoted (1 mL), and stored at -20° C.

Lysogeny Broth (LB) Agar

10 g tryptone peptone, 5 g yeast extract, and 5 g sodium chloride were dissolved in deionized water, brought to a final volume of 1 L and the pH was adjusted to 7.0 with NaOH. Batches of 300 mL were prepared, to which 4.8 g Agar-Agar was added. Subsequently, the agar was autoclaved and stored at RT.

Lysogeny Broth (LB) medium

10 g tryptone peptone, 5 g yeast extract, and 5 g sodium chloride were dissolved in deionized water and brought to a final volume of 1 L. Post pH adjustment to 7.0 the medium was autoclaved and stored at 4° C.

PBS-buffer

8.0 g sodium chloride, 0.2 g potassium chloride, 1.7 g disodium hydrogen phosphate, and 0.163 g potassium dihydrogen phosphate were dissolved in deionized water and brought to a final volume of 1 L. The pH was adjusted to 7.4 with HCl, the solution autoclaved, and finally stored at 4° C.

PBS-T

Phosphate Buffered Saline (PBS) containing 0.1 %, 0.5 %, or 1.0 % (v/v) Tween 20.

PEG/NaCl solution

100 g polyethylene glycol-8000 (20 %) and 73 g sodium chloride (2.5 M) were dissolved in deionized water, brought to a final volume of 500 mL and stored at 4° C.

SOC medium

4 g tryptone peptone, 1 g yeast extract, and 0.1 g sodium chloride were dissolved in 150 mL deionized water. 2 mL 0.25 M potassium chloride and 2 mL 1 M magnesium chloride were added. After the pH was adjusted to 7.0, the medium was brought to a final volume of 200 mL. Subsequently, the medium was divided into 50 mL aliquots and autoclaved. Finally, 200 μ L 1 M glucose solution (filter-sterilized; 0.22 μ m) were added, and the medium was stored at 4° C.

Super Broth (SB) medium

10 g MOPS, 20 g yeast extract, and 30 g tryptone peptone were dissolved in deionized water, and brought to a final volume of 1 L. After the pH was adjusted to 7.0, the medium was autoclaved and stored at 4° C.

TBS-buffer

8.7 g sodium chloride and 6.1 g Tris base were dissolved in deionized water and brought to a final volume of 1 L. After the pH was adjusted to 7.4 with HCl, the solution was autoclaved and stored at RT.

7.3.2 General Techniques

DNA-Preparation and purification

DNA-purification was performed by agarose-gel electrophoresis, which separates molecules according to their size, utilizing *PerfectBlue* electrophoresis chambers (Peqlab Biotechnologie) and a *PowerPac 300* power supply (Bio-Rad Laboratories Inc.). To assign DNA bands to a specific molecular weight range, the 2-log DNA ladder from New England Biolabs was simultaneously run on each gel. DNA extraction from the agarose-gel was carried out with a *peqGOLD* Gel Extraction Kit (Peqlab Biotechnologie) according to the manufacturer's instructions. After purification, DNA-quantification was performed with a *NanoDrop 2000* spectrophotometer (Thermo Fisher Scientific Inc.). The purity of the DNA samples was assessed via determination of the optical density (OD) ratio at 260 nm to 280 nm (OD_{260} / OD_{280} -ratio). Values between 1.7 and 2.0 indicate sufficient purity of the prepared DNA.

Plasmid preparation

Plasmid DNA was isolated from *E. coli* cell cultures. Therefore, 5 mL prewarmed LB-medium containing 100 µg/mL carbenicillin was inoculated with a single bacterial colony from a carbenicillin-agar culture dish and the culture was grown for 16 h at 37° C and 200 rpm. The bacteria were pelleted using a *Heraeus Fresco 17* microcentrifuge (Thermo Fisher Scientific Inc.) in 1.5 mL centrifuge tubes by spinning for 1 min at 10000 rpm. Supernatant was discarded and plasmid DNA was isolated and purified using a *peqGOLD* Plasmid Miniprep-Kit (Peqlab Biotechnologie), according to the manufacturer's instructions. DNA was eluted from the spin column in 50 µL deionized water and stored at -20° C.

Colony PCR and DNA-Sequencing

The polymerase chain reaction (PCR) is a commonly applied method to amplify DNA by several orders of magnitude. DNA sequencing is used to determine the primary structure (nucleotide sequence) of a given DNA fragment. The following primers were purchased from *biomers.net* GmbH and used for colony PCR as well as sequencing of pIII-fusion DNA inserts:

ompseq (sense primer): 5'-AAG ACA GCT ATC GCG ATT GCA G-3'

gback (backward primer): 5'-GCC CCC TTA TTA GCG TTT GCC ATC-3'

PCR-reactions were performed in 0.2 mL centrifuge tubes utilizing a *Primus 25 advanced* thermocycler (Peqlab Biotechnologie GmbH). The amplification of DNA encoding for pIII-fusion protein inserts was carried out according to the following PCR protocol:

PCR sample:

3 μ L bacteria suspension
 5 μ L 10 x PCR buffer
 5.25 μ L dNTP-mix (2.5 mM)
 0.5 μ L ompseq (250 ng/ μ L)
 0.5 μ L gback (250 ng/ μ L)
 0.25 μ L *Taq*-polymerase
 35.5 μ L deionized water

PCR program:

94° C for 2min
 94° C for 0,5min
 55° C for 0.5min
 72° C for 0.5 min
 72° C for 10min
 4° C until further processing

} 35 cycles

PCR products were visualized via agarose gel electrophoresis. Bands at ~ 300 bp correspond to gene III fusion protein inserts (Figure 7.5).

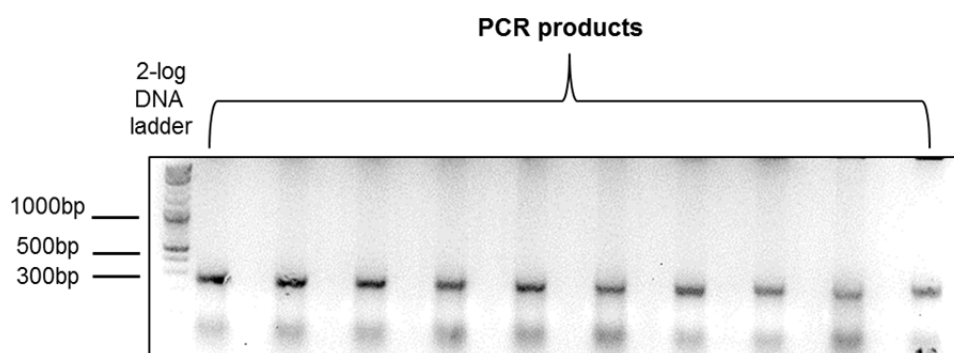


Figure 7.5: Example of colony PCR products analysed by gel electrophoresis

To obtain the exact nucleotide sequences of the gene III fusion inserts, samples to be subjected to DNA sequencing were submitted to *Seqlab* Sequence Laboratories Göttingen GmbH. Therefore, DNA extraction from the agarose-gel was carried out with a *peqGOLD Gel Extraction Kit* (Peqlab Biotechnologie) according to the instruction manual and DNA sequencing samples were prepared as follows:

- 0.7 μ L Tris/HCl buffer (pH 8.5, 10 mM)
- 0.7 μ L gback primer (20 pM)
- x μ L DNA (100 ng)
- x μ L H₂O (total volume 7 μ L)

Preparation of *E. coli* ER2738 cells

E. coli ER2738 cells were purchased from New England Biolabs and streaked out on an agar plate, which was incubated overnight at 37° C. Subsequently, 100 mL prewarmed LB medium was inoculated with a single bacterial colony from the plate. The culture was then incubated for 16 h at 37° C with 200 rpm agitation. 500 μ L of the culture was transferred into 500 mL prewarmed SB medium, in duplicate, and allowed to grow at 37° C with 200 rpm agitation to an OD₅₉₅ of 0.5 (approx. 3h). Afterwards, the cell cultures were divided and

transferred into four centrifuge bottles that were precooled on ice. The cells were then centrifuged for 10 min at 3000 g and 4° C using an *Allegra X-15* centrifuge (Beckman Coulter) equipped with a *SX4750A* rotor. The supernatant was discarded and several washing steps were performed on ice. Therefore, bacterial pellets were suspended in 5 mL ice-cold, deionized water, transferred in two centrifuge beakers and filled to approx. 75 % filling level with water. After centrifugation for 10 min at 3000 g and 4° C, cells were resuspended in 5 mL water containing 10 % glycerol, transferred into 50 mL centrifuge tubes and filled to 20 mL with 10 % glycerol. After another round of centrifugation (10 min at 3000 g and 4° C), both cell pellets were suspended in 5 mL 10 % glycerol each, and combined in one 50 mL centrifuge tube, which was subsequently filled to 30 mL with 10 % glycerol. Again, this tube was centrifuged (10 min at 3000 g / 4° C) and cells were suspended in 1 mL 10 % glycerol and diluted with additional 10 % glycerol to an OD₅₉₅ of 1.0, which was measured in 1:100 dilutions with a *BioPhotometer plus* (Eppendorf AG). Finally, cells were divided into 55 µL and 300 µL aliquots, immediately frozen in liquid nitrogen and stored at -80° C.

7.3.3 Molecular cloning and phage library preparation

The following DNA oligomers coding for desired peptide libraries were purchased from *biomers*. Randomised positions are marked in green and sticky ends for successful cloning into *pComb3HSS* plasmid are underlined:

Sense strand of VPE library used for biopanning against VPK-a₁₆:

5' phosphate- CGGCC GAG GTT AGC GCG CTG GAA AAG GAG GTG GCC AGT **NNK**
GAG AAA GAG **NNK** AGT GCC **NNK** GAA AAG AAA GTA GCG AGC CTG AAA AAG GAG
GTA AGT GCG TTA GAA GGCCAGGC 3'

Antisense strand of VPE library used for biopanning against VPK-a₁₆:

5' phosphate- TGGCC TTC TAA CGC ACT TAC CTC CTT TTT CAG GCT CGC TAC TTT
CTT TTC **MNN** GGC ACT **MNN** CTC TTT CTC **MNN** ACT GGC CAC CTC CTT TTC CAG
CGC GCT AAC CTC GGCCGCCT 3'

Sense strand of VPE library used for biopanning against VPK-d₁₉:

5' phosphate - CGGCC GAG GTT AGC GCG CTG GAA AAG GAG GTG GCC AGT TTA
GAG AAA GAG **NNK** AGT GCC **NNK** GAA AAG AAA **NNK** GCG AGC CTG AAA AAG GAG
GTA AGT GCG TTA GAA GGCCAGGC 3'

Antisense strand of VPE library used for biopanning against VPK-d₁₉:

5' phosphate - TGGCC TTC TAA CGC ACT TAC CTC CTT TTT CAG GCT CGC **MNN**
TTT CTT TTC **MNN** GGC ACT **MNN** CTC TTT CTC TAA ACT GGC CAC CTC CTT TTC
CAG CGC GCT AAC CTC GGCCGCCT 3'

Sense strand of Acid-pp library:

5' phosphate - CGGCC CTG AGC GCG CTG GAG AAG GAG CTG GCG AGC CTG GAG AAG GAG **NNK** AGC GCG **NNK NNK** AAG **NNK** CTG GCG AGC CTG GAG AAG GAG CTG AGC GCG CTG GAG AAG GAG GGCCAGGC – 3'

Antisense strand Acid-pp library:

5' phosphate - TGGCC CTC CTT CTC CAG CGC GCT CAG CTC CTT CTC CAG GCT CGC CAG **MNN** CTT **MNN MNN** CGC GCT **MNN** CTC CTT CTC CAG GCT CGC CAG CTC CTT CTC CAG CGC GCT CAG GGCCGCCT – 3'

N: A,C,G, or T

M: A or C

K: G or T

DNA-Annealing

Sense and antisense DNA-strand libraries were dissolved in deionized water and dilutions of 15 ng/ μ L (500 μ L) were prepared. Subsequently, 3 \times 80 μ L of each were mixed and incubated together with 20 μ L annealing-buffer (10 \times) and 20 μ L deionized water for 10 min at 95 $^{\circ}$ C (thermocycler). Subsequently, the reaction mixture was gradually cooled down to 25 $^{\circ}$ C in (0.5 $^{\circ}$ C/min, thermocycler). The reaction products were purified by agarose-gel electrophoresis (2 % agarose, Figure 7.6), bands corresponding to annealed DNA were cut out and purified using a *peqGOLD* gel extraction kit according to manufactures manual (Peqlab Biotechnologie).

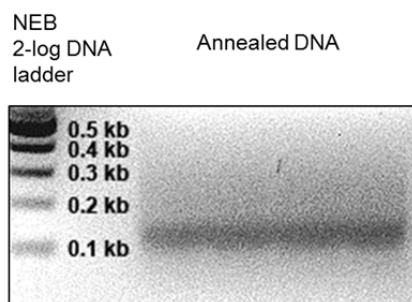


Figure 7.6: Gel electrophoresis example of an annealed DNA sense and antisense strand

Plasmid digestion

3 \times 6 μ L *pComb3HSS* vector (kindly provided by the Barbas lab; Scripps Research Institute, La Jolla, CA, USA) were incubated for 5 h at 50 $^{\circ}$ C in 6 μ L NEBuffer 2 (10 \times), 0.6 μ L BSA, 2 μ L Sfi1 (40 U) and 45.4 μ L deionized water. The reaction mixture was then separated by gel electrophoresis (0.7 % agarose, Figure 7.7), the DNA-band corresponding to the digested (linearized) vector was cut out and extracted with *peqGOLD* gel extraction kit according to the manufacturer's instructions.

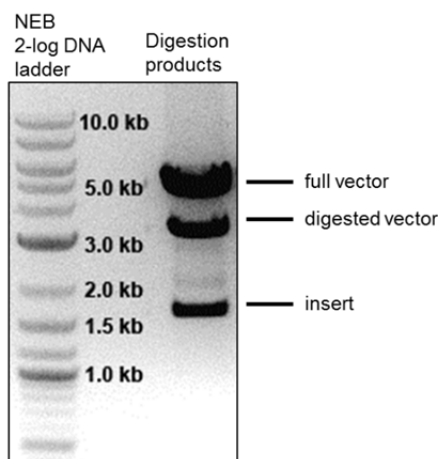


Figure 7.7: Agarose gel of the digested *pComb3HSS* vector

Ligation

1 μg double stranded DNA was ligated into linearized *pComb3HSS* by incubating it with 1.4 μg of the digested vector and 4000 U *T4 DNA ligase* (10 μL) in 200 μL *T4 DNA ligase* reaction buffer at 16° C overnight. As negative control, 140 μg digested vector with 400 U *T4 DNA ligase* in 20 μL buffer was subjected to the same conditions. After taking 1 μL of each for test-electroporation, the reaction mixtures were concentrated utilizing a *peqGOLD* miniprep kit by dilution with a five-fold excess of PW buffer, and elution with 20 μL deionized water.

Transformation of *E. coli*

Transformation of *E. coli* cells with ligated plasmid vectors carrying DNA-libraries was performed using an *Electroporator 2510* (Eppendorf AG). Electrocompetent *E. coli ER2738* cells (competence $\geq 2 \times 10^{10}$ cfu/ μg DNA) for phage display with the following genotype were purchased from Lucigen Corp..

[F'*proA*⁺*B*⁺ *lacI*^q Δ (*lacZ*)*M15* *zzf::Tn10* (tet^r)] *fhuA2 glnV* Δ (*lac-proAB*) *thi-1* Δ (*hdsS-mcrB*)5

For the transformation of *E. coli* cells with DNA-libraries, the entire eluate from plasmid purification of ligated vectors was mixed with 60 μL electrocompetent cells and electroporated with 2.5 kV in a previously cooled cuvette. Immediately afterwards, the cells were collected with 1 mL of warm SOC-medium and transferred into a 12 mL incubation tube with 2 mL of warm SOC. Subsequently, the cuvette was washed two times with 1 mL SOC (total culture volume of 5 mL). This culture was incubated for 1 h at 37° C and 200 rpm. Afterwards, the bacterial suspension was transferred into 10 mL warm SB-medium, 50 μL of a 10^{-2} , 10^{-3} , and 10^{-4} dilutions were streaked out onto Carbenicillin-Agar plates and incubated overnight at 37° C.

3 μL Carbenicillin (100 mg/mL) was added to the remaining bacterial culture, which was then shaken at 37° C/200 rpm for 1 h. After adding another 4.5 μL of Carb., the culture was

grown for an additional hour. In order to generate recombinant bacteriophages with the transformed bacteria, the cells were then poured into a previously warmed (37° C) centrifuge tube containing 183 mL SB-medium and 92.5 µL Carb.. Subsequently, 2 mL of *VCSM13* helper phage were added. The bacteria culture was now incubated for 2 h at 37° C/200 rpm to allow for phage infection. Finally, 280 µL Kanamycin (50 mg/mL) were added to select infected bacteria and the culture was shaken overnight at 37° C / 200 rpm.

20 colonies were picked from the plates, the gene III fusion protein insert amplified by colony-PCR, and separated via agarose gel electrophoresis. DNA bands corresponding to the insert size were cut out, purified with *peqGOLD* gel extraction kit (Peqlab Biotechnologie). In order to analyze the diversity in the randomized positions (see Table 9.1, p. 107; Table 9.2, p. 109; Table 9.4, p. 114), samples were sequenced by *Seqlab* Sequence Laboratories Göttingen GmbH.

Preparation of phage library

The overnight bacterial culture was centrifuged at 3000 g at 4° C for 30 min with a *Allegra X-15* centrifuge (Beckman Coulter) equipped with a *SX4750A* rotor. Pellets and debris were discarded. To precipitate the phages, the supernatant was transferred into a fresh centrifuge tube containing 40 mL ice cold PEG 8000 (20 %)/NaCl (2.5 M) solution, and stored on ice for 30 min. Afterwards, the phage were centrifuged for 30 min at 5250 g (maximum of *Allegra X-15*) at 4°C. The supernatant was discarded, and the precipitate was air dried by putting the beaker upside-down on a paper towel for 10 min. Pelleted phages were resuspended with 2 mL PBS and sterile filtered with *Rotilabo*-syringe filters (CME, 0.22 µm pore size, Carl Roth). For long term storage at 4°C, sodium azide was added to a final concentration of 0.02 % (w/v).

7.3.4 Phage display

Infection of E. coli

In order to generate a fresh phage suspension with newly expressed pIII-fusions protein on their surface, suitable for panning experiments (round 1), *E. coli ER2738* cultures were prepared and infected with the phage library. Therefore, 50 mL warm SB-medium were inoculated with 50 µL *E. coli* and the culture grown at 37° C/200 rpm to an OD₆₀₀ ≈ 1 (approx. 3 h). Subsequently, the cultures were infected with 50 µL phage library suspension for 15 min at RT by gently agitating every minute for 15 sec. Afterwards, 50 µL of 10³-, 10⁴- and 10⁵-dilutions (in SB) were streaked out on carbenicillin containing agar plates and incubated overnight at 37° C.

10 µL Carb. (100 mg/mL) was added to the remaining culture, and shaken for 1 h at 37° C/200 rpm. Afterwards, 15 µL Carb. was added and the cells were allowed to grow for an additional hour. The cell cultures were then transferred to centrifuge beakers containing

148 mL prewarmed SB-medium and 75 μ L Carb. (100 mg/mL). At this point the cells were infected with 2 mL *M13*-helper phages. Subsequently, the cultures were incubated for 2 h at 37° C/200 rpm before 280 μ L Kanamycin (50 mg/mL) was added and shaking was continued overnight.

Phage preparation

The overnight culture was centrifuged at 3000 g/4° C for 30 min, and the supernatant was transferred into a centrifuge tube with 40 mL (round 1) and 20 mL (round 2-6) cold 20 % PEG 8000 / NaCl (2.5 M) solution, respectively. To precipitate the phages, the centrifuge tube was kept on ice for 30 min. Afterwards, the phages were pelleted by centrifugation at 5250 g/4° C, and, after discarding the supernatant, air dried for 10 min. The precipitate was resuspended in 2 mL (round 1) and 1 mL (round 2-6) PBS, and sterile filtered (0.22 μ m filter). The phages were stored on ice until immediate use in panning.

Panning

For selection of binding partners of the synthesized biotinylated peptides, 30 μ L *Dynabeads M-280 Streptavidin* magnetic particles (Invitrogen) were mixed with 5.5 μ L peptide solution (5 μ g/ μ L) and 500 μ L PBS (negative control without peptide) in 1.5 mL tubes, and rotated on a *Stuart Rotator SB2* (Bibby Scientific Ltd.) at RT for 45 min. Afterwards, the tubes were placed for 4 min on a *Roti-Mag* separator (Carl Roth GmbH). With the help of this external magnet the particles were collected on the reaction vessel wall and phages with no or low affinity fusion peptides are washed off. Therefore, after the supernatant was discarded, the magnetic beads were washed two times with 500 μ L PBS-T (0.1 %) by inverting the tube 10 times and allowing 4 min of separation time on the magnet post for every wash. 500 μ L of 5 % milk powder in PBS were added to block remaining free binding sites, and the tubes were rotated for a further 45 min.

Subsequently, the beads were separated from the milk solution on the magnetic separator (4 min), the milk was discarded, and a mix of 250 μ L freshly prepared phage solution (see *phage preparation*) and 250 μ L PBS were added. For biopanning via coiled-coil formation, the tubes were rotated for 90 min. Afterwards, unbound phage were washed off the magnetic beads with specific solutions depending on the panning round and the optimized protocol for the target sequence (Table 7.5; Table 7.6). Each washing step was followed by 4 min fixation of the magnetic particles on the magnet.

Table 7.5: Washing steps after biopanning against VPK variants

| Panning round | Washing solutions | Volume |
|---------------|----------------------------------------------------|----------------------------------------|
| 1 | PBS-T (0.1 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 2 | PBS-T (1.0 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 3 | PBS-T (1.0 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 4 | PBS-T (1.0 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 5 | PBS-T (1.0 % Tween 20) 1 M GndHCl in PBS TBS | 3 × 500 µL 3 × 500 µL 1 × 500 µL |

Table 7.6: Washing steps after biopanning against B3β2γ variants.

| Panning round | Washing solutions | Volume |
|---------------|----------------------------------------------------|----------------------------------------|
| 1 | PBS-T (0.1 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 2 | PBS-T (0.5 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 3 | PBS-T (0.5 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 4 | PBS-T (1.0 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 5 | PBS-T (1.0 % Tween 20) TBS | 4 × 500 µL 1 × 500 µL |
| 6 | PBS-T (1.0 % Tween 20) 1 M GndHCl in PBS TBS | 3 × 500 µL 3 × 500 µL 1 × 500 µL |

Reinfection of *E. coli*

Bacterial cultures for reinfection were prepared by inoculation of 5 mL warm SB-medium with 10 µL *E. coli* ER2738 each, and incubation for 2 h at 37° C ($OD_{600} \approx 1$). To enable the evaluation of infection efficiency of the phage suspension prior to biopanning (input-titer), 50 µL bacterial culture were infected with 1 µL 10^6 -diluted phage preparation (in PBS) by incubation at RT/750 rpm for 20 min in a *Thriller Thermoshaker Incubator* (PeqLab Biotechnologie). The entire suspension was then streaked out on Carb.-Agar plates, incubated overnight at 37° C, and colonies were counted after overnight incubation.

The phage that were selected during biopanning and subsequently thoroughly washed were eluted from the magnetic beads by trypsin digestion. Therefore, the beads were suspended in 25 µL freshly prepared trypsin solution (10 mg/mL) and shaken at RT/750 rpm for 30 min in the thermoshaker. The digestion was quenched by the addition of 80 µL SB-medium. After separation of the suspensions on the magnet, the supernatant was used for reinfection of *E. coli* by adding it to 5 mL of prepared bacterial cultures and shaking at 37° C/200 rpm for 30 min. A 10^{-2} and a 10^{-3} dilution in SB of every sample and negative

control were streaked out on Carb.-Agar plates and incubated overnight at 37° C. Colonies grown on these plates were counted on the next day in order to evaluate the enrichment of selected binding partners (output titre).

5 mL warm SB and 2 µL Carb. (100 mg/mL) were added to the remaining bacterial suspensions, followed by 1 h shaking at 37° C/200 rpm. The cultures were then poured into 90 mL prewarmed SB containing 46 µL Carb. (100 mg/mL), in centrifuge tubes, and infected with 1 mL *M13* helper phage. After further shaking for 90 min at 37° C/200 rpm, 140 µL Kanamycin (50 mg/mL) was added in order to amplify solely the helper phage infected bacteria, and the cultures were incubated overnight at 37° C/200 rpm. Overnight bacterial cultures were prepared for biopanning rounds 2 – 6 by repeating the procedure starting from phage preparation.

Colonies were screened by PCR for the pIII fusion protein insert and prepared with the *peqGOLD* miniprep kit according to the manufacturer's instructions. Plasmid DNA sequencing was carried out by *Seqlab* Sequence Laboratories Göttingen GmbH.

9 Supplementary Data

9.1 Phage Display with Fluorinated Amino Acids

9.1.1 Selection against VPK-X₁₆

Table 9.1: Sequencing results after test infection of *E. coli* with generated VPE_{a16} library

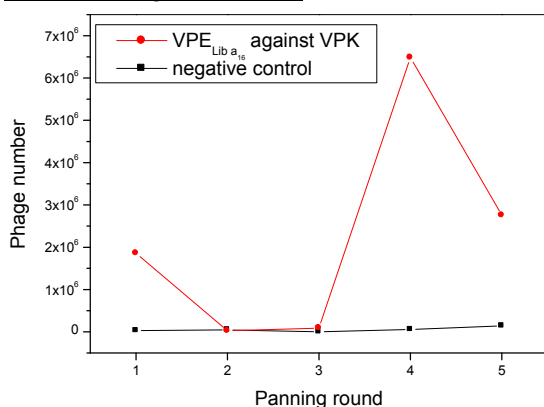
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/TTG | Asn/AAT | Gly/GGT |
| 2 | Leu/CTT | Leu/CTC | Leu/CTT |
| 3 | Ile/ATC | Ile/ATT | His/CAC |
| 4 | Pro/CCA | His/CAT | Pro/CCC |
| 5 | Asn/AAT | His/CAC | Arg/CGA |

The number of clones free from defects in the phage library was calculated from the colony number found through a test infection of an *E. coli* ER2738 culture (1.7×10^7 mutation-free clones). The library screened was sufficient to multiply include every possible combination of amino acids at the three randomized positions (theoretical size: $20^3 = 8000$).

Selection against VPK and fluorinated variants of Abu in VPK's position a₁₆

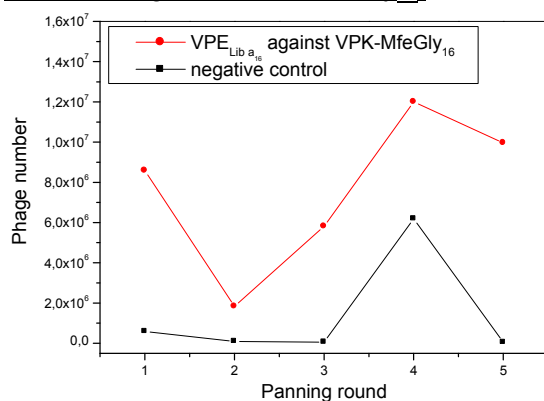
In the following phage enrichment is depicted as the number of selected phages. Moreover, phenotypes and genotypes of the three randomized VPE positions (d'_{12} , a'_{16} , and d'_{19}) after the 5th panning round against the different VPK-X₁₆-variants are presented as sequenced.

Selection against VPK:



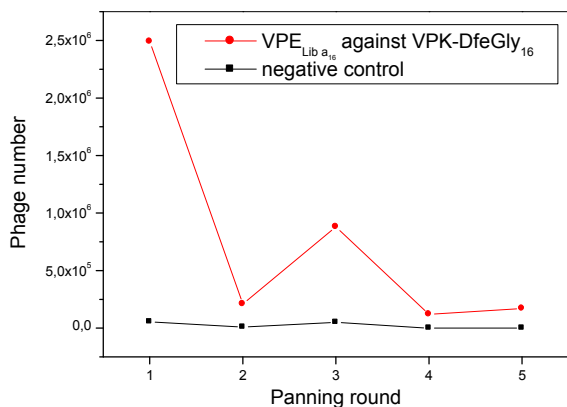
| clone | VPE-d' ₁₂ | VPE-a' ₁₆ | VPE-d' ₁₉ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTC | Leu/CTT | Leu/CTT |
| 2 | Leu/CTC | Leu/CTC | Leu/CTT |
| 3 | Leu/CTC | Leu/CTT | Leu/CTC |
| 4 | Leu/CTT | Ile/ATT | Leu/CTT |
| 5 | Leu/CTC | Leu/CTT | Leu/CTC |
| 6 | Leu/CTA | Ile/ATC | Leu/CTG |
| 7 | Leu/CTT | Leu/CTT | Leu/CTT |
| 8 | Tyr/TAT | Ile/ATT | Leu/CTT |
| 9 | Leu/CTG | Leu/CTT | Leu/CTT |

Selection against VPK-MfeGly₁₆:



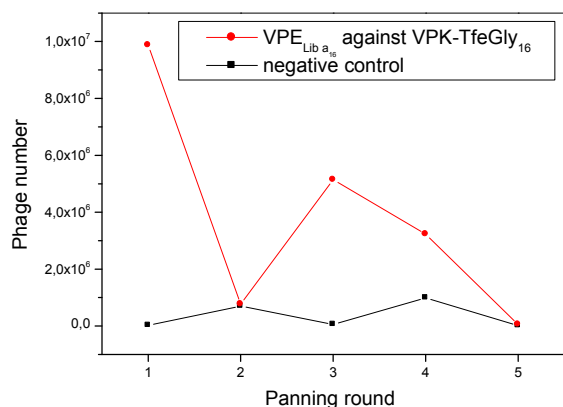
| clone | VPE-d' ₁₂ | VPE-a' ₁₆ | VPE-d' ₁₉ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTT | Leu/CTC | Leu/CTG |
| 2 | Leu/CTG | Ile/ATT | Leu/CTC |
| 3 | Leu/CTT | Ile/ATT | Leu/CTT |
| 4 | Leu/CTC | Ile/ATT | Leu/CTG |
| 5 | Leu/CTG | Ile/ATA | Leu/CTC |
| 6 | Leu/CTT | Ile/ATC | Leu/CTG |
| 7 | Leu/CTC | Ile/ATC | Leu/CTG |

Selection against VPK-DfeGly₁₆:



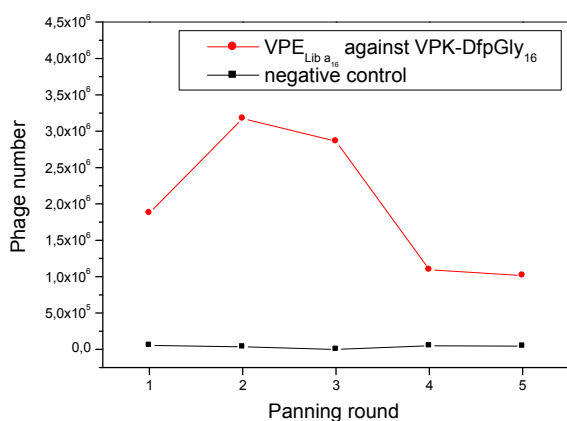
| clone | VPE-d' ₁₂ | VPE-a' ₁₆ | VPE-d' ₁₉ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTC | Ile/ATT | Leu/CTT |
| 2 | Leu/CTC | Leu/CTC | Leu/CTC |
| 3 | Leu/CTC | Ile/ATA | Leu/CTC |
| 4 | Ile/ATC | Leu/CTA | Leu/CTC |
| 5 | Leu/CTT | Ile/ATC | Leu/CTT |
| 6 | Leu/CTG | Leu/CTG | Leu/CTG |
| 7 | Leu/TTG | Ile/ATC | Leu/CTC |

Selection against VPK-TfeGly₁₆:



| clone | VPE-d' ₁₂ | VPE-a' ₁₆ | VPE-d' ₁₉ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTG | Arg/CGT | Leu/CTT |
| 2 | Leu/CTT | Leu/CTT | Leu/CTC |
| 3 | Leu/CTT | Leu/CTT | Leu/CTC |
| 4 | Trp/TGG | Ile/ATA | Leu/CTT |
| 5 | Leu/CTT | Leu/CTT | Leu/CTG |
| 6 | Leu/CTT | Ile/ATC | Leu/CTG |

Selection against VPK-DfpGly₁₆:



| clone | VPE-d' ₁₂ | VPE-a' ₁₆ | VPE-d' ₁₉ |
|-------|----------------------|----------------------|----------------------|
| 1 | Ile/ATT | Ile/ATT | Leu/CTT |
| 2 | Leu/CTG | Ile/ATC | Leu/CTC |
| 3 | Leu/CTT | Ile/ATC | Leu/CTT |
| 4 | Leu/CTT | Ile/ATT | Leu/CTT |
| 5 | Leu/CTC | Leu/CTT | Leu/CTC |
| 6 | Leu/CTT | Leu/CTT | Leu/CTT |
| 7 | Leu/CTG | Leu/CTT | Leu/CTT |
| 8 | Leu/CTG | Leu/CTT | Leu/CTT |

9.1.2 Selection against VPK-X₁₉

Table 9.2: Sequencing results after test infection of *E. coli* with generated VPE_{d₁₉} library

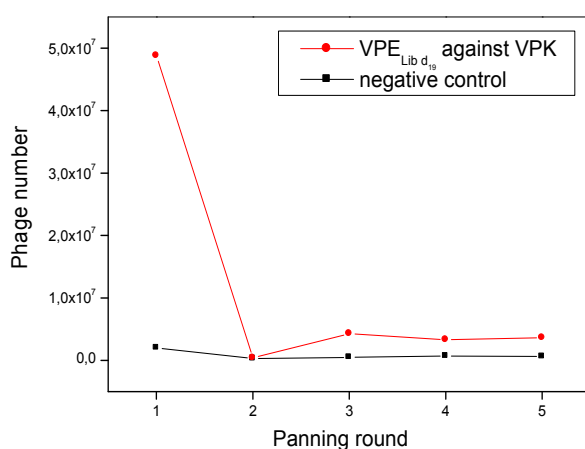
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Lys/AAG | Met/ATG | Thr/ACG |
| 2 | Phe/TTT | Ser/TGG | Val/GTT |
| 3 | Glu/GAG | Glu/GAG | Arg/CGT |
| 4 | Gly/GGT | Leu/TTG | Gln/CAG |
| 5 | Pro/CCT | Arg/CGG | Pro/CCG |

The number of clones free from defects in the phage library was calculated from the colony number found through a test infection of an *E. coli* ER2738 culture (1.2×10^7 mutation-free clones). The library screened was sufficient to multiply include every possible combination of amino acids at the three randomized positions (theoretical size: size: $20^3 = 8000$).

Selection against Fluorinated Variants of Abu in VPK's Position d₁₉

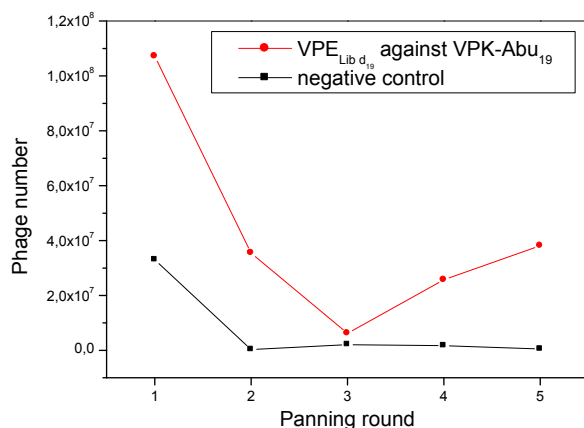
In the following phage enrichment is depicted as the number of selected phages. Moreover, mutation free phenotypes and genotypes of the three randomized VPE positions (a'₁₆, d'₁₉, and a'₂₃) after the 5th panning round against the different VPK-X₁₉-variants are presented.

Selection against VPK:



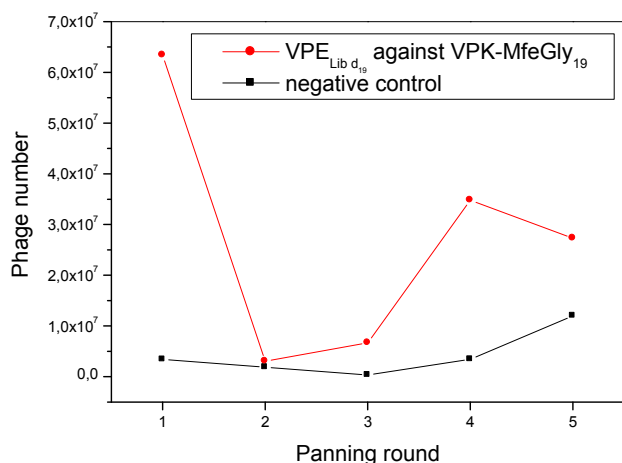
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 2 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 3 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 4 | Leu/CTT | Leu/CTG | Tyr/TAT |
| 5 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 6 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 7 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 8 | Leu/CTG | Leu/CTT | Tyr/TAT |
| 9 | Leu/CTG | Leu/CTT | Tyr/TAT |

Selection against VPK-Abu₁₉:



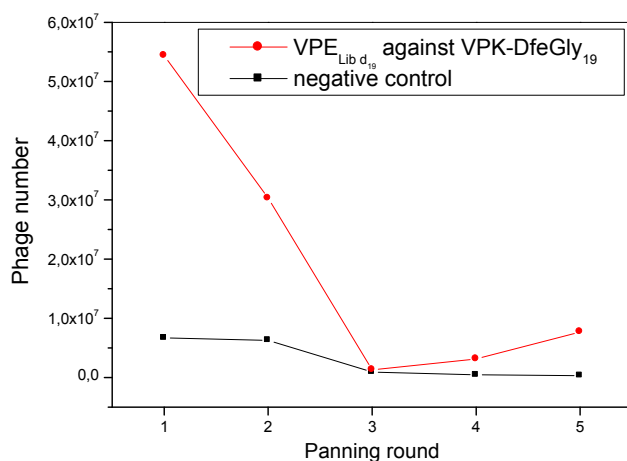
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/TTG | Leu/CTT | Ile/ATT |
| 2 | Leu/CTT | Leu/CTG | Leu/CTG |
| 3 | Val/GTG | Leu/CTG | Leu/TTG |
| 4 | Leu/CTT | Leu/CTG | Leu/CTT |
| 5 | Leu/CTT | Leu/CTT | Ile/ATT |
| 6 | Leu/CTT | Leu/CTT | Leu/TTG |
| 7 | Leu/CTG | Leu/CTG | Leu/TTG |
| 8 | Leu/CTG | Leu/CTG | Leu/TTG |
| 9 | Leu/CTT | Leu/CTT | Leu/CTT |
| 10 | Ile/ATT | Leu/CTT | Ile/ATT |

Selection against VPK-MfeGly₁₉:



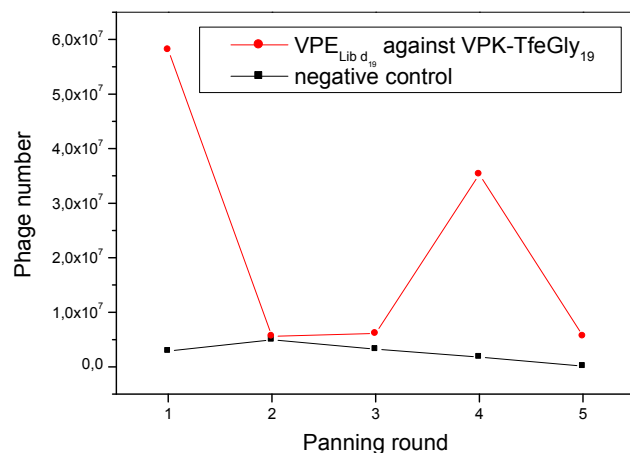
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/CTT | Leu/CTG | Leu/CTG |
| 2 | Leu/CTT | Leu/CTG | Leu/CTG |
| 3 | Leu/CTG | Leu/TTG | Leu/CTT |
| 4 | Leu/TTG | Leu/CTG | Leu/CTT |
| 5 | Leu/CTG | Leu/TTG | Ile/ATT |
| 6 | Leu/CTG | Leu/CTG | Leu/TTG |
| 7 | Leu/CTG | Leu/CTT | Leu/CTT |
| 8 | Leu/CTG | Leu/CTG | Ile/ATT |

Selection against VPK-DfeGly₁₉:



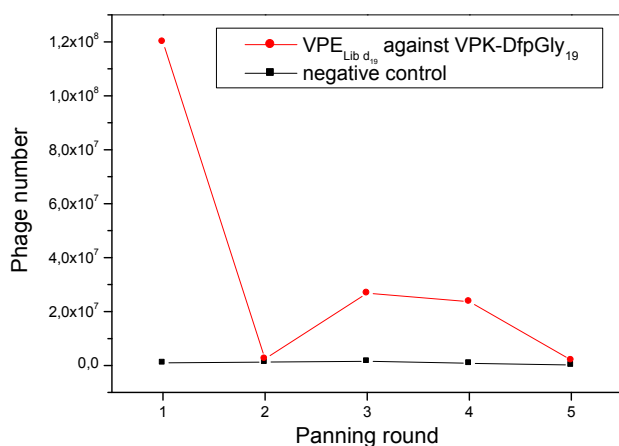
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/TTG | Leu/CTT | Leu/CTG |
| 2 | Leu/CTT | Leu/TTG | Leu/CTG |
| 3 | Leu/CTG | Leu/CTG | Leu/TTG |
| 4 | Leu/CTG | Leu/CTG | Ile/ATT |
| 5 | Leu/CTT | Leu/CTT | Leu/TTG |
| 6 | Leu/CTG | Leu/CTT | Ile/ATT |
| 7 | Leu/CTG | Leu/CTT | Leu/TTG |
| 8 | Leu/CTT | Leu/CTG | Leu/CTG |
| 9 | Leu/TTG | Leu/CTT | Leu/CTT |

Selection against VPK-TfeGly₁₉:



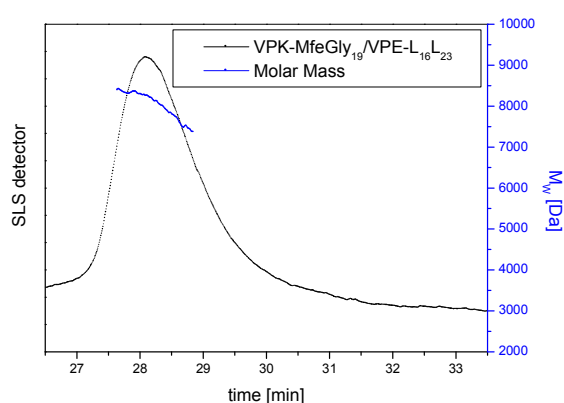
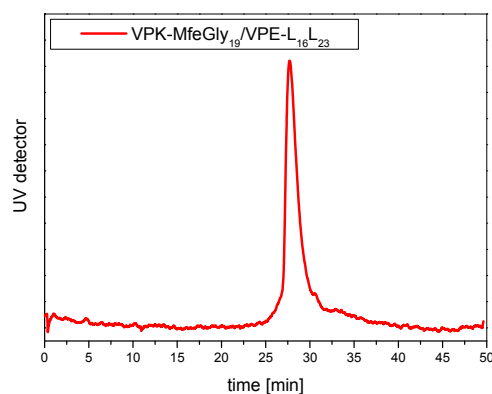
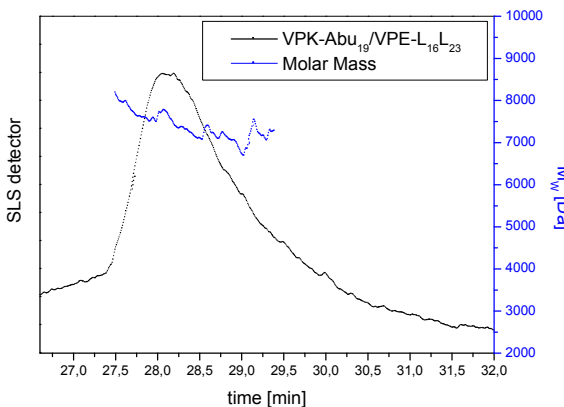
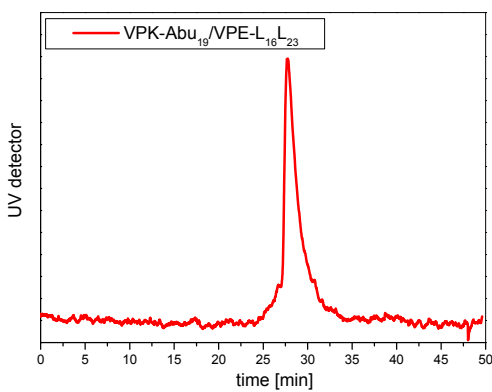
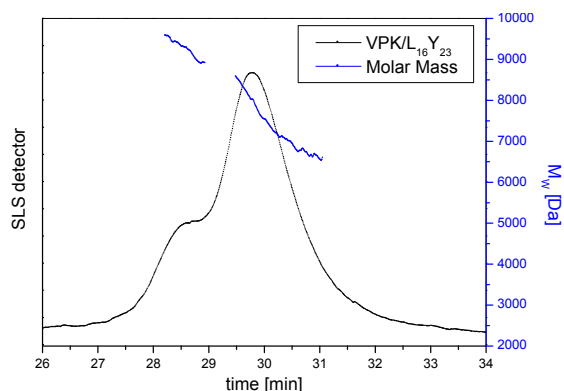
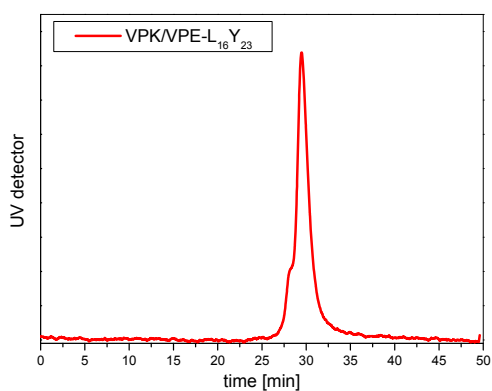
| clone | VPE-a' ₁₆ | VPE-d' ₁₉ | VPE-a' ₂₃ |
|-------|----------------------|----------------------|----------------------|
| 1 | Leu/TTG | Leu/CTG | Leu/CTG |
| 2 | Leu/CTT | Leu/CTG | Leu/CTG |
| 3 | Leu/CTT | Leu/TTG | Leu/CTT |
| 4 | Leu/CTT | Leu/CTT | Leu/CTG |
| 5 | Leu/CTT | Leu/CTT | Leu/CTG |
| 6 | Leu/CTT | Leu/CTT | Leu/CTG |
| 7 | Leu/TTG | Leu/TTG | Leu/CTT |
| 8 | Leu/CTT | Leu/CTT | Leu/CTT |
| 9 | Leu/CTG | Leu/CTT | Leu/TTG |

Selection against VPK-DfpGly₁₉:



| clone | VPE-a ¹⁶ | VPE-d ¹⁹ | VPE-a ²³ |
|-------|---------------------|---------------------|---------------------|
| 1 | Leu/CTG | Leu/CTT | Leu/CTG |
| 2 | Leu/TTG | Leu/CTG | Leu/CTG |
| 3 | Leu/TTG | Leu/CTG | Ile/ATT |
| 4 | Leu/CTG | Leu/CTT | Leu/CTG |
| 5 | Leu/TTG | Leu/CTT | Leu/TTG |
| 6 | Leu/CTT | Leu/CTT | Leu/CTG |
| 7 | Leu/TTG | Leu/CTG | Leu/TTG |
| 8 | Leu/CTT | Leu/CTT | Leu/CTT |
| 9 | Leu/TTG | Leu/CTT | Leu/CTT |

SEC/SLS to Determine the Oligomerization State of Selected Bundles



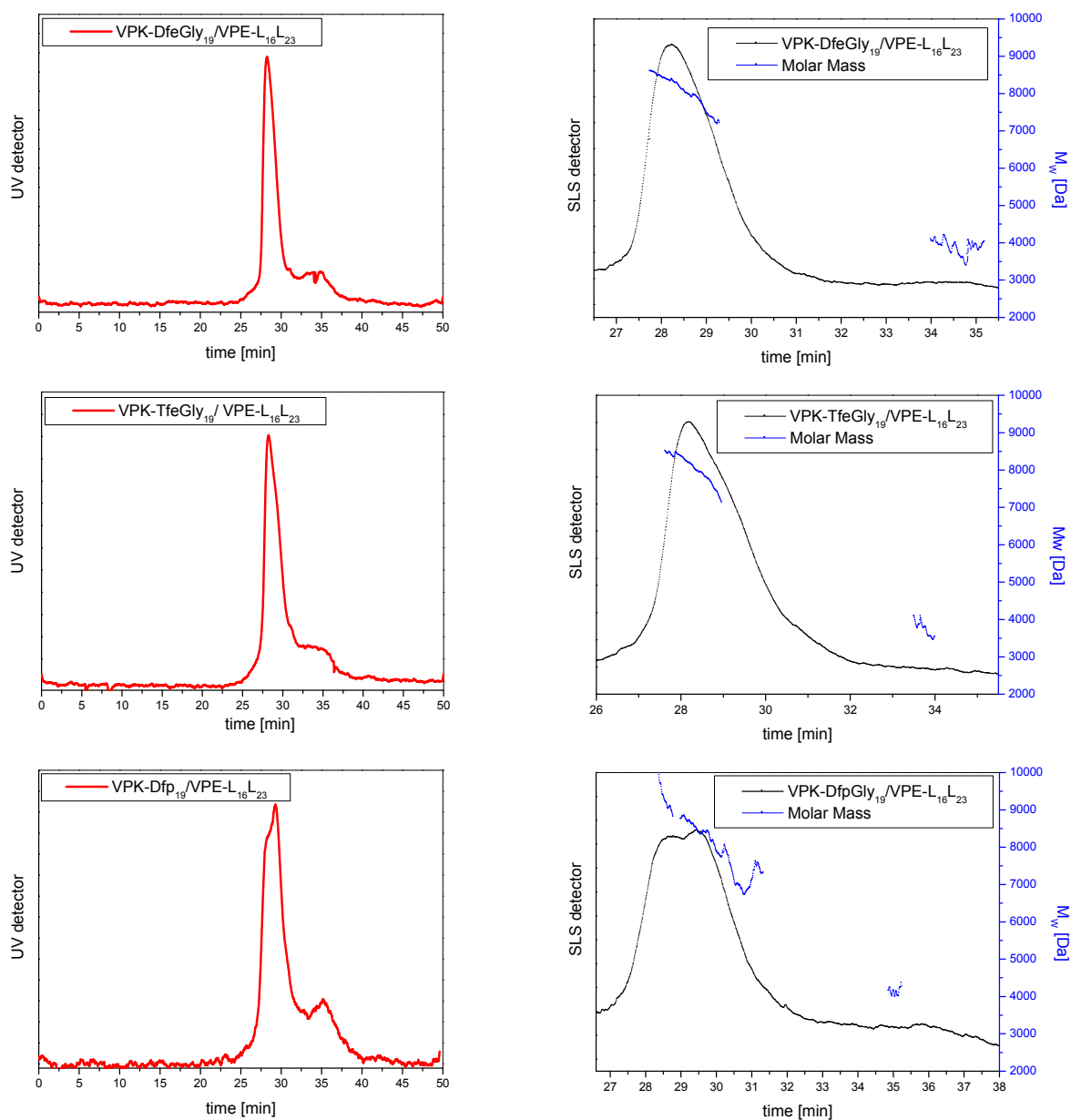


Figure 9.1: SLS chromatograms depicted as the mean of 3 experiments. Respective SEC chromatograms are given to show species singularity. Peptide conc.:60 μ M; eluent: 10 mM PBS; flow rate: 0.3 mL/min; pH 7.4. Measurements were taken from a 1:1 mixture of VPK(X₁₉)/VPE_{L16L23} as indicated in chromatogram.

CD Measurements of Selected Bundles

Biopanning against VPK wild type:

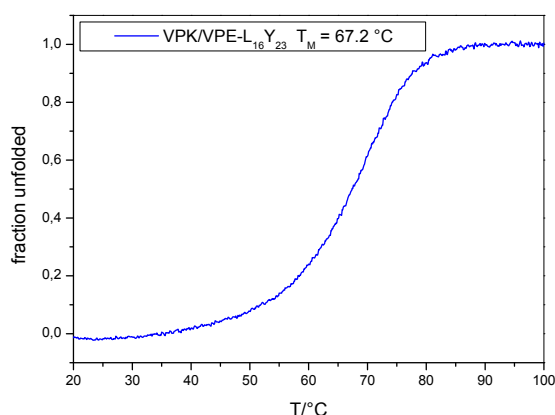


Figure 9.2: Thermal denaturation spectra of equimolar mixtures of **VPK/VPE_{L16Y23}**. The spectra were recorded in phosphate buffer 100 mM, at pH 7.4. Standard deviations from three independent measurements were ± 0.1 °C.

Biopanning against VPK containing fluorinated amino acids at position d_{19} :

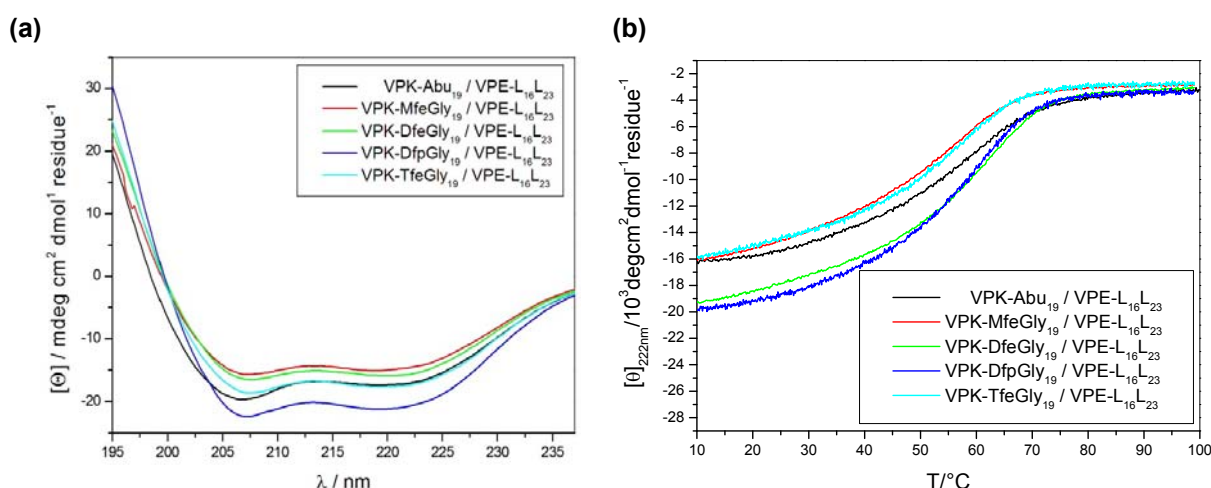


Figure 9.3: (a) CD spectra of equimolar mixtures of **VPK(X₁₉)/VPE_{L16L23}** as depicted in figure legend. (b) Thermal denaturation spectra of equimolar mixtures of **VPK(X₁₉)/VPE_{L16L23}** as depicted in figure legend. The spectra were recorded in phosphate buffer 100 mM, at pH 7.4. Standard deviations from three independent measurements were ± 0.1 °C.

Table 9.3: T_M of **VPK(X₁₉)/VPE_{L16L23}** in comparison to **VPK(X₁₉)/VPE_{WT}**. Values are given in °C.

| | VPE-L _{16L23} | VPE _{WT} | ΔT |
|----------------------------|------------------------|-------------------|------------|
| VPK-Abu ₁₉ : | 57.6 | 53.7 | 3.9 |
| VPK-MfeGly ₁₉ : | 57.2 | 52.0 | 5.2 |
| VPK-DfeGly ₁₉ : | 62.7 | 56.9 | 5.8 |
| VPK-TfeGly ₁₉ : | 58.3 | 55.3 | 3.0 |
| VPK-DfpGly ₁₉ : | 60.1 | 57.5 | 2.6 |

9.2 Phage Display with $\beta\gamma$ -Foldamers

Table 9.4: Sequencing results after test infection of *E. coli* with generated Acid-pp library

| clone | a' ₁₅ | d' ₁₈ | e' ₁₉ | g' ₂₁ |
|-------|------------------|------------------|------------------|------------------|
| 1 | Arg/AGG | Leu/TTG | His/CAC | Arg/AGG |
| 2 | Val/GTT | Ser/TCT | Tyr/TAT | Arg/AGA |
| 3 | Ile/ATT | Asp/GAT | Trp/TGG | Gly/GGT |
| 4 | Gly/GGA | Ser/TCT | Gly/GGC | Asn/AAC |
| 5 | Ser/AGC | Phe/TTT | Phe/TTC | Leu/CTA |
| 6 | Glu/GAA | Arg/CGT | Asn/AAT | Arg/AGG |

The number of clones free from defects in the phage library was calculated from the colony number found through a test infection of an *E. coli* ER2738 culture (2.22×10^8 mutation-free clones). The library screened was sufficient to multiply include every possible combination of amino acids at the three randomized positions (theoretical size: $20^4 = 1.6 \times 10^5$).

9.2.1 Parallel Helix Alignment

Selection against B3 β 2 γ and B3 β 2 γ -variant1

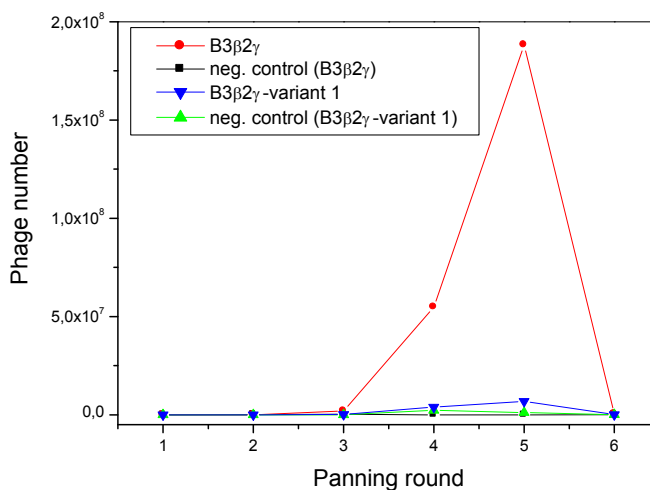


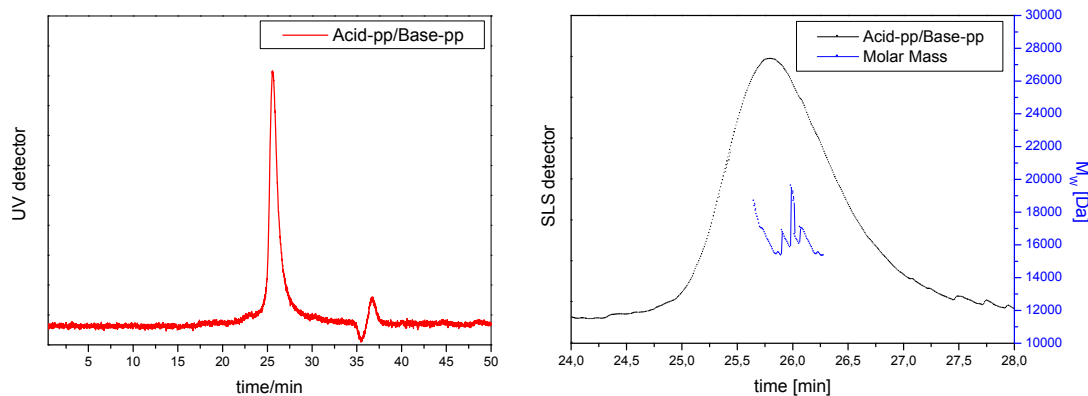
Figure 9.4: Phage enrichment depicted as the phage number throughout each panning round with B3 β 2 γ and B3 β 2 γ -variant1 and negative controls of the respective panning round.

Table 9.5: Sequencing results of mutation-free vectors found post panning with B3 β 2 γ and B3 β 2 γ -variant1 are given.

| Binding partner | Panning round | Amino acids found in randomized positions (Amino acid / DNA codon) | | | |
|----------------------------------|---------------|----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|----------------------------------------------------------------------------------------------------------------------|
| | | Position a' ₁₅ | Position d' ₁₈ | Position e' ₁₉ | Position g' ₂₁ |
| B3 β 2 γ | 4 | Cys / TGT Val / GTT | Phe / TTT Cys / TGT | Leu / TTG Glu / GAG | Glu / GAG Phe / TTT |
| | 5 | Ile / ATT Ile / ATT Cys / TGT | Cys / TGT Cys / TGT Phe / TTT | Glu / GAG Glu / GAG Leu / CTT | Phe / TTT Phe / TTT Asp / GAT |
| | 6 | Cys / TGT Cys / TGT Cys / TGT Cys / TGT Cys / TGT Val / GTT | Phe / TTT Phe / TTT Phe / TTT Phe / TTT Phe / TTT Leu / CTG | Leu / TTG Leu / CTT Leu / CTT Leu / CTG Leu / CTG Leu / TTG | Glu / GAG Glu / GAG Glu / GAG Asp / GAT Asp / GAT Asp / GAT |
| B3 β 2 γ -variant 1 | 4 | Ala / GCT Gln / TAG Leu / CTT Thr / ACG Gly / GGT | Ser / TCT Val / GTG Cys / TGT Ser / AGT Pro / CCG | Gln / TAG Tyr / TAT Asn / AAT Leu / CTG Pro / CCG | Ser / TCG Ala / GCG Val / GTG Val / GTT Ser / TCG |
| | 5 | Arg / CGG | Val / GTT | His / CAT | Gln / TAG |
| | 6 | Thr / ACT Thr / ACG | Arg / CGT Arg / CGG | Leu / TTG Ala / GCG | Leu / TTG Gln / CAG |

Table 9.6: SEC/SLS results of B3 β 2 γ and Base-pp variants. Standard deviations determined from 3 independent measurements.

| | Theoretical Tetramer Mass [Da] | Determined Mass [Da] |
|----------------------------------|--------------------------------|----------------------|
| Acid-pp/Base-pp | 15954 | 16503 \pm 1933 |
| Acid-pp/ B3 β 2 γ | 15804 | 16683 \pm 1356 |
| Acid-CFLE/ B3 β 2 γ | 15878 | 15010 \pm 1521 |
| Acid-ICEF/ B3 β 2 γ | 15878 | 15127 \pm 560 |



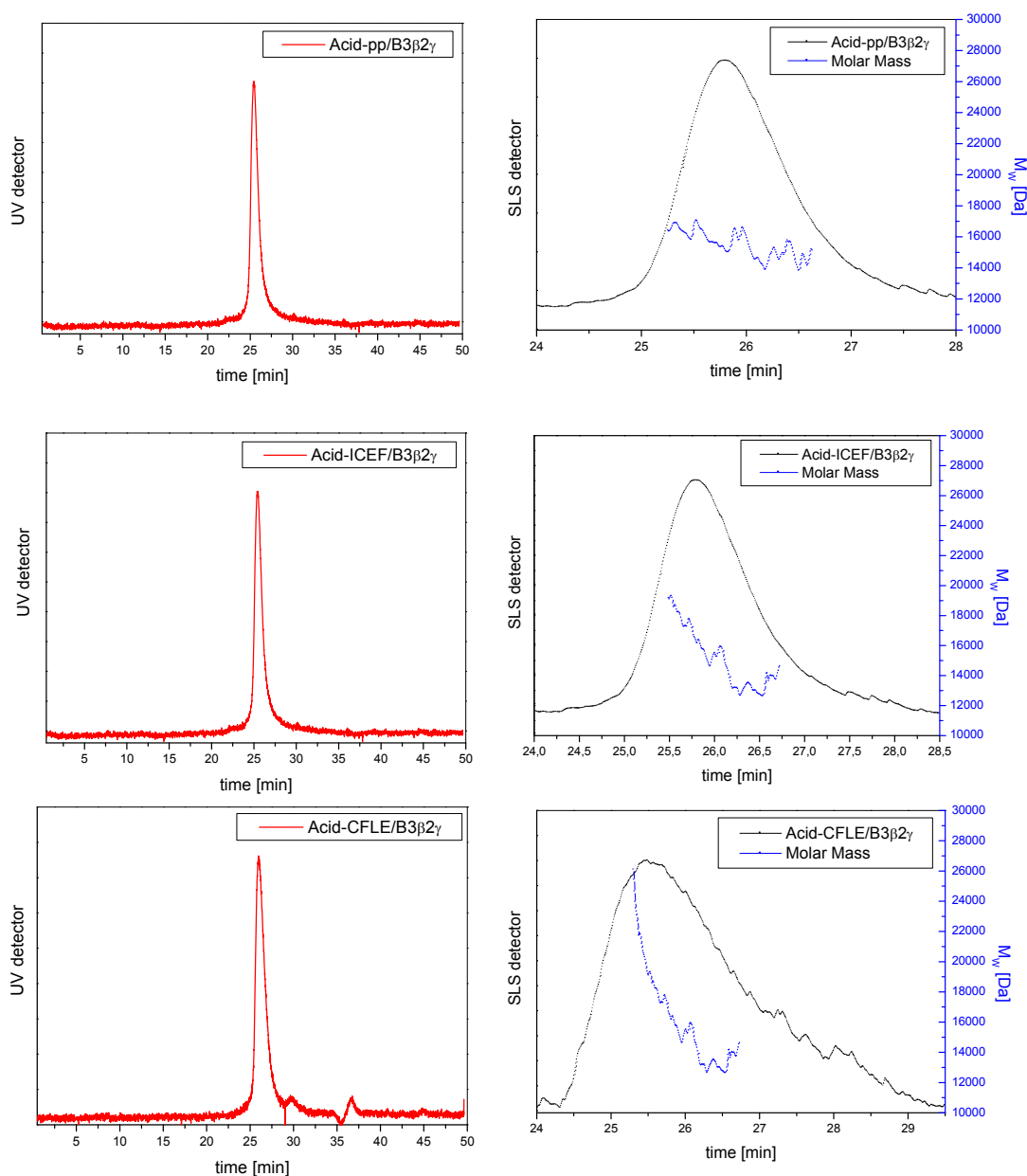


Figure 9.5: SLS chromatograms depicted as the mean of 3 experiments. Respective SEC chromatograms are given to show species singularity. Peptide conc.: 60 μM; eluent: 10 mM PBS; flow rate: 0.3 mL/min; pH 7.4. Measurements were taken from a 1:1 mixture of peptide analogues as indicated in the chromatograms.

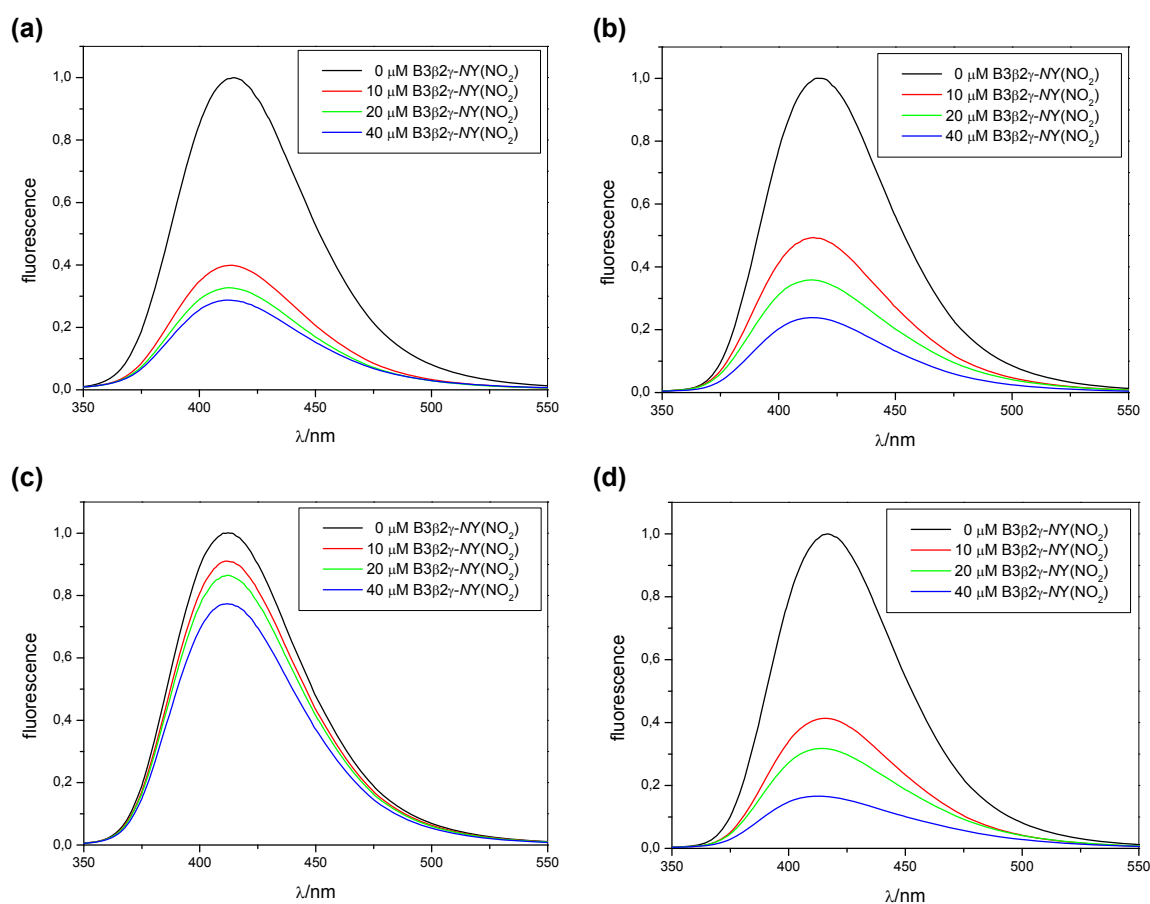


Figure 9.6: Fluorescence spectra of Acid-pp variants at pH 7.4 (10 mM phosphate buffer), in the presence of different concentrations of B3β2γ-NY(NO₂) as depicted in respective figure legend. (a) 20 μM Acid-pp-CAbz; (b) 20 μM Acid-pp-NAbz; (c) 20 μM Acid-ICEF-CAbz; (d) 20 μM Acid-ICEF-NAbz. The spectra were normalized.

As described in the main text the decrease in fluorescence intensity shown in Figure 9.6 is the result of quaternary structure formation. Accordingly, the fluorescence should be recovered upon chemically induced unfolding. Figure 9.7 shows the fluorescence intensity of a mixture of Acid-pp variants and B3β2γ-NY(NO₂) at different concentrations of guanidine hydrochloride. GndHCl was taken from an 8 M stock solution in the corresponding phosphate buffer.

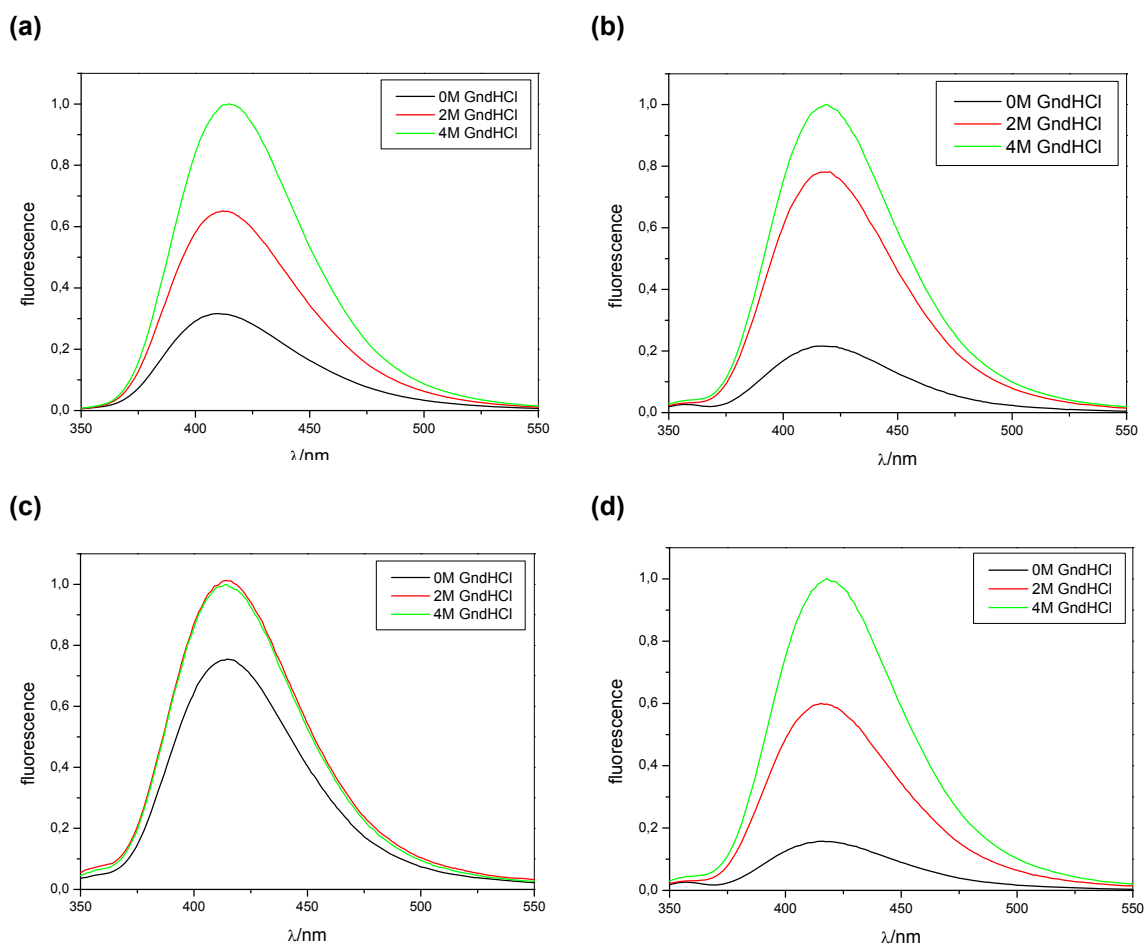


Figure 9.7: Fluorescence spectra at different concentrations of GdnHCl as depicted in figure legend. (a) 20 μ M Acid-pp-CAbz/40 μ M B3 β 2 γ , (b) 20 μ M Acid-pp-NAbz/40 μ M B3 β 2 γ , (c) 20 μ M Acid-ICEF-CAbz/40 μ M B3 β 2 γ , (d) 20 μ M Acid-ICEF-NAbz/40 μ M B3 β 2 γ . The spectra were normalized.

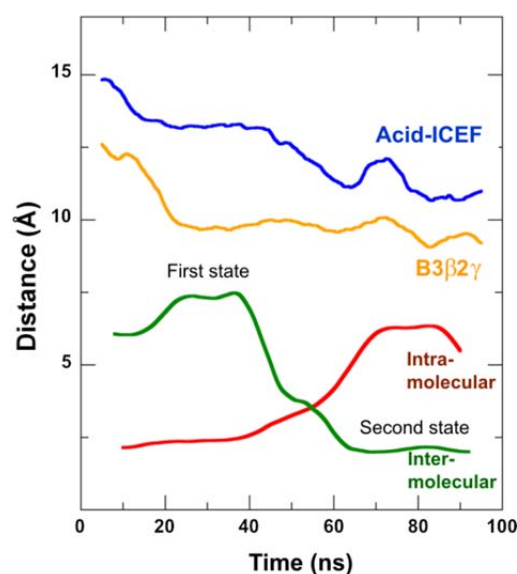
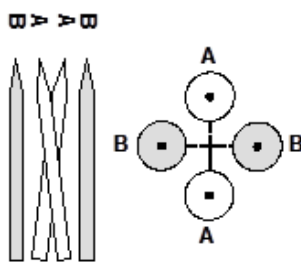


Figure 9.8: Plot illustrating the distances between the two Acid-ICEF helices (blue) and the two B3 β 2 γ helices (orange), SH Cys18 - CO Glu14 (red) and SH Cys18 - CO β L17 (green).

Table 9.7: The calculated distances between the centers of the masses of two basic molecules in Å (B) and two acidic molecules (A) for each studied system are summarized.

| Distances (?) | START | | END | |
|-------------------|-------|-----|-----|-----|
| | B-B | A-A | B-B | A-A |
| Acid-ICEF /B3β2γ | 14 | 14 | 10 | 12 |
| Acid-ICEF /Basepp | 14 | 14 | 12 | 13 |
| Acidpp/B3β2γ | 14 | 14 | 12 | 16 |
| Acidpp/Basepp | 14 | 15 | 12 | 13 |



9.2.2 Antiparallel Helix Alignment

Selection against B3β2γ and B3β2γ-Variant1

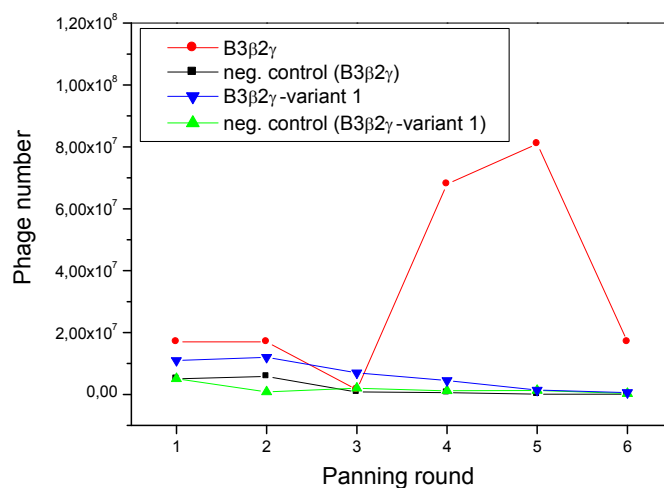


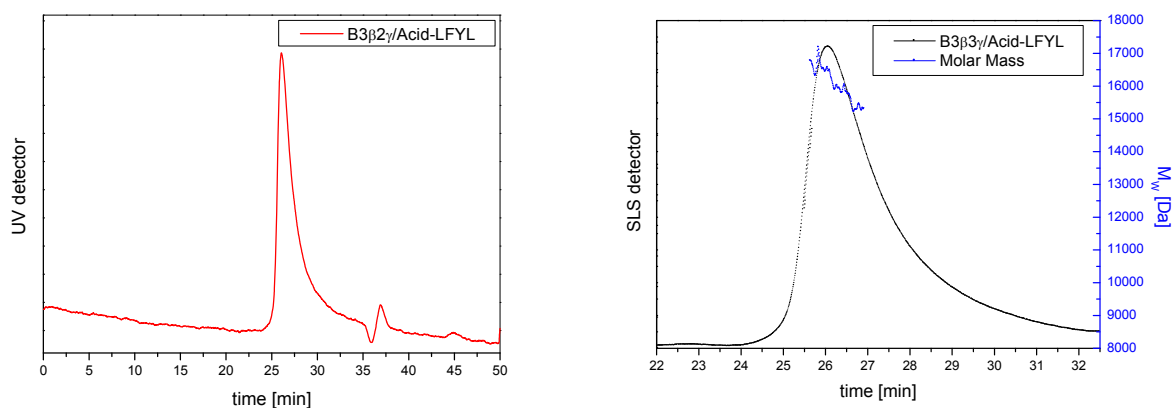
Figure 9.9: Phage enrichment depicted as the phage number throughout each panning round with B3β2γ, B3β2γ-variant1, and negative controls of the respective panning round.

Table 9.8: Sequencing results of mutation-free vectors found post panning with B3β2γ and B3β2γ-variant1 are given.

| Binding partner | Panning round | Amino acids found in randomized positions (Amino acid / DNA codon) | | | |
|-----------------|---------------|--------------------------------------------------------------------|---------------|---------------|---------------|
| | | Position a'15 | Position d'18 | Position e'19 | Position g'21 |
| B3β2γ | 5 | His / CAT | Cys / TGT | Ala / GCG | Asn / ATT |
| | | His / CAT | Cys / TGT | Ala / GCG | Asn / ATT |
| | | His / CAT | Cys / TGT | Ala / GCG | Asn / ATT |
| | | Met / ATG | Thr / ACG | Glu / GAG | Arg / CGT |
| | | Met / ATG | Thr / ACG | Glu / GAG | Arg / CGT |
| | | Leu / TTG | Leu / CTG | Leu / CTG | Leu / CTT |
| | 6 | Leu / TTG | Leu / CTG | Leu / CTG | Leu / CTT |
| | | Leu / TTG | Leu / CTG | Leu / CTG | Leu / CTT |
| | | Leu / TTG | Leu / CTG | Leu / CTG | Leu / CTT |
| | | His / CAT | Cys / TGT | Ala / GCG | Asn / AAT |
| | | His / CAT | Cys / TGT | Ala / GCG | Asn / AAT |
| | | Met / ATG | Thr / ACG | Glu / GAG | Arg / CGT |
| B3β2γ- variant1 | 5 | Leu / CTG | Asp / GAT | Leu / TTG | Gln / TAG |
| | | Gln / TAG | Asp / GAT | Leu / CTG | Val / GTG |
| | 6 | His / CAT | Pro / CCT | Leu / CTG | Pro / CTT |
| | | Ile / ATT | Pro / CCG | Met / ATG | Trp / TGG |

Table 9.9: SEC/SLS results of B3β2γ and Base-pp variants. Standard deviations determined from 3 independent measurements.

| | Theoretical Mass [Da] | Determined Mass [Da] |
|------------------|-----------------------|----------------------|
| Acid-LFYI/ B3β2γ | 15900 (tetramer) | 16603 ± 790 |
| Acid-LLLL/ B3β2γ | 15732 (tetramer) | 15790 ± 682 |
| Acid-HCAN/ B3β2γ | 7838 (dimer) | 8222 ± 508 |
| | 3929 (monomer) | 4108 ± 192 |
| Acid-MTER/ B3β2γ | 7931 (dimer) | 7633 ± 326 |



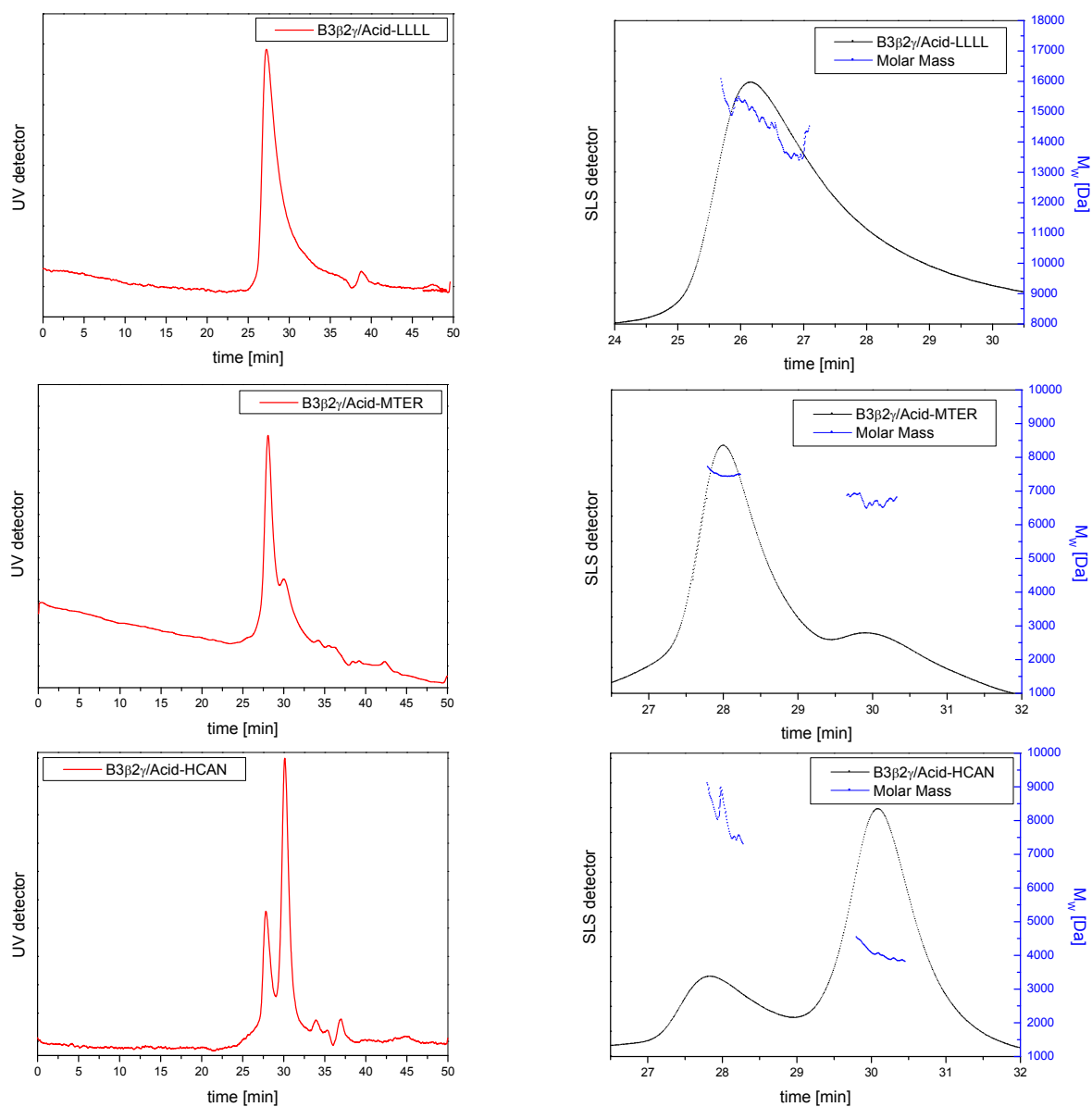


Figure 9.10: SLS chromatograms depicted as the mean of 3 experiments. Respective SEC chromatograms are given to show species singularity. Peptide conc.: 60 μ M; eluent: 10 mM PBS; flow rate: 0.3 mL/min; pH 7.4. Measurements were taken from a 1:1 mixture of peptide analogues as indicated in the chromatograms.

10 Literature

1. Branden, C.; Tooze, J. *Introduction to protein structure*. Garland New York: **1991**; Vol. 2.
2. Voet, D.; Voet, J. G.; Pratt, C. W.; Beck-Sickinger, A.; Hahn, U. *Lehrbuch der Biochemie*. Wiley VCH: **2010**.
3. Stefani, M.; Dobson, C. M. Protein aggregation and aggregate toxicity: new insights into protein folding, misfolding diseases and biological evolution. *Journal of Molecular Medicine* **2003**, 81, 678-699.
4. Eckert, D. M.; Kim, P. S. Mechanisms of viral membrane fusion and its inhibition. *Annual Review of Biochemistry* **2001**, 70, 777-810.
5. Barlow, D.; Thornton, J. Helix geometry in proteins. *Journal of Molecular Biology* **1988**, 201, 601-619.
6. Kim, Y.-W.; Grossmann, T. N.; Verdine, G. L. Synthesis of all-hydrocarbon stapled [alpha]-helical peptides by ring-closing olefin metathesis. *Nature Protocols* **2011**, 6, 761-771.
7. Henchey, L. K.; Jochim, A. L.; Arora, P. S. Contemporary strategies for the stabilization of peptides in the α -helical conformation. *Current Opinion in Chemical Biology* **2008**, 12, 692-697.
8. Hook, D. F.; Gessier, F.; Noti, C.; Kast, P.; Seebach, D. Probing the Proteolytic Stability of β -Peptides Containing α -Fluoro- and α -Hydroxy- β -Amino Acids. *ChemBioChem* **2004**, 5, 691-706.
9. Hackenberger, C. P.; Schwarzer, D. Chemoselective ligation and modification strategies for peptides and proteins. *Angewandte Chemie International Edition* **2008**, 47, 10030-10074.
10. Arnold, U. Incorporation of non-natural modules into proteins: structural features beyond the genetic code. *Biotechnology Letters* **2009**, 31, 1129-1139.
11. Budisa, N. Expanded genetic code for the engineering of ribosomally synthesized and post-translationally modified peptide natural products (RiPPs). *Current Opinion in Biotechnology* **2013**, 4, 591-598.
12. Young, T. S.; Schultz, P. G. Beyond the canonical 20 amino acids: expanding the genetic lexicon. *Journal of Biological Chemistry* **2010**, 285, 11039-11044.
13. Wiltschi, B.; Budisa, N. Natural history and experimental evolution of the genetic code. *Applied Microbiology and Biotechnology* **2007**, 74, 739-753.
14. Salwiczek, M.; Nyakatura, E. K.; Gerling, U. I.; Ye, S.; Koksich, B. Fluorinated amino acids: compatibility with native protein structures and effects on protein-protein interactions. *Chemical Society Reviews* **2012**, 41, 2135-2171.
15. Müller, K.; Faeh, C.; Diederich, F. Fluorine in pharmaceuticals: looking beyond intuition. *Science* **2007**, 317, 1881-1886.
16. Buer, B. C.; de la Salud-Bea, R.; Al Hashimi, H. M.; Marsh, E. N. G. Engineering protein stability and specificity using fluororous amino acids: The importance of packing effects. *Biochemistry* **2009**, 48, 10810-10817.
17. Buer, B. C.; Meagher, J. L.; Stuckey, J. A.; Marsh, E. N. G. Structural basis for the enhanced stability of highly fluorinated proteins. *Proceedings of the National Academy of Sciences* **2012**, 109, 4810-4815.
18. Lee, H.-Y.; Lee, K.-H.; Al-Hashimi, H. M.; Marsh, E. N. G. Modulating protein structure with fluororous amino acids: increased stability and native-like structure conferred on a 4-helix bundle protein by hexafluoroleucine. *Journal of the American Chemical Society* **2006**, 128, 337-343.
19. Lee, K.-H.; Lee, H.-Y.; Slutsky, M. M.; Anderson, J. T.; Marsh, E. N. G. Fluorous effect in proteins: de novo design and characterization of a four- α -helix bundle protein containing hexafluoroleucine. *Biochemistry* **2004**, 43, 16277-16284.
20. Bilgiçer, B.; Fichera, A.; Kumar, K. A coiled coil with a fluororous core. *Journal of the American Chemical Society* **2001**, 123, 4393-4399.
21. Bilgiçer, B.; Kumar, K. De novo design of defined helical bundles in membrane environments. *Proceedings of the National Academy of Sciences of the United States of America* **2004**, 101, 15324-15329.
22. Yoder, N. C.; Kumar, K. Fluorinated amino acids in protein design and engineering. *Chemical Society Reviews* **2002**, 31, 335-341.
23. Jäckel, C.; Salwiczek, M.; Koksich, B. Fluorine in a Native Protein Environment—How the Spatial Demand and Polarity of Fluoroalkyl Groups Affect Protein Folding. *Angewandte Chemie International Edition* **2006**, 45, 4198-4203.

24. Jäckel, C.; Seufert, W.; Thust, S.; Kokschi, B. Evaluation of the Molecular Interactions of Fluorinated Amino Acids with Native Polypeptides. *ChemBioChem* **2004**, *5*, 717-720.
25. Salwiczek, M.; Samsonov, S.; Vagt, T.; Nyakatura, E.; Fleige, E.; Numata, J.; Cölfen, H.; Pisabarro, M. T.; Kokschi, B. Position-Dependent Effects of Fluorinated Amino Acids on the Hydrophobic Core Formation of a Heterodimeric Coiled Coil. *Chemistry – A European Journal* **2009**, *15*, 7628-7636.
26. Cheng, R. P.; Gellman, S. H.; DeGrado, W. F. Beta-peptides: from structure to function. *Chemical Reviews* **2001**, *101*, 3219-3232.
27. Gellman, S. H. Foldamers: A Manifesto. *Accounts of Chemical Research* **1998**, *31*, 173-180.
28. Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. A Field Guide to Foldamers. *Chemical Reviews* **2001**, *101*, 3893-4012.
29. Seebach, D.; Beck, A. K.; Bierbaum, D. J. The World of β - and γ -Peptides Comprised of Homologated Proteinogenic Amino Acids and Other Components. *Chemistry & Biodiversity* **2004**, *1*, 1111-1239.
30. Vasudev, P. G.; Chatterjee, S.; Shamala, N.; Balaram, P. Structural Chemistry of Peptides Containing Backbone Expanded Amino Acid Residues: Conformational Features of beta, gamma, and Hybrid Peptides. *Chemical Reviews* **2011**, *111*, 657-687.
31. Crick, F. H. C. Is [alpha]-Keratin a Coiled Coil? *Nature* **1952**, *170*, 882-883.
32. J. Liu, Q. Z., Y. Deng, C.-S. Cheng, N. R. Kallenbach, M. Lu. A seven-helix coiled coil. *Proceedings of the National Academy of Sciences* **2006**, *103*, 15457-15462.
33. Kim, P. S.; Berger, B.; Wolf, E. MultiCoil: A program for predicting two- and three-stranded coiled coils. *Protein Science* **1997**, *6*, 1179-1189.
34. Moutevelis, E.; Woolfson, D. N. A Periodic Table of Coiled-Coil Protein Structures. *Journal of Molecular Biology* **2009**, *385*, 726-732.
35. Rose, A.; Meier, I. Scaffolds, levers, rods and springs: diverse cellular functions of long coiled-coil proteins. *Cellular and Molecular Life Sciences* **2004**, *61*, 1996-2009.
36. Mason, J. M.; Arndt, K. M. Coiled Coil Domains: Stability, Specificity, and Biological Implications. *ChemBioChem* **2004**, *5*, 170-176.
37. Burkhard, P.; Stetefeld, J.; Strelkov, S. V. Coiled coils: a highly versatile protein folding motif. *Trends in Cell Biology* **2001**, *11*, 82-88.
38. Pauling, L.; Corey, R. Compound helical configurations of polypeptide chains: structure of proteins of the alpha-keratin type. *Nature* **1953**, *171*, 59-61.
39. J. Sodek, R. S. H., L. B. Smillie, L. Jurasek. Amino-Acid Sequence of Rabbit Skeletal Tropomyosin and Its Coiled-Coil Structure. *Proceedings of the National Academy of Sciences* **1972**, *69*, 3800-3804.
40. O'Shea, E. K.; Klemm, J. D.; Kim, P. S.; Alber, T. X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. *Science* **1991**, *254*, 539-544.
41. Cohen, C.; Parry, D. A. D. α -Helical coiled coils and bundles: How to design an α -helical protein. *Proteins: Structure, Function, and Bioinformatics* **1990**, *7*, 1-15.
42. Woolfson, D. N.; David A. D. Parry, a. J. M. S. The Design of Coiled-Coil Structures and Assemblies. In *Advances in Protein Chemistry*, Academic Press: 2005; Vol. Volume 70, pp 79-112.
43. Harbury, P. B.; Zhang, T.; Kim, P. S.; Alber, T. A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. *Science* **1993**, *262*, 1401-1407.
44. Harbury, P. B.; Kim, P. S.; Alber, T. Crystal structure of an isoleucine-zipper trimer. *Nature* **1994**, *371*, 80-83.
45. Monera, O. D.; Zhou, N. E.; Kay, C. M.; Hodges, R. S. Comparison of antiparallel and parallel two-stranded alpha-helical coiled-coils. Design, synthesis, and characterization. *Journal of Biological Chemistry* **1993**, *268*, 19218-19227.
46. Gonzalez, L.; Woolfson, D. N.; Alber, T. Buried polar residues and structural specificity in the GCN4 leucine zipper. *Nature Structural Biology* **1996**, *3*, 1011-1018.
47. Oakley, M. G.; Kim, P. S. A Buried Polar Interaction Can Direct the Relative Orientation of Helices in a Coiled Coil *Biochemistry* **1998**, *37*, 12603-12610.
48. McClain, D. L.; Binfet, J. P.; Oakley, M. G. Evaluation of the energetic contribution of interhelical coulombic interactions for coiled coil helix orientation specificity. *Journal of Molecular Biology* **2001**, *313*, 371-383.

49. Monera, O. D.; Kay, C. M.; Hodges, R. S. Electrostatic Interactions Control the Parallel and Antiparallel Orientation of α -Helical Chains in Two-Stranded α -Helical Coiled-Coils. *Biochemistry* **1994**, *33*, 3862-3871.
50. Zhou, N. E.; Kay, C. M.; Hodges, R. S. The Role of Interhelical Ionic Interactions in Controlling Protein Folding and Stability: De Novo Designed Synthetic Two-stranded α -Helical Coiled-Coils. *Journal of Molecular Biology* **1994**, *237*, 500-512.
51. Kohn, W. D.; Kay, C. M.; Hodges, R. S. Protein destabilization by electrostatic repulsions in the two-stranded α -helical coiled-coil/leucine zipper. *Protein Science* **1995**, *4*, 237-250.
52. Zhou, N. E.; Kay, C. M.; Hodges, R. S. The net energetic contribution of interhelical electrostatic attractions to coiled-coil stability. *Protein Engineering* **1994**, *7*, 1365-1372.
53. Horne, W. S.; Johnson, L. M.; Ketas, T. J.; Klasse, P. J.; Lu, M.; Moore, J. P.; Gellman, S. H. Structural and biological mimicry of protein surface recognition by α/β -peptide foldamers. *Proceedings of the National Academy of Sciences* **2009**, *106*, 14751-14756.
54. Watanabe, S.; Takada, A.; Watanabe, T.; Ito, H.; Kida, H.; Kawaoka, Y. Functional Importance of the Coiled-Coil of the Ebola Virus Glycoprotein. *Journal of Virology* **2000**, *74*, 10194-10201.
55. Harrison, S. C. Mechanism of membrane fusion by viral envelope proteins. *Advances in Virus Research* **2005**, *64*, 231-261.
56. Chan, D. C.; Fass, D.; Berger, J. M.; Kim, P. S. Core Structure of gp41 from the HIV Envelope Glycoprotein. *Cell* **1997**, *89*, 263-273.
57. Chan, D. C.; Kim, P. S. HIV entry and its inhibition. *Cell* **1998**, *93*, 681.
58. Zhang, L.; Huang, Y.; He, T.; Cao, Y.; Ho, D. D. HIV-1 subtype and second-receptor use. *Nature* **1996**, *383*, 768-768.
59. Weissenhorn, W.; Dessen, A.; Harrison, S. C.; Skehel, J. J.; Wiley, D. C. Atomic structure of the ectodomain from HIV-1 gp41. *Nature* **1997**, *387*, 426-430.
60. Buzon, V.; Natrajan, G.; Schibli, D.; Campelo, F.; Kozlov, M. M.; Weissenhorn, W. Crystal structure of HIV-1 gp41 including both fusion peptide and membrane proximal external regions. *PLoS pathogens* *6*, e1000880.
61. Muñoz-Barroso, I.; Durell, S.; Sakaguchi, K.; Appella, E.; Blumenthal, R. Dilation of the human immunodeficiency virus-1 envelope glycoprotein fusion pore revealed by the inhibitory action of a synthetic peptide from gp41. *The Journal of Cell Biology* **1998**, *140*, 315-323.
62. Wild, C. T.; Shugars, D. C.; Greenwell, T. K.; McDanal, C. B.; Matthews, T. J. Peptides corresponding to a predictive alpha-helical domain of human immunodeficiency virus type 1 gp41 are potent inhibitors of virus infection. *Proceedings of the National Academy of Sciences* **1994**, *91*, 9770-9774.
63. Melikyan, G. B.; Markosyan, R. M.; Hemmati, H.; Delmedico, M. K.; Lambert, D. M.; Cohen, F. S. Evidence that the transition of HIV-1 gp41 into a six-helix bundle, not the bundle configuration, induces membrane fusion. *The Journal of Cell Biology* **2000**, *151*, 413-424.
64. Kliger, Y.; Gallo, S. A.; Peisajovich, S. G.; Muñoz-Barroso, I.; Avkin, S.; Blumenthal, R.; Shai, Y. Mode of action of an antiviral peptide from HIV-1 inhibition at a post-lipid mixing stage. *Journal of Biological Chemistry* **2001**, *276*, 1391-1397.
65. Lu, M.; Blacklow, S. C.; Kim, P. S. A trimeric structural domain of the HIV-1 transmembrane glycoprotein. *Nature Structural & Molecular Biology* **1995**, *2*, 1075-1082.
66. Weissenhorn, W.; Wharton, S. A.; Calder, L. J.; Earl, P. L.; Moss, B.; Aliprandis, E.; Skehel, J. J.; Wiley, D. C. The ectodomain of HIV-1 env subunit gp41 forms a soluble, alpha-helical, rod-like oligomer in the absence of gp120 and the N-terminal fusion peptide. *The EMBO journal* **1996**, *15*, 1507.
67. Esté, J. A.; Telenti, A. HIV entry inhibitors. *The Lancet* *370*, 81-88.
68. Bird, G. H.; Madani, N.; Perry, A. F.; Princiotta, A. M.; Supko, J. G.; He, X.; Gavathiotis, E.; Sodroski, J. G.; Walensky, L. D. Hydrocarbon double-stapling remedies the proteolytic instability of a lengthy peptide therapeutic. *Proceedings of the National Academy of Sciences* **2010**, *107*, 14093-14098.
69. Johnson, L. M.; Horne, W. S.; Gellman, S. H. Broad distribution of energetically important contacts across an extended protein interface. *Journal of the American Chemical Society* **2011**, *133*, 10038-10041.
70. Bilgiçer, B.; Kumar, K. Synthesis and thermodynamic characterization of self-sorting coiled coils. *Tetrahedron* **2002**, *58*, 4105-4112.

71. Cheng, R. P.; Girinath, P.; Suzuki, Y.; Kuo, H.-T.; Hsu, H.-C.; Wang, W.-R.; Yang, P.-A.; Gullickson, D.; Wu, C.-H.; Koyack, M. J. Positional effects on helical Ala-based peptides. *Biochemistry* **2010**, *49*, 9372-9384.
72. Tang, Y.; Ghirlanda, G.; Vaidehi, N.; Kua, J.; Mainz, D. T.; Goddard, W. A.; DeGrado, W. F.; Tirrell, D. A. Stabilization of coiled-coil peptide domains by introduction of trifluoroleucine. *Biochemistry* **2001**, *40*, 2790-2796.
73. Cheng, R. P.; DeGrado, W. F. Long-range interactions stabilize the fold of a non-natural oligomer. *Journal of the American Chemical Society* **2002**, *124*, 11564-11565.
74. Price, J. L.; Horne, W. S.; Gellman, S. H. Structural consequences of β -amino acid preorganization in a self-assembling α/β -peptide: fundamental studies of foldameric helix bundles. *Journal of the American Chemical Society* **2010**, *132*, 12378-12387.
75. Qiu, J. X.; Petersson, E. J.; Matthews, E. E.; Schepartz, A. Toward β -amino acid proteins: a cooperatively folded β -peptide quaternary structure. *Journal of the American Chemical Society* **2006**, *128*, 11338-11339.
76. Raguse, T. L.; Lai, J. R.; LePlae, P. R.; Gellman, S. H. Toward β -peptide tertiary structure: Self-association of an amphiphilic 14-helix in aqueous solution. *Organic Letters* **2001**, *3*, 3963-3966.
77. Litowski, J. R.; Hodges, R. S. Designing heterodimeric two-stranded α -helical coiled-coils: the effect of chain length on protein folding, stability and specificity. *Journal of Biological Chemistry* **2002**, *277*, 37272-37279.
78. O'Shea, E. K.; Lumb, K. J.; Kim, P. S. Peptide 'Velcro': design of a heterodimeric coiled coil. *Current Biology* **1993**, *3*, 658-667.
79. Tripet, B.; Wagschal, K.; Lavigne, P.; Mant, C. T.; Hodges, R. S. Effects of side-chain characteristics on stability and oligomerization state of a de Novo-designed model coiled-coil: 20 amino acid substitutions in position "d". *Journal of Molecular Biology* **2000**, *300*, 377-402.
80. Moitra, J.; Szilák, L.; Krylov, D.; Vinson, C. Leucine is the most stabilizing aliphatic amino acid in the d position of a dimeric leucine zipper coiled coil. *Biochemistry* **1997**, *36*, 12567-12573.
81. Chou, P. Y.; Fasman, G. D. Conformational parameters for amino acids in helical, β -sheet, and random coil regions calculated from proteins. *Biochemistry* **1974**, *13*, 211-222.
82. Chou, P. Y.; Fasman, G. D. Prediction of protein conformation. *Biochemistry* **1974**, *13*, 222-245.
83. Araghi, R. R. α -Helical Coiled Coil Mimicry by Alternating Sequences of β - and γ -Amino Acids. Doctoral Dissertation, Freie Universität Berlin, 2011.
84. Rezaei Araghi, R.; Jäckel, C.; Cölfen, H.; Salwiczek, M.; Völkel, A.; Wagner, S. C.; Wieczorek, S.; Baldauf, C.; Kokschi, B. A β/γ Motif to Mimic α -Helical Turns in Proteins. *ChemBioChem* **2010**, *11*, 335-339.
85. Merkel, L.; Budisa, N. Organic fluorine as a polypeptide building element: in vivo expression of fluorinated peptides, proteins and proteomes. *Organic & Biomolecular Chemistry* **2012**, *10*, 7241-7261.
86. Arkin, M. R.; Wells, J. A. Small-molecule inhibitors of protein-protein interactions: progressing towards the dream. *Nature Reviews Drug Discovery* **2004**, *3*, 301-317.
87. Tamara L. Hendrickson; Valérie de Crécy-Lagard; Schimmel, P. Incorporation of nonnatural amino acids into proteins. *Annual Review of Biochemistry* **2004**, *73*, 147-176.
88. Smart, B. E. Fluorine substituent effects (on bioactivity). *Journal of Fluorine Chemistry* **2001**, *109*, 3-11.
89. Silverstein, T. P. The Real Reason Why Oil and Water Don't Mix. *Journal of Chemical Education* **1998**, *75*, 116.
90. Kauzmann, W.; C.B. Anfinsen, M. L. A. K. B. a. J. T. E. Some Factors in the Interpretation of Protein Denaturation. In *Advances in Protein Chemistry*, Academic Press: 1959; Vol. Volume 14, pp 1-63.
91. Pace, C. N.; Shirley, B. A.; McNutt, M.; Gajiwala, K. Forces contributing to the conformational stability of proteins. *The FASEB Journal* **1996**, *10*, 75-83.
92. Salonen, L. M.; Ellermann, M.; Diederich, F. Aromatic rings in chemical and biological recognition: energetics and structures. *Angewandte Chemie International Edition* **2011**, *50*, 4808-4842.
93. Pace, C. J.; Gao, J. Exploring and Exploiting Polar- π Interactions with Fluorinated Aromatic Amino Acids. *Accounts of Chemical Research* **2012**.

94. Bott, G.; Field, L. D.; Sternhell, S. Steric effects. A study of a rationally designed system. *Journal of the American Chemical Society* **1980**, 102, 5618-5626.
95. Nagai, T.; Nishioka, G.; Koyama, M.; Ando, A.; Miki, T.; Kumadaki, I. Reactions of trifluoromethyl ketones. IX. Investigation of the steric effect of a trifluoromethyl group based on the stereochemistry on the dehydration of trifluoromethyl homoallyl alcohols. *Journal of Fluorine Chemistry* **1992**, 57, 229-237.
96. Samsonov, S. A.; Salwiczek, M.; Anders, G.; Koksich, B.; Pisabarro, M. T. Fluorine in protein environments: a QM and MD study. *The Journal of Physical Chemistry B* **2009**, 113, 16400-16408.
97. Salwiczek, M. Biophysical Aspects of Single Fluoroalkylamino Acid Substitutions within a Natural Polypeptide Environment. Doctoral Dissertation, Freie Universität Berlin, Berlin, 2010.
98. Kovacs, J. M.; Mant, C. T.; Hodges, R. S. Determination of intrinsic hydrophilicity/hydrophobicity of amino acid side chains in peptides in the absence of nearest-neighbor or conformational effects. *Peptide Science* **2006**, 84, 283-297.
99. Modeling and molecular dynamics simulations conducted by Dr. Jérémie Mortier. In 2011-2013.
100. Woll, M. G.; Hadley, E. B.; Mecozzi, S.; Gellman, S. H. Stabilizing and destabilizing effects of phenylalanine → F5-phenylalanine mutations on the folding of a small protein. *Journal of the American Chemical Society* **2006**, 128, 15932-15933.
101. Pauling, L. *The Nature of the Chemical Bond. Chap 2*. Cornell University Press: Ithaca, New York, **1939**.
102. Neil, E.; Marsh, G. Towards the nonstick egg: designing fluororous proteins. *Chemistry & Biology* **2000**, 7, 153-157.
103. O'Hagan, D. Understanding organofluorine chemistry. An introduction to the C-F bond. *Chemical Society Reviews* **2008**, 37, 308-319.
104. Danielson, M. A.; Falke, J. J. Use of ¹⁹F NMR to probe protein structure and conformational changes. *Annual Review of Biophysics and Biomolecular Structure* **1996**, 25, 163.
105. Zheng, H.; Comeforo, K.; Gao, J. Expanding the Fluorous Arsenal: Tetrafluorinated Phenylalanines for Protein Design. *Journal of the American Chemical Society* **2008**, 131, 18-19.
106. Hunter, C. A.; Sanders, J. K. M. The nature of pi-pi interactions. *Journal of the American Chemical Society* **1990**, 112, 5525-5534.
107. Zheng, H.; Gao, J. Highly Specific Heterodimerization Mediated by Quadrupole Interactions. *Angewandte Chemie International Edition* **2010**, 49, 8635-8639.
108. Ponzini, F.; Zagha, R.; Hardcastle, K.; Siegel, J. S. Phenyl/Pentafluorophenyl Interactions and the Generation of Ordered Mixed Crystals: sym-Triphenethynylbenzene and sym-Tris(perfluorophenethynyl)benzene *Angewandte Chemie International Edition* **2000**, 39, 2323-2325.
109. Waters, M. L. Aromatic interactions in model systems. *Current Opinion in Chemical Biology* **2002**, 6, 736-741.
110. West, A. P.; Mecozzi, S.; Dougherty, D. A. Theoretical Studies of the Supramolecular Synthon Benzene ··· Hexafluorobenzene. *Journal of Physical Organic Chemistry* **1997**, 10, 347-350.
111. Chiu, H.-P.; Cheng, R. P. Chemoenzymatic synthesis of (S)-hexafluoroleucine and (S)-tetrafluoroleucine. *Organic Letters* **2007**, 9, 5517-5520.
112. Chiu, H.-P.; Suzuki, Y.; Gullickson, D.; Ahmad, R.; Kokona, B.; Fairman, R.; Cheng, R. P. Helix propensity of highly fluorinated amino acids. *Journal of the American Chemical Society* **2006**, 128, 15556-15557.
113. Son, S.; Tanrikulu, I. C.; Tirrell, D. A. Stabilization of bzip peptides through incorporation of fluorinated aliphatic residues. *ChemBioChem* **2006**, 7, 1251-1257.
114. Tang, Y.; Ghirlanda, G.; Petka, W. A.; Nakajima, T.; DeGrado, W. F.; Tirrell, D. A. Fluorinated coiled-coil proteins prepared in vivo display enhanced thermal and chemical stability. *Angewandte Chemie* **2001**, 113, 1542-1544.
115. Tang, Y.; Tirrell, D. A. Biosynthesis of a highly stable coiled-coil protein containing hexafluoroleucine in an engineered bacterial host. *Journal of the American Chemical Society* **2001**, 123, 11089-11090.
116. Pendley, S. S.; Yu, Y. B.; Cheatham, T. E. Molecular dynamics guided study of salt bridge length dependence in both fluorinated and non-fluorinated parallel dimeric coiled-coils. *Proteins: Structure, Function, and Bioinformatics* **2009**, 74, 612-629.

117. Waters, M. L. Aromatic interactions in peptides: impact on structure and function. *Peptide Science* **2004**, 76, 435-445.
118. Robson Marsden, H.; Fraaije, J. G. E. M.; Kros, A. Introducing Quadrupole Interactions into the Peptide Design Toolkit. *Angewandte Chemie International Edition* **2010**, 49, 8570-8572.
119. Cejas, M. A.; Kinney, W. A.; Chen, C.; Vinter, J. G.; Almond, H. R.; Balss, K. M.; Maryanoff, C. A.; Schmidt, U.; Breslav, M.; Mahan, A.; Lacy, E.; Maryanoff, B. E. Thrombogenic collagen-mimetic peptides: Self-assembly of triple helix-based fibrils driven by hydrophobic interactions. *Proceedings of the National Academy of Sciences* **2008**, 105, 8513-8518.
120. Nieswandt, B.; Watson, S. P. Platelet-collagen interaction: is GPVI the central receptor? *Blood* **2003**, 102, 449-61.
121. Appella, D. H.; Christianson, L. A.; Karle, I. L.; Powell, D. R.; Gellman, S. H. beta-peptide foldamers: Robust Helix formation in a new family of beta-amino acid oligomers. *Journal of the American Chemical Society* **1996**, 118, 13071-13072.
122. Horne, W. S.; Gellman, S. H. Foldamers with heterogeneous backbones. *Accounts of chemical research* **2008**, 41, 1399-1408.
123. Venkatraman, J.; Shankaramma, S. C.; Balaram, P. Design of folded peptides. *Chemical Reviews* **2001**, 101, 3131-3152.
124. Seebach, D.; Overhand, M.; Kühnle, F. N.; Martinoni, B.; Oberer, L.; Hommel, U.; Widmer, H. β -Peptides: Synthesis by Arndt-Eistert homologation with concomitant peptide coupling. Structure determination by NMR and CD spectroscopy and by X-ray crystallography. Helical secondary structure of a β -hexapeptide in solution and its stability towards pepsin. *Helvetica Chimica Acta* **1996**, 79, 913-941.
125. Hintermann, T.; Gademann, K.; Jaun, B.; Seebach, D. γ -Peptides Forming More Stable Secondary Structures than α -Peptides: Synthesis and helical NMR-solution structure of the γ -hexapeptide analog of H-(Val-Ala-Leu) 2-OH. *Helvetica Chimica Acta* **1998**, 81, 983-1002.
126. Applequist, J.; Glickson, J. Conformation of poly- β -alanine in aqueous solution from proton magnetic resonance and deuterium exchange studies. *Journal of the American Chemical Society* **1971**, 93, 3276-3281.
127. Goodman, C. M.; Choi, S.; Shandler, S.; DeGrado, W. F. Foldamers as versatile frameworks for the design and evolution of function. *Nature Chemical Biology* **2007**, 3, 252-262.
128. Seebach, D.; Hook, D. F.; Glättli, A. Helices and other secondary structures of β - and γ -peptides. *Peptide Science* **2006**, 84, 23-37.
129. Daniels, D. S.; Petersson, E. J.; Qiu, J. X.; Schepartz, A. High-resolution structure of a β -peptide bundle. *Journal of the American Chemical Society* **2007**, 129, 1532-1533.
130. Ananda, K.; Vasudev, P. G.; Sengupta, A.; Poopathi Raja, K. M.; Shamala, N.; Balaram, P. Polypeptide helices in hybrid peptide sequences. *Journal of the American Chemical Society* **2005**, 127, 16668-16674.
131. Baldauf, C.; Günther, R.; Hofmann, H.-J. Helix Formation in α , γ - and β , γ -Hybrid Peptides: Theoretical Insights into Mimicry of α - and β -Peptides. *The Journal of Organic Chemistry* **2006**, 71, 1200-1208.
132. Choi, S. H.; Guzei, I. A.; Gellman, S. H. Crystallographic characterization of the α/β -peptide 14/15-helix. *Journal of the American Chemical Society* **2007**, 129, 13780-13781.
133. Molski, M. A.; Goodman, J. L.; Craig, C. J.; Meng, H.; Kumar, K.; Schepartz, A. β -Peptide Bundles with Fluorous Cores. *Journal of the American Chemical Society* **2010**, 132, 3658-3659.
134. Horne, W. S.; Price, J. L.; Keck, J. L.; Gellman, S. H. Helix bundle quaternary structure from α/β -peptide foldamers. *Journal of the American Chemical Society* **2007**, 129, 4178-4180.
135. Horne, W. S.; Boersma, M. D.; Windsor, M. A.; Gellman, S. H. Sequence-Based Design of α/β -Peptide Foldamers That Mimic BH3 Domains. *Angewandte Chemie International Edition* **2008**, 47, 2853-2856.
136. Qian, H.; Schellman, J. A. Helix-coil theories: a comparative study for finite length polypeptides. *The Journal of Physical Chemistry* **1992**, 96, 3987-3994.
137. Scholtz, J. M.; Qian, H.; Robbins, V. H.; Baldwin, R. L. The energetics of ion-pair and hydrogen-bonding interactions in a helical peptide. *Biochemistry* **1993**, 32, 9668-9676.
138. Phelan, J. C.; Skelton, N. J.; Braisted, A. C.; McDowell, R. S. A general method for constraining short peptides to an α -helical conformation. *Journal of the American Chemical Society* **1997**, 119, 455-460.

139. Jackson, D. Y.; King, D. S.; Chmielewski, J.; Singh, S.; Schultz, P. G. General approach to the synthesis of short. α -helical peptides. *Journal of the American Chemical Society* **1991**, *113*, 9391-9392.
140. Vijayalakshmi, S.; Rao, R. B.; Karle, I.; Balaram, P. Comparison of helix-stabilizing effects of α , α -dialkyl glycines with linear and cycloalkyl side chains. *Biopolymers* **2000**, *53*, 84-98.
141. Blackwell, H. E.; Grubbs, R. H. Highly efficient synthesis of covalently cross-linked peptide helices by ring-closing metathesis. *Angewandte Chemie International Edition* **1998**, *37*, 3281-3284.
142. Schafmeister, C. E.; Po, J.; Verdine, G. L. An all-hydrocarbon cross-linking system for enhancing the helicity and metabolic stability of peptides. *Journal of the American Chemical Society* **2000**, *122*, 5891-5892.
143. Fletcher, S.; Hamilton, A. D. Targeting protein-protein interactions by rational design: mimicry of protein surfaces. *Journal of the Royal Society Interface* **2006**, *3*, 215-233.
144. Kim, Y.-W.; Kutchukian, P. S.; Verdine, G. L. Introduction of all-hydrocarbon i, i+ 3 staples into α -helices via ring-closing olefin metathesis. *Organic Letters* **2010**, *12*, 3046-3049.
145. Bernal, F.; Tyler, A. F.; Korsmeyer, S. J.; Walensky, L. D.; Verdine, G. L. Reactivation of the p53 tumor suppressor pathway by a stapled p53 peptide. *Journal of the American Chemical Society* **2007**, *129*, 2456-2457.
146. Walensky, L. D.; Kung, A. L.; Escher, I.; Malia, T. J.; Barbuto, S.; Wright, R. D.; Wagner, G.; Verdine, G. L.; Korsmeyer, S. J. Activation of apoptosis in vivo by a hydrocarbon-stapled BH3 helix. *Science Signaling* **2004**, *305*, 1466.
147. Danial, N. N.; Walensky, L. D.; Zhang, C.-Y.; Choi, C. S.; Fisher, J. K.; Molina, A. J.; Datta, S. R.; Pitter, K. L.; Bird, G. H.; Wikstrom, J. D. Dual role of proapoptotic BAD in insulin secretion and beta cell survival. *Nature Medicine* **2008**, *14*, 144-153.
148. Moellering, R. E.; Cornejo, M.; Davis, T. N.; Del Bianco, C.; Aster, J. C.; Blacklow, S. C.; Kung, A. L.; Gilliland, D. G.; Verdine, G. L.; Bradner, J. E. Direct inhibition of the NOTCH transcription factor complex. *Nature* **2009**, *462*, 182-188.
149. Patgiri, A.; Jochim, A. L.; Arora, P. S. A hydrogen bond surrogate approach for stabilization of short peptide sequences in α -helical conformation. *Accounts of chemical research* **2008**, *41*, 1289-1300.
150. Wang, D.; Liao, W.; Arora, P. S. Enhanced Metabolic Stability and Protein-Binding Properties of Artificial α Helices Derived from a Hydrogen-Bond Surrogate: Application to Bcl-xL. *Angewandte Chemie International Edition* **2005**, *44*, 6525-6529.
151. Wang, D.; Lu, M.; Arora, P. S. Inhibition of HIV-1 Fusion by Hydrogen-Bond-Surrogate-Based α Helices. *Angewandte Chemie International Edition* **2008**, *47*, 1879-1882.
152. Barbas III, C. F.; Burton, D. R.; Scott, J. K.; Silverman, G. J. *Phage Display: A Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press: New York, 2001.
153. Yuan, L.; Kurek, I.; English, J.; Keenan, R. Laboratory-directed protein evolution. *Microbiology and Molecular Biology Reviews* **2005**, *69*, 373-392.
154. Brustad, E. M.; Arnold, F. H. Optimizing non-natural protein function with directed evolution. *Current Opinion in Chemical Biology* *15*, 201-210.
155. FitzGerald, K. In vitro display technologies - new tools for drug discovery. *Drug Discovery Today* **2000**, *5*, 253-258.
156. Schaffitzel, C.; Plückthun, A. Protein-fold evolution in the test tube. *Trends in Biochemical Sciences* **2001**, *26*, 577-579.
157. Jung, S.; Honegger, A.; Plückthun, A. Selection for improved protein stability by phage display. *Journal of Molecular Biology* **1999**, *294*, 163-180.
158. Roberts, R. W.; Szostak, J. W. RNA-peptide fusions for the in vitro selection of peptides and proteins. *Proceedings of the National Academy of Sciences* **1997**, *94*, 12297-12302.
159. Smith, G. P. Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **1985**, *228*, 1315-1317.
160. Smith, G. P.; Petrenko, V. A. Phage display. *Chemical Reviews* **1997**, *97*, 391-410.
161. Kehoe, J. W.; Kay, B. K. Filamentous phage display in the new millennium. *Chemical Reviews* **2005**, *105*, 4056-4072.
162. Vagt, T.; Jäckel, C.; Samsonov, S.; Teresa Pisabarro, M.; Kokschi, B. Selection of a buried salt bridge by phage display. *Bioorganic & Medicinal Chemistry Letters* **2009**, *19*, 3924-3927.

163. Hagemann, U. B.; Mason, J. M.; Müller, K. M.; Arndt, K. M. Selectional and mutational scope of peptides sequestering the Jun/Fos coiled-coil domain. *Journal of Molecular Biology* **2008**, 381, 73-88.
164. Lai, J. R.; Fisk, J. D.; Weisblum, B.; Gellman, S. H. Hydrophobic core repacking in a coiled-coil dimer via phage display: Insights into plasticity and specificity at a protein-protein interface. *Journal of the American Chemical Society* **2004**, 126, 10514-10515.
165. Matsuura, T.; Yomo, T. In vitro evolution of proteins. *Journal of Bioscience and Bioengineering* **2006**, 101, 449-456.
166. Marvin, D. A. Filamentous phage structure, infection and assembly. *Current Opinion in Structural Biology* **1998**, 8, 150-158.
167. Terry, T. D.; Malik, P.; Perham, R. N. Accessibility of Peptides Displayed on Filamentous Bacteriophage Virions: Susceptibility to Proteinases. *Biological Chemistry* **1997**, 378, 523-530.
168. Lubkowski, J.; Hennecke, F.; Plückthun, A.; Wlodawer, A. Filamentous phage infection: crystal structure of g3p in complex with its coreceptor, the C-terminal domain of TolA. *Structure* **1999**, 7, 711-722.
169. Stassen, A. P. M.; Folmer, R. H. A.; Hilbers, C. W.; Konings, R. N. H. Single-stranded DNA binding protein encoded by the filamentous bacteriophage M13: structural and functional characteristics. *Molecular Biology Reports* **1994**, 20, 109-127.
170. Guan, Y.; Zhang, H.; Wang, A. H. Electrostatic potential distribution of the gene V protein from Ff phage facilitates cooperative DNA binding: A model of the GVP-ssDNA complex. *Protein Science* **1995**, 4, 187-197.
171. Russel, M.; Model, P. Genetic analysis of the filamentous bacteriophage packaging signal and of the proteins that interact with it. *Journal of Virology* **1989**, 63, 3284-3295.
172. Gao, C.; Mao, S.; Lo, C.-H. L.; Wirsching, P.; Lerner, R. A.; Janda, K. D. Making artificial antibodies: a format for phage display of combinatorial heterodimeric arrays. *Proceedings of the National Academy of Sciences* **1999**, 96, 6025-6030.
173. Gao, C.; Mao, S.; Kaufmann, G.; Wirsching, P.; Lerner, R. A.; Janda, K. D. A method for the generation of combinatorial antibody libraries using pIX phage display. *Proceedings of the National Academy of Sciences* **2002**, 99, 12612-12616.
174. Hufton, S. E.; Moerkerk, P. T.; Meulemans, E. V.; de Bruine, A.; Arends, J.-W.; Hoogenboom, H. R. Phage display of cDNA repertoires: the pVI display system and its applications for the selection of immunogenic ligands. *Journal of Immunological Methods* **1999**, 231, 39-51.
175. Zhong, G.; Smith, G. P.; Berry, J.; Brunham, R. C. Conformational mimicry of a chlamydial neutralization epitope on filamentous phage. *Journal of Biological Chemistry* **1994**, 269, 24183-24188.
176. McConnell, S. J.; Hoess, R. H. Tendamistat as a scaffold for conformationally constrained phage peptide libraries. *Journal of Molecular Biology* **1995**, 250, 460-470.
177. Jäckel, C. Entwicklung eines Screening-Systems zur systematischen Untersuchung der Wechselwirkungseigenschaften fluoralkyl-substituierter Aminosäuren mit Polypeptiden. Freie Universität Berlin, Dissertation 2006.
178. Vagt, D.-B. T. Entwicklung eines coiled coil-basierten Screeningsystems zur Bestimmung spezifischer Wechselwirkungspartner fluoralkylsubstituierter Aminosäuren. Freie Universität Berlin, Dissertation 2009.
179. Barbas, C. F.; Kang, A. S.; Lerner, R. A.; Benkovic, S. J. Assembly of combinatorial antibody libraries on phage surfaces: the gene III site. *Proceedings of the National Academy of Sciences* **1991**, 88, 7978-7982.
180. Wiesehan, K.; Buder, K.; Linke, R. P.; Patt, S.; Stoldt, M.; Unger, E.; Schmitt, B.; Bucci, E.; Willbold, D. Selection of D-Amino-Acid Peptides That Bind to Alzheimer's Disease Amyloid Peptide A β 1-42 by Mirror Image Phage Display. *ChemBioChem* **2003**, 4, 748-753.
181. Field, G.; Noble, R. Solid peptide synthesis utilizing 9-fluorenylmethoxycarbonyl amino acid. *International Journal of Peptide and Protein Research* **1990**, 35, 161-214.
182. Piran, U.; Riordan, W. J. Dissociation rate constant of the biotin-streptavidin complex. *Journal of Immunological Methods* **1990**, 133, 141-143.
183. Vagt, T.; Nyakatura, E.; Salwiczek, M.; Jäckel, C.; Koksche, B. Towards identifying preferred interaction partners of fluorinated amino acids within the hydrophobic environment of a dimeric coiled coil peptide. *Organic & Biomolecular Chemistry* **2010**, 8, 1382-1386.

184. Lupas, A.; Van Dyke, M.; Stock, J. Predicting coiled coils from protein sequences. *Science* **1991**, 252, 1162.
185. Wagschal, K.; Tripet, B.; Mant, C.; Hodges, R. S.; Lavigne, P. The role of position a in determining the stability and oligomerization state of α -helical coiled coils: 20 amino acid stability coefficients in the hydrophobic core of proteins. *Protein Science* **1999**, 8, 2312-2329.
186. Reimann, O. Identification of preferred interaction partners for fluoroalkyl-substituted amino acids in a protein environment by phage display and their thermal characterization. Freie Universität Berlin, Master of Science Thesis 2011.
187. Padmanabhan, S.; Baldwin, R. L. Helix-stabilizing interaction between tyrosine and leucine or valine when the spacing is $i, i+4$. *Journal of Molecular Biology* **1994**, 241, 706-713.
188. McGregor, M. J.; Islam, S. A.; Sternberg, M. J. Analysis of the relationship between side-chain conformation and secondary structure in globular proteins. *Journal of molecular biology* **1987**, 198, 295-310.
189. Zhao, Y. H.; Abraham, M. H.; Zissimos, A. M. Fast calculation of van der Waals volume as a sum of atomic and bond contributions and its application to drug compounds. *The Journal of Organic Chemistry* **2003**, 68, 7368-7373.
190. Minks, C.; Huber, R.; Moroder, L.; Budisa, N. Atomic Mutations at the Single Tryptophan Residue of Human Recombinant Annexin V: Effects on Structure, Stability, and Activity. *Biochemistry* **1999**, 38, 10649-10659.
191. Steiner, T.; Hess, P.; Bae, J. H.; Wiltschi, B.; Moroder, L.; Budisa, N. Synthetic biology of proteins: tuning GFPs folding and stability with fluoroproline. *PLoS One* **2008**, 3, e1680.
192. Wieczorek, S. Selecting preferred interaction partners of homologated amino acids in a chimeric coiled coil motif: An application of phage display. Freie Universität Berlin, Master of Science Thesis 2010.
193. Rezaei Araghi, R. personal communication. 2012.
194. Duus, J.; Meldal, M.; Winkler, J. R. Fluorescence energy-transfer probes of conformation in peptides: the 2-aminobenzamide/nitrotyrosine pair. *The Journal of Physical Chemistry B* **1998**, 102, 6413-6418.
195. Gray, T.; Matthews, B. Intrahelical hydrogen bonding of serine, threonine and cysteine residues within α -helices and its relevance to membrane-bound proteins. *Journal of Molecular Biology* **1984**, 175, 75-81.
196. Gregoret, L. M.; Rader, S. D.; Fletterick, R. J.; Cohen, F. E. Hydrogen bonds involving sulfur atoms in proteins. *Proteins: Structure, Function, and Bioinformatics* **1991**, 9, 99-107.
197. Kyte, J.; Doolittle, R. F. A simple method for displaying the hydropathic character of a protein. *Journal of Molecular Biology* **1982**, 157, 105-132.
198. Nagano, N.; Ota, M.; Nishikawa, K. Strong hydrophobic nature of cysteine residues in proteins. *FEBS letters* **1999**, 458, 69-71.
199. Zhou, G.; Wu, D.; Snyder, B.; Ptak, R. G.; Kaur, H.; Gochin, M. Development of indole compounds as small molecule fusion inhibitors targeting HIV-1 glycoprotein-41. *Journal of Medicinal Chemistry* **2011**, 54, 7220-7231.
200. He, Y.; Cheng, J.; Lu, H.; Li, J.; Hu, J.; Qi, Z.; Liu, Z.; Jiang, S.; Dai, Q. Potent HIV fusion inhibitors against Enfuvirtide-resistant HIV-1 strains. *Proceedings of the National Academy of Sciences* **2008**, 105, 16332-16337.
201. Veiga, S.; Henriques, S.; Santos, N. C.; Castanho, M. Putative role of membranes in the HIV fusion inhibitor enfuvirtide mode of action at the molecular level. *Biochemical Journal* **2004**, 377, 107.
202. Chong, H.; Yao, X.; Qiu, Z.; Qin, B.; Han, R.; Waltersperger, S.; Wang, M.; Cui, S.; He, Y. Discovery of critical residues for viral entry and inhibition through structural insight of HIV-1 fusion inhibitor CP621-652. *Journal of Biological Chemistry* **2012**, 287, 20281-20289.
203. Rimsky, L. T.; Shugars, D. C.; Matthews, T. J. Determinants of human immunodeficiency virus type 1 resistance to gp41-derived inhibitory peptides. *Journal of Virology* **1998**, 72, 986-993.
204. Sia, S. K.; Carr, P. A.; Cochran, A. G.; Malashkevich, V. N.; Kim, P. S. Short constrained peptides that inhibit HIV-1 entry. *Proceedings of the National Academy of Sciences* **2002**, 99, 14664-14669.
205. Naito, T.; Izumi, K.; Kodama, E.; Sakagami, Y.; Kajiwara, K.; Nishikawa, H.; Watanabe, K.; Sarafianos, S. G.; Oishi, S.; Fujii, N. SC29EK, a peptide fusion inhibitor with enhanced α -helicity,

- inhibits replication of human immunodeficiency virus type 1 mutants resistant to enfuvirtide. *Antimicrobial Agents and Chemotherapy* **2009**, 53, 1013-1018.
206. Nishikawa, H.; Nakamura, S.; Kodama, E.; Ito, S.; Kajiwara, K.; Izumi, K.; Sakagami, Y.; Oishi, S.; Ohkubo, T.; Kobayashi, Y. Electrostatically constrained α -helical peptide inhibits replication of HIV-1 resistant to enfuvirtide. *The International Journal of Biochemistry & Cell Biology* **2009**, 41, 891-899.
207. Gochin, M.; Cai, L. The role of amphiphilicity and negative charge in glycoprotein 41 interactions in the hydrophobic pocket. *Journal of Medicinal Chemistry* **2009**, 52, 4338-4344.
208. Chan, D. C.; Chutkowski, C. T.; Kim, P. S. Evidence that a prominent cavity in the coiled coil of HIV type 1 gp41 is an attractive drug target. *Proceedings of the National Academy of Sciences* **1998**, 95, 15613-15617.
209. Mo, H.; Konstantinidis, A. K.; Stewart, K. D.; Dekhtyar, T.; Ng, T.; Swift, K.; Matayoshi, E. D.; Kati, W.; Kohlbrenner, W.; Molla, A. Conserved residues in the coiled-coil pocket of human immunodeficiency virus type 1 gp41 are essential for viral replication and interhelical interaction. *Virology* **2004**, 329, 319-327.
210. Wang, S.; York, J.; Shu, W.; Stoller, M. O.; Nunberg, J. H.; Lu, M. Interhelical interactions in the gp41 core: implications for activation of HIV-1 membrane fusion. *Biochemistry* **2002**, 41, 7283-7292.
211. He, Y.; Liu, S.; Jing, W.; Lu, H.; Cai, D.; Chin, D. J.; Debnath, A. K.; Kirchhoff, F.; Jiang, S. Conserved residue Lys574 in the cavity of HIV-1 Gp41 coiled-coil domain is critical for six-helix bundle stability and virus entry. *Journal of Biological Chemistry* **2007**, 282, 25631-25639.
212. Jiang, S.; Debnath, A. K. A salt bridge between an N-terminal coiled coil of gp41 and an antiviral agent targeted to the gp41 core is important for anti-HIV-1 activity. *Biochemical and Biophysical Research Communications* **2000**, 270, 153-157.
213. He, Y.; Cheng, J.; Li, J.; Qi, Z.; Lu, H.; Dong, M.; Jiang, S.; Dai, Q. Identification of a critical motif for the human immunodeficiency virus type 1 (HIV-1) gp41 core structure: implications for designing novel anti-HIV fusion inhibitors. *Journal of Virology* **2008**, 82, 6349-6358.
214. Caffrey, M. Model for the structure of the HIV gp41 ectodomain: insight into the intermolecular interactions of the gp41 loop. *Biochimica et Biophysica Acta -Molecular Basis of Disease* **2001**, 1536, 116-122.
215. Bellows, M.; Taylor, M.; Cole, P.; Shen, L.; Siliciano, R.; Fung, H.; Floudas, C. Discovery of entry inhibitors for HIV-1 via a new de novo protein design framework. *Biophysical Journal* **2010**, 99, 3445-3453.
216. Gopalan, R. D.; Del Borgo, M. P.; Bergman, Y. E.; Unabia, S.; Mulder, R. J.; Wilce, M. C.; Wilce, J. A.; Aguilar, M.-I.; Perlmutter, P. Conformational stability studies of a stapled hexa- β -peptide library. *Organic & Biomolecular Chemistry* **2012**, 10, 1802-1806.
217. Huhmann, S. Master thesis. 2013.
218. Erdbrink, H.; Peuser, I.; Gerling, U. I.; Lentz, D.; Koksche, B.; Czekelius, C. Conjugate hydrotrifluoromethylation of α , β -unsaturated acyl-oxazolidinones: synthesis of chiral fluorinated amino acids. *Organic & Biomolecular Chemistry* **2012**, 10, 8583-8586.
219. Chakrabarty, A.; Kortemme, T.; Baldwin, R. L. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Science* **1994**, 3, 843-852.
220. Andersen, N. H.; Tong, H. Empirical parameterization of a model for predicting peptide helix/coil equilibrium populations. *Protein Science* **1997**, 6, 1920-1936.
221. Gerling, U. I. M. In *Personal Communication* **2013**.
222. Merrifield, R. B. Solid phase peptide synthesis. I. The synthesis of a tetrapeptide. *Journal of the American Chemical Society* **1963**, 85, 2149-2154.
223. Gude, M.; Ryf, J.; White, P. D. An accurate method for the quantitation of Fmoc-derivatized solid phase supports. *Letters in Peptide Science* **2002**, 9, 203-206.
224. Kaiser, E.; Colescott, R.; Bossinger, C.; Cook, P. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Analytical Biochemistry* **1970**, 34, 595-598.
225. Beck, W.; Jung, G. Convenient reduction of S-oxides in synthetic peptides, lipopeptides and peptide libraries. *Letters in Peptide Science* **1994**, 1, 31-37.
226. Pace, C. N.; Vajdos, F.; Fee, L.; Grimsley, G.; Gray, T. How to measure and predict the molar absorption coefficient of a protein. *Protein Science* **1995**, 4, 2411-2423.

-
227. Ahn, J.-M.; Boyle, N. A.; MacDonald, M. T.; Janda, K. D. Peptidomimetics and peptide backbone modifications. *Mini Reviews in Medicinal Chemistry* **2002**, *2*, 463-473.
228. Greenfield, N. J. Using circular dichroism spectra to estimate protein secondary structure. *Nature Protocols* **2007**, *1*, 2876-2890.
229. Fasman, G. *Circular dichroism and the conformational analysis of biomolecules*. Plenum Press: New York: 1996.
230. Greenfield, N. J. Methods to estimate the conformation of proteins and polypeptides from circular dichroism data. *Analytical Biochemistry* **1996**, *235*, 1-10.
231. Chen, Y.-H.; Yang, J. T.; Chau, K. H. Determination of the helix and β form of proteins in aqueous solution by circular dichroism. *Biochemistry* **1974**, *13*, 3350-3359.
232. Gans, P. J.; Lyu, P. C.; Manning, M. C.; Woody, R. W.; Kallenbach, N. R. The helix-coil transition in heterogeneous peptides with specific side-chain interactions: Theory and comparison with CD spectral data. *Biopolymers* **1991**, *31*, 1605-1614.
233. Chakrabarty, A.; Doig, A. J.; Baldwin, R. L. Helix capping propensities in peptides parallel those in proteins. *Proceedings of the National Academy of Sciences* **1993**, *90*, 11332-11336.
234. Chakrabarty, A.; Kortemme, T.; Baldwin, R. L. Helix propensities of the amino acids measured in alanine-based peptides without helix-stabilizing side-chain interactions. *Protein Science* **1994**, *3*, 843-852.
235. Allen, D. L.; Pielak, G. J. Baseline length and automated fitting of denaturation data. *Protein Science* **1998**, *7*, 1262-1263.
236. Förster, T. 10th Spiers Memorial Lecture. Transfer mechanisms of electronic excitation. *Discussions of the Faraday Society* **1959**, *27*, 7-17.
237. Sapsford, K. E.; Berti, L.; Medintz, I. L. Materials for Fluorescence Resonance Energy Transfer Analysis: Beyond Traditional Donor-Acceptor Combinations. *Angewandte Chemie International Edition* **2006**, *45*, 4562-4589.
238. Foltá-Stogniew, E. *Methods in Molecular Biology: New and emerging proteomic techniques*. Springer: 2006; Vol. 328.
239. Liu, S.; Zhao, Q.; Jiang, S. Determination of the HIV-1 gp41 fusogenic core conformation modeled by synthetic peptides: applicable for identification of HIV-1 fusion inhibitors. *Peptides* **2003**, *24*, 1303-1313.