Development of stabilized FK506 binding protein 51 variants & generation of ligand-based affinity chromatography beads

Entwicklung von stabilisierten FK506 bindendes Protein 51 Varianten & Generierung von Liganden-basierten Affinitätschromatographie Beads

Zur Erlangung des Grades eines Doktors der Naturwissenschaften (Dr. rer. nat.)
Genehmigte Dissertation von Anna Charalampidou aus Drama, Griechenland
Tag der Einreichung: 28.09.2023, Tag der Prüfung: 20.11.2023

1. Gutachten: Prof. Dr. Felix Hausch
2. Gutachten: Prof. Dr. Viktor Stein
Darmstadt, Technische Universität Darmstadt

TECHNISCHE UNIVERSITÄT DARMSTADT

Fachbereich Chemie
Clemens-Schöpf Institut
Arbeitsgruppe Hausch

Development of stabilized FK506 binding protein 51 variants & generation of ligand-based affinity chromatography beads
Entwicklung von stabilisierten FK506 bindendes Protein 51 Varianten & Generierung von Liganden-basierten Affinitätschromatographie Beads

Genehmigte Dissertation von Anna Charalampidou

Tag der Einreichung: 28.09.2023
Tag der Prüfung: 20.11.2023

Darmstadt, Technische Universität Darmstadt

Bitte zitieren Sie dieses Dokument als:
URN: urn:nbn:de:tuda-tuprints-263516
URL: http://tuprints.ulb.tu-darmstadt.de/26351
Jahr der Veröffentlichung auf TUprints: 2023

Dieses Dokument wird bereitgestellt von tuprints,
E-Publishing-Service der TU Darmstadt
http://tuprints.ulb.tu-darmstadt.de
tuprints@ulb.tu-darmstadt.de

Η παρούσα διδακτορική διατριβή είναι αφιερωμένη στη μνήμη του πατέρα μου Δαμιανού Χαραλαμπίδη.

Erklärungen laut Promotionsordnung

§ 8 Abs. 1 lit. c PromO

Ich versichere hiermit, dass die elektronische Version meiner Dissertation mit der schriftlichen Version übereinstimmt.

§ 8 Abs. 1 lit. d PromO

Ich versichere hiermit, dass zu einem vorherigen Zeitpunkt noch keine Promotion versucht wurde. In diesem Fall sind nähere Angaben über Zeitpunkt, Hochschule, Dissertationsthema und Ergebnis dieses Versuchs mitzuteilen.

§ 9 Abs. 1 PromO

Ich versichere hiermit, dass die vorliegende Dissertation selbstständig und nur unter Verwendung der angegebenen Quellen verfasst wurde.

§ 9 Abs. 2 PromO

Die Arbeit hat bisher noch nicht zu Prüfungszwecken gedient.

Darmstadt, 28.09.2023

A. Charalampidou

Danksagung

Zunächst möchte ich mich bei meiner gesamten Familie und Freunden bedanken für ihre Unterstützung während meiner gesamten Ausbildung und meiner Doktorarbeit. Выражаю Вам огромную благодарность за неустанную и продуктивную поддержку. Ein besondere Dank gilt dabei meinen Eltern Marina Vamvakidou und Andreas Kaiser, sowie meiner γιαγιά Anna Vamvakidou und meinem Partner Fabian Knaup.

Des Weiteren danke ich meinem Supervisor, Prof. Dr. Felix Hausch für die Aufnahme in seine Arbeitsgruppe und die Vergabe der abwechslungsreichen und interessanten Themen. Durch die Herausforderungen habe ich mich nicht nur fachlich weiterentwickelt, sondern bin auch persönlich gewachsen. Besonders schätze ich die Unterstützung bezüglich des MIT Forschungsaufenthalts.

Ein besonderer Dank gilt auch Dr. Christian Meyners, der mich während meines Forschungspraktikums und später bei der Masterarbeit betreut hat. Danke, dass Du dir immer Zeit genommen hast, mir die verschiedensten Methodiken zu erklären und auch während meiner Doktorarbeit, immer Zeit gefunden hast, über meine Problemstellungen zu diskutieren und mir zu helfen.

Außerdem bedanke ich mich bei meinen Kooperationspartnern und dem gesamten TRABITA-Konsortium. Die regelmäßigen Treffen waren immer sehr lehrreich, aber der Spaß kam auch nie zu kurz. Insbesondere bei Thomas Nehls möchte ich mich für jegliche MS-Messungen und für die gelungene Zusammenarbeit bedanken. Mein Dank gilt auch Dr. Tim Heymann für seine unermüdliche Unterstützung und die MS-Messungen.

Moreover, I have to thank Dr. Sebastian Pomplun and Prof. Bradley L. Pentelute for inviting me to work on my project in the Pentelute Lab. Dr. Satish Gandhesiri, Lia Lozano Salazar, Charlotte Farquhar, Alex Callahan, Dr. Joseph Brown, and Brenno Masina made me feel welcome in the new environment and I am grateful for their support inside and outside the Lab.

Weiterhin bedanke ich mich bei meinen Studenten Tibor Herz, Marco Pieroni, Simon Reiners und Paula Henke für ihren Beitrag zu dieser Arbeit.

Abschließend möchte ich meinen Dank an den gesamten Arbeitskreis richten. Die Arbeitsatmosphäre war sehr angenehm, und selbst wenn Experimente schief gelaufen sind, konnte man sich sicher sein, dass man trotzdem noch etwas zu Lachen hat. Ein besonderer Dank geht an die Organiker - Dr. Fabian Knaup, Dr. Michael Walz, Vanessa Buffa und Robin Deutscher -, die mich bei meinen Synthesen unterstützt und mit mir ihre Erfahrungen und Abzüge geteilt haben. Darüber hinaus möchte ich mich insbesondere bei meinen Biochemiker-Kollegen - Monika Gnatzy, Martha Taubert, Thomas Geiger, Sarah Engel, Dr. Asat Baischew, Maximilian Repity, Sabine Schäfer, Wisely Oki Sugiarto, Steffen Hartmann, Dr. Christian Meyners und Dr. Stephanie Merz - für die Ermutigungen und die gemeinsamen Kaffeepäuschen mit Tee und zahlreiche wissenschaftlichen... aber auch nicht wissenschaftlichen Diskussionen bedanken.

Zusammenfassung

Die vorliegende Arbeit umfasst zwei Projekte: die Stabilisierung der F67-out Konformation von FKBP51 (FK506 bindendes Protein 51) und die Entwicklung einer Liganden-basierten Affinitäts-Chromatographie-Matrix.

FKBP51 ist ein bekanntes Beispiel für eine transiente Bindungstasche, die für die Selektivität bei der Ligandenentwicklung und die Unterscheidung von strukturell ähnlichen homologen Proteinen entscheidend ist. In FKBP51 ist die transiente Bindungstasche durch das Ausklappen des Phenyl-Rests F67 aus der Bindungstasche gekennzeichnet und wird auch als F67-out-Konformation bezeichnet. Ziel war es, Proteinkonstrukte mit einer stabilisierten F67-out-Konformation zu erzeugen, um Hits in Screenings zu identifizieren, die diese spezifische Proteinkonformation binden. Durch rationales Protein-Engineering ist es gelungen, FKBP51-Varianten mit einer stabilisierten F67-out-Konformation zu erzeugen. Zu den Techniken, die zur kovalenten Fixierung der Konformation in der F67-out Konformation eingesetzt wurden, gehörten Crosslinking-Ansätze wie Photocrosslinking, Click-Chemie oder Cystein-basiertes Crosslinking sowie die chemische Proteinsynthese. Die Varianten zeigten eine verbesserte Bindungsaffinität zu einem konformationsselektiven Liganden-Tracer im Fluoreszenzpolarisations-Assay (FP-Assay), wobei die $FKBP51^{FK1}$ F67C/K60C Variante mit einer Disulfidbrücke sich als die am besten geeignete Variante für ein anschließendes Fragment-Screening erwies. Die Gründe dafür sind die einfache Herstellung der Proteinvariante in hoher Menge und Reinheit, die nachgewiesene Verbesserung der Bindungsaffinität für konformationsselektive Liganden und die Möglichkeit, das Protein in seiner Apo-Form zu kristallisieren. Ein erstes fragmentbasiertes Screening mittels Thermoshift-Assay, gefolgt von einem FP-Assay, ergab vielversprechende Hits mit einer aminsubstituierten Chinolin oder Isochinolin-Kernstruktur. Dies ist der erste Schritt auf einem neuen Weg zur Identifizierung von neuen Leitstrukturen.

Darüber hinaus wurde eine Liganden-basierte Affinitäts-Chromatographie-Matrix hergestellt. Dies soll ermöglichen FKBP51 aus komplexen biologischen Mischungen anzureichern sowie neue Interaktionspartner und/oder Off-Targets zu identifizieren. Dabei gelang es vier Liganden zu immobilisieren und zu testen, wobei Beladungs- und Elutionsbedingungen optimiert wurden. Zwei Systeme erwiesen sich als geeignet für die Anreicherung von FKBP51 aus komplexen biologischen Mischungen. Entweder wird SAFit1 immobilisiert und die Elution findet unter denaturierenden Bedingungen statt, oder es wird ein SAFit1-Analogon mit geringer Bindungsaffinität (hier SAFit-LA genannt) immobilisiert, so dass die Elution kompetitiv mit SAFit1 erfolgen kann.

Abstract

The presented work covers two projects: the stabilization of the F67-out conformation for the FK506 binding protein 51 (FKBP51) and the generation of a ligand-based affinity matrix.

FKBP51 is a prominent example of a transient binding pocket that is key to achieve selectivity in ligand development. In FKBP51, the transient binding pocket is characterized by a rearrangement of residue F67 outward from the binding pocket, also called the F67-out conformation. The aim was to generate protein constructs with a stabilized F67-out conformation to identify hits in screenings targeting this specific protein conformation. By rational protein engineering, I succeeded to generate FKBP51 variants with a stabilized F67-out conformation. Techniques used to covalently fix the conformation in this out conformation include crosslinking approaches such as photocrosslinking, click chemistry, or cysteine-directed crosslinking, as well as chemical protein synthesis. The obtained variants showed improvement in binding affinity to a conformationally selective ligand tracer in the fluorescence polarization (FP) assay. Overall, the FKBP51^{FK1} F67C/K60C variant with a disulfide bridge emerged as the most suitable variant for fragment screening. The reasons for this are the facile production of the protein variant in high quantity and purity, the improvement in binding affinity for conformationally selective ligands, and the possibility of crystallizing the protein in its apo form in a F67-out-like conformation. Initial fragment-based screening by thermoshift assay followed by FP assay resulted in promising hits with an amine-substituted quinoline and isoquinoline-core structure. This is the first step on a new path to identifying new lead structures.

In addition, a ligand-based affinity chromatography matrix was generated. This should allow FKBP51 to be enriched from complex biological mixtures as well as identify new interaction partners and/or off-targets. In this work, four ligands were successfully immobilized and tested, with loading and elution conditions optimized. Two systems were found to be suitable for the enrichment of FKBP51 from complex biological mixtures. Either SAFit1 is directly immobilized and elution takes place under harsh conditions, or a SAFit1 analog with low affinity (here called SAFit-LA) is immobilized so that elution can be performed competitively with SAFit1.

Inhaltsverzeichnis

1 Introduction **1**
- 1.1 Protein binding pockets . . . 1
  - 1.1.1 Binding mechanism . . . 2
  - 1.1.2 Dynamics of protein binding pockets . . . 5
  - 1.1.3 Transient and cryptic binding pockets in drug discovery . . . 9
- 1.2 Chemical protein modification . . . 12
  - 1.2.1 Protein modification via canonical amino acids . . . 12
  - 1.2.2 Protein modification via unnatural amino acids . . . 15
- 1.3 Ligand-based affinity chromatography . . . 19
  - 1.3.1 Generation of an affinity matrix . . . 19
  - 1.3.2 Principles of affinity chromatography . . . 20

2 Aim **22**

3 Results & Discussion **24**
- 3.1 Stabilization of the FKBP51^{FK1} F67-out conformation . . . 24
  - 3.1.1 Stabilization by point mutations . . . 24
  - 3.1.2 Stabilization by photocrosslinking . . . 30
  - 3.1.3 Stabilization by click chemistry . . . 34
  - 3.1.4 Stabilization by lactam bridge formation . . . 41
  - 3.1.5 Stabilization by cysteine-directed crosslinking . . . 49
  - 3.1.6 Fragment library screening with a stabilized FKBP51^{FK1} F67-out variant . . . 62
- 3.2 Ligand based affinity chromatography . . . 68
  - 3.2.1 Generation and examination of ligand-based affinity matrices . . . 69
  - 3.2.2 Application of ligand-based affinity chromatography . . . 79

4 Conclusion & Outlook **87**

5 Experimental section **90**
- 5.1 Site directed mutagenesis . . . 90
  - 5.1.1 SPRINP-PCR . . . 90
  - 5.1.2 Transformation . . . 91
  - 5.1.3 Plasmid Preparation . . . 92
- 5.2 Protein expression . . . 92
  - 5.2.1 Generation of chemically competent cells . . . 92
  - 5.2.2 Subsequential transformation for amber codon suppression . . . 92
  - 5.2.3 Protein expression . . . 92
  - 5.2.4 Lysis of bacterial cells . . . 93
  - 5.2.5 Purification of proteins . . . 93

5.3 Biophysical methods . . . 94
5.3.1 Fluorescence polarization assay . . . 94
5.3.2 Thermal Shift Assay . . . 96
5.4 Protein characterization . . . 97
5.4.1 SDS-PAGE . . . 97
5.4.2 Western Blotting . . . 98
5.4.3 LC-MS . . . 98
5.4.4 Analytical HPLC . . . 99
5.4.5 ETD measurements . . . 99
5.4.6 SEC-MS – native & denatured . . . 99
5.4.7 IM-MS and CCS . . . 100
5.4.8 CIU . . . 100
5.4.9 Protein Crystallization . . . 100
5.5 Protein crosslinking . . . 102
5.5.1 Photocrosslinking . . . 102
5.5.2 CuAAC . . . 102
5.5.3 PEGylation . . . 102
5.5.4 Crosslinking with DBCO-MI . . . 102
5.5.5 Cysteine directed crosslinking with BMB . . . 102
5.5.6 Cysteine directed crosslinking with DBA . . . 103
5.5.7 Disulfid bridge formation . . . 103
5.6 Protein Synthesis . . . 104
5.6.1 Automated peptide flow synthesis (AFPS) . . . 104
5.6.2 Biotinylation . . . 104
5.6.3 Orthogonal on-bead deprotection of Alloc and OAll groups . . . 105
5.6.4 Lactamization . . . 106
5.6.5 Cleavage . . . 106
5.6.6 Preparative mass directed RP-HPLC . . . 106
5.6.7 Refolding of the synthesized FKBP51 variants . . . 106
5.7 Ligand-based affinity chromatography . . . 107
5.7.1 Synthesis . . . 107
5.7.2 Generation of cell lysates . . . 112
5.7.3 Immobilization of a ligand on a solid support . . . 113
5.7.4 Protocol for affinity chromatography . . . 114

6 Abbreviations **116**

7 Appendix **119**
7.1 Summary of chemical structures **1**-**13** . . . 119
7.2 Thermal shift assay data . . . 120

8 References **125**

1 Introduction

The human genome project was an international collaboration aimed to create a draft sequence of the human genome.[1] Surprisingly, about 30,000 protein-coding genes were identified, instead of the estimated 100,000 genes.[1–3] This highlights the complexity of the biological machinery and the diverse biological functions of protein variations achieved through alternative splicing and post-translational modifications (PTMs) rather than large numbers of different genes. To describe the different ”protein forms”, the term proteoform was first introduced by Smith and Kelleher in 2013 to highlight the different molecular protein forms as the product of a single gene including variations due to alternative splicing and PTM's.[2] New tools are needed for a deeper understanding of the role of proteoforms in multifunctional complexes, as well as for the identification of their localization, abundance, and characteristics. Knowledge of proteoforms is a key to understand biological mechanisms in health, as well as in disease, when it comes to drug discovery.[3,4] The first step in drug discovery is the identification of a possible drug target which is druggable.[5] In 2002 Hopkins and Groom introduced the concept of the ”druggable genome”, which refers to the set of human genes that code for proteins that can bind drug-like molecules.[6] A few years later, Russ and Lampel estimated the number of protein-coding genes with druggable pocket at about 3000.[7] Nowadays, a druggable protein target is defined as a protein that has a binding site to which a small, selective, orally available, non-toxic molecule can bind that serves as a molecular modulator for altering dysfunctional biological mechanisms caused by disease.[8] One of the first models for the binding of proteins to ligands was described by Fischer, who postulated that the protein binding pocket acts like a rigid lock into which only a precisely fitting substrate can fit like a key.[9] Later, it became clear that proteins in solution exhibit flexible and dynamic behavior, which will be discussed in more detail in section 1.1.1.[10]

1.1 Protein binding pockets

The investigation of protein-protein or protein-ligand interactions is crucial for a deeper understanding of cellular processes in living organisms.[11] Proteins exhibit a wide range of functions, from structural building blocks and mechanical sensors to biochemical signaling. In this regard, proteins interact with various (macro)molecules such as other proteins, peptides, nucleic acids, membranes, and small molecules through molecular recognition via so called protein binding pockets.[12,13] The protein binding sites are characterized by cavities on the surface of proteins, that can accommodate other proteins or small molecules, resulting in a binding interaction through three main types of interactions: Van-der-Waals, hydrogen bond and electrostatic.[14] The complex formations are highly dynamic and can be distinguished into reactive or programmed changes. Reactive interactions are reactions initiated by exogenous stimuli, whereas programmed changes are caused by endogenous signals, e.g. in cell cycle dynamics.[12]

1.1.1 Binding mechanism

The binding of a protein and a ligand can be described by a simple equation formulated as follows:

$$P + L \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}}}{\rightleftharpoons}} PL \tag{1.1}$$

where P is defined as the protein, L as the ligand, and PL as the associated protein-ligand complex. The kinetic rate constant of association is described by k_{on} and that of dissociation by k_{off}. When the association of protein and ligand is in balance with the dissociation of the complex, the system is in equilibrium. Thereby, the binding constant K_b is defined as the ratio of k_{on} and k_{off}, which in turn can be described as the ratio of the equilibrium concentration of the complex divided by the product of the equilibrium concentrations of free protein and ligand. The reciprocal of the binding constant (K_b) is the dissociation constant K_d with a unit of M.[13]

$$K_\mathrm{b} = \frac{k_\mathrm{on}}{k_\mathrm{off}} = \frac{[PL]}{[P][L]} = \frac{1}{K_\mathrm{d}} \tag{1.2}$$

Many experimental methods are available to determine the binding affinity of a protein-ligand complex (K_d) by e.g. fluorescence polarization (FP)[15,16], nuclear magnetic resonance spectroscopy (NMR)[17], isothermal calorimetric titration (ITC)[18] and surface plasmon resonance spectroscopy (SPR)[19]. However, for the rational design of drugs , it is not only crucial to determine the binding affinity of protein and ligand, but also to understand the underlying mechanism of the recognition of the interacting partners.[20] As mentioned in the beginning, the first binding mechanism model was introduced by Fischer in 1894. It states that a substrate only binds to its enzyme and triggers a chemical reaction if the reactants fit together like a key to a lock.[9] Thus, in this simplest case, the formation of the protein-ligand complex occurs in one step, described as follows:[20]

$$P \underset{\mathrm{k_{off}}}{\overset{\mathrm{k_{on}[L]}}{\rightleftharpoons}} PL \tag{1.3}$$

The change in the concentration of the protein or protein-ligand complex over time can be formulated in the differential equations 1.4 and 1.5:[20]

$$d[P]/dt = -k_\mathrm{on}[L] + k_\mathrm{off} \tag{1.4}$$

$$d[PL]/dt = k_\mathrm{on}[L] - k_\mathrm{off} \tag{1.5}$$

or:

$$\begin{pmatrix} d[P]/dt \\ d[PL]/dt \end{pmatrix} = \begin{pmatrix} -k_\mathrm{on}[L] & k_\mathrm{off} \\ k_\mathrm{on}[L] & -k_\mathrm{off} \end{pmatrix} \begin{pmatrix} [P] \\ [PL] \end{pmatrix}$$

Determining the non-zero eigenvalues of the 2x2 matrix leads to equation 1.6, which shows the linear relationship between the observed rate constant k_{obs} and the equilibrium ligand concentration [L]. The slope is k_{on} (association rate constant) and k_{off} (dissociation

rate constant) is the intercept, from which the dissociation constant K_d can be easily calculated.[20]

$$k_{obs} = k_{off} + k_{on}[L] \tag{1.6}$$

Unfortunately, the one-step binding mechanism is not a realistic model. Koshland pointed out that a non-substrate can bind an enzyme but does not necessarily induce a chemical reaction, which the lock-and-key model could not explain. These limitations of the model led to the postulated mechanism of induced fit, which takes into account protein plasticity.[21,22] The induced fit mechanism is characterized by a two-step process. First, the protein binds weakly to the ligand, inducing a conformational change in the protein binding pocket, which subsequently leads to tight binding. Another model introduced in the late 60s is the so called conformational selection binding mechanism. In case of a conformational selection mechanism the conformational rearrangement of the 3D protein structure first occurs, before the ligand binds specifically to the protein.[23,24] The induced fit mechanism can be represented as in equation 1.7, where k_r is the rate constant of the conformational rearrangement of the protein and the reverse reaction is described by the rate constant k_{-r}:

$$P + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} (PL) \underset{k_{-r}}{\overset{k_r}{\rightleftharpoons}} PL \tag{1.7}$$

For the conformational selection mechanism on the other hand equation 1.8 applies:

$$P \underset{k_{-r}}{\overset{k_r}{\rightleftharpoons}} (P) + L \underset{k_{off}}{\overset{k_{on}}{\rightleftharpoons}} PL \tag{1.8}$$

Differentiating between the induced fit and conformational selection binding mechanism is not a trivial task. One experimental approach is to determine the relaxation rate constant (k_{obs}, obs for observed) as a function of ligand concentration, e.g., using the stopped-flow method. Under the rapid equilibrium approximation, one can assume that a hyperbolic decrease in the k_{obs} value as a function of ligand concentration indicates that a conformational selection binding mechanism has occurred, whereas a hyperbolic increase in k_{obs} indicates an induced fit mechanism. The underlying equations are 1.9 for induced fit and 1.10 for conformational selection.[20,25,26]

$$k_{obs} = k_{-r} + k_r \frac{[L]}{K_d + [L]} \tag{1.9}$$

$$k_{obs} = k_r + k_{-r} \frac{K_d}{K_d + [L]} \tag{1.10}$$

However, the rapid equilibrium approximation assumes that the binding and dissociation steps are much faster compared to the conformational rearrangement processes, which is not necessarily true.[26]

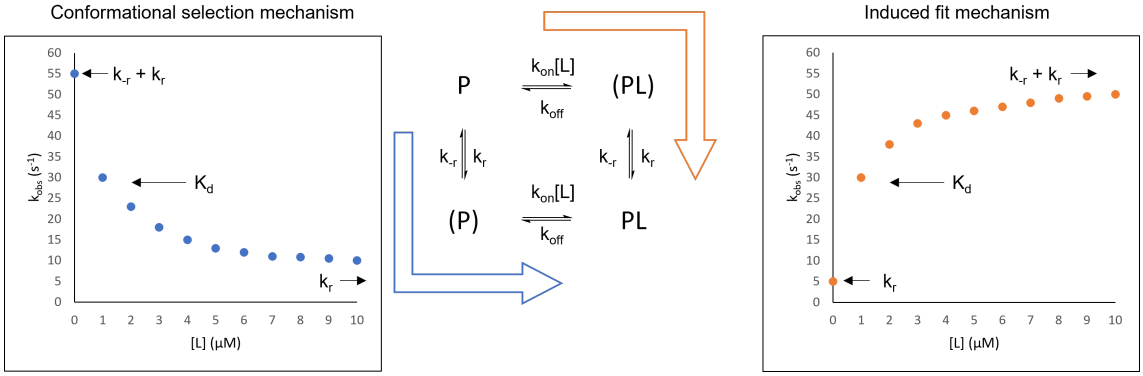


Figure 1.1: Formation of protein-ligand complexes by conformational selection (in blue) compared to induced fit mechanisms (in orange), where P stands for protein, L for ligand, and PL for the protein-ligand complex. Brackets indicate a conformational change in the protein. Also, k_{on} represents the association rate constant and k_{off} stands for the dissociation rate constant. [L] is the ligand concentration, k_r is the rate constant of the conformational rearrangement, and k_{-r} represents the reverse reaction. According to the rapid equilibrium approximation model, k_{obs} decreases with [L] when following the binding mechanism of conformational selection, as shown in the modeled hyperbolic curve (left, in blue) according to the equation 1.10. For the induced conformational model, the equation 1.9 applies, leading to an increase in k_{obs} with [L], as shown in the modeled curve on the right, in orange. Based on [20].

Di Cera and Vogt showed in their publications[20,25,26] that the simplified rapid equilibrium approximation led to misleading interpretation of various binding mechanism studies.[27] When no assumptions are made about relative binding rates and conformational rearrangements, the equation 1.11 applies to induced fit and the equation 1.12 applies to conformational selection mechanism.

$$-\lambda_{1,2} = \frac{k_{-r} + k_r + k_{off} + k_{on}[L] \pm \sqrt{(k_{off} + k_{on}[L] - k_{-r} - k_r)^2 + 4k_r k_{off}}}{2} \tag{1.11}$$

$$-\lambda_{1,2} = \frac{k_{-r} + k_r + k_{off} + k_{on}[L] \pm \sqrt{(k_{off} + k_{on}[L] - k_{-r} - k_r)^2 + 4k_{-r} k_{on}[L]}}{2} \tag{1.12}$$

The eigenvalues are the non-zero solutions of a set of differential equations describing the change in concentration of protein, ligand, and the resulting complex, taking into account all conformational transitions over time depending on the assumed binding mechanism. The equations can be expressed in matrix form, where the concentrations of the species are the elements of a column vector and the rate constants are the elements of the matrix. The large eigenvalue $-\lambda_2$ represents the fastest rate, while the smaller eigenvalue $-\lambda_1$ is usually directly related to k_{obs}, which can be determined by fast kinetic techniques such as stopped flow. The only difference between the 1.11 and 1.12 equations is the last term in the square root, which results in k_{obs} always increasing with [L] in an induced fit mechanism, whereas in conformational selection k_{obs} can increase, decrease, or be independent of [L] depending on the relative values of k_{off} and k_r. That is, if k_{obs}

decreases with [L] or is independent, the binding event under study can be assumed to proceed with a conformational selection mechanism, but an increase in k_{obs} with [L] does not automatically mean that the underlying binding mechanism is the induced fit mechanism.[20] An important piece of the puzzle when investigating the underlying binding mechanism of two interacting partners is to consider structural data such as X-ray or NMR studies. The presence of multiple ligand-free protein conformations may indicate that the underlying binding mechanism is conformational selection. Multiple ligand-protein complexes on the other hand may indicate that the underlying binding mechanism is induced fit.[25]

1.1.2 Dynamics of protein binding pockets

As already established, proteins are dynamic structures, and the ability of protein binding pockets to change their conformation is essential for binding specificity. Protein binding pockets of a temporal nature can also be referred to as transient binding pockets. The Wade group classified the different shapes of protein binding pockets into five groups: (1) subpockets, (2) adjacent pockets, (3) binding pockets with a breathing motion, (4) channel or tunnel pockets and (5) allosteric pockets (see figure 1.2).[28] The characteristics and some example proteins for each class are presented below. However, it is important to note that the characteristics of these groups may overlap.

In general, an adjacent pocket is characterized by the appearance or disappearance of an additional space in close proximity to the original pocket of the protein.[28] For example, in interleukin-2, rotation of the side chain and flexibility of the backbone near the original binding site results in an adjacent pocket that can accommodate a small molecular ligand.[29,30] In contrast, protein binding breathing is vibrational motion caused by the flexibility of the side chains and protein backbone.[28] An example of this is the protein B-cell lymphoma-extra-large (BCL-XL), in which the α-helices in the protein binding site exhibit a contractive motion.[31–33] Another class of dynamic protein binding pockets are the channel or tunnel pockets, which close or open dynamically.[28] Tassin-Moindrot et al. investigated the complex formation of nonspecific lipid transfer protein (ns-LTP) with prostaglandin B_2 by NMR. The NMR data revealed different binding modes depending on the protein and/or the ligand.[34] An allosteric transient pocket is a dynamic appearing or disappearing pocket that is distant from the original pocket but can change it through ligand binding.[28] The highly conserved DFG motif in the active site of the P38 mitogen-activated protein kinase (P38 MAPK) is able to open an allosteric binding pocket by flipping the F residue toward the ATP/ADP binding site. This conformational change allows for modulation of binding events.[35]

A transient subpocket is characterized by the appearance or disappearance of a subpocket in the original binding pocket and allows the spatial expansion of the binding site.[28] An example of a transient binding subpocket in a protein is the FK506 binding protein 51 (FKBP51). Since this is the model protein of this work, the protein is introduced here in more detail.

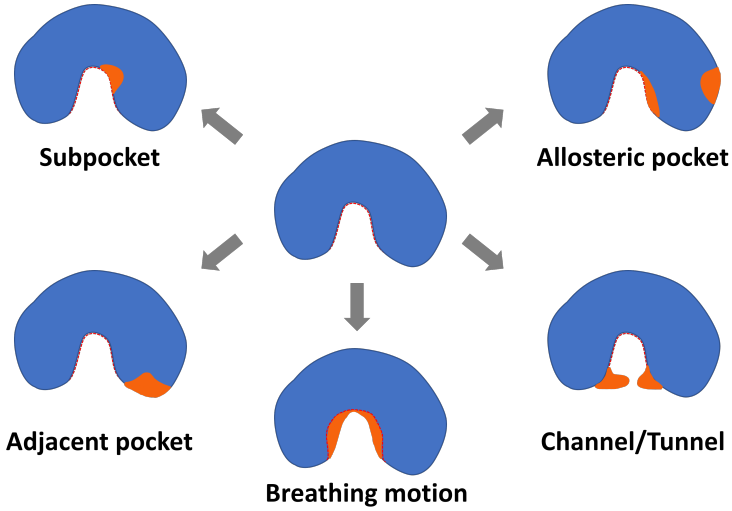


Figure 1.2: The five protein binding pocket dynamic classes defined by STANK ET AL. The reference structure is shown in the middle with the red dotted line representing the protein binding pocket. The orange regions show the difference in protein binding shape depending on the class of protein dynamics. Based on [28].

Structure and biological function of FKBP51

FK506-binding proteins (FKBPs) belong to the immunophilin family and bind the natural, immunosuppressive drugs rapamycin and FK506.[36] In humans, there are over 15 FKBP proteins with a highly conserved sequence compared to FKBP12, the first discovered and best studied FKBP member.[37,38] In this work, the protein of interest is FKBP51, which is encoded by the gene *fkbp5* and plays a major role in the regulation of steroid hormorn receptors and the modulation of the stress response.[39–41] Structurally, FKBP51 is composed of an FK1, an FK2 and a TPR (tetratricopeptiderepeat) domain.[42,43] The structurally best studied domain FK1 of FKBP51 is situated at the N-terminus. The FK fold consists of five antiparallel β-strands bent around a central α-helix.[43,44] At the N-terminus of the FK1 domain is an additional α helix (α1), followed by the three β strands β1, β2, and β3, with the β3 strand being separated by a bulge region forming a β3a and β3b strand. This is followed by the α2 helix and finally the β4 and β5 strands (see figure 1.3).[45]

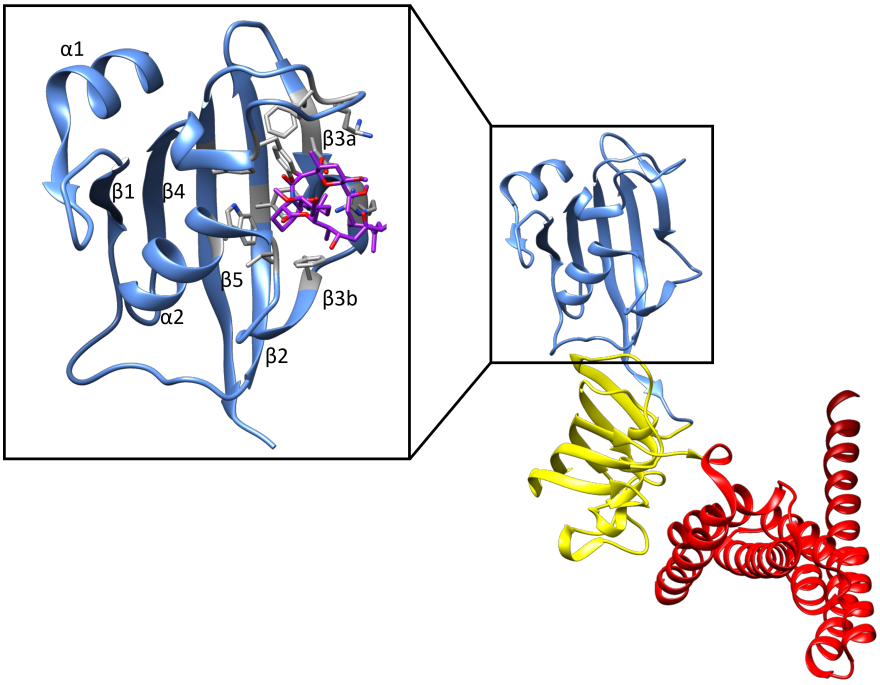


Figure 1.3: Structure of full-length FKBP51 (PDB: 5OMP. Blue: FK1 domain, yellow: FK2 domain, red: TPR domain) with zoom on the FK1 domain (PDB: 3O5E) in complex with the natural product FK506 (purple) and the side chains of the binding pocket in gray.

The FK2 domain exhibits 19 % sequence identity with the FK1 domain, but still shows typical FK folding. Exemplary differences between the two domains are the absence of the 40s-loop between the β3a and β3b strands, resulting in a continuous β3 strand and that only 6 of 12 residues forming the binding pocket are identical.[44] While the FK1 domain has a petidyl-prolyl-*cis-trans* isomerase activity, which is inhibited by binding to FK506 or Rapamycin, the FK2 domain is catalytically inactive and not able to bind aforementioned substances.[40,43] At the C-terminus is the tetratricopeptide repeat domain TPR located, which consists of seven antiparallel α-helices made up of repeats of the 34 amino acid long TPR motif and an additional seventh helix at the end. FKBP51 is best known for being a co-chaperone of heat shock protein 90 (Hsp90) in the steroid hormone receptor (SHR) maturation cycle, with the TPR domain serving as a binding site for Hsp90 by recognizes the MEEDV motif of Hsp90. The TPR domain is structurally similar to other Hsp90 binding proteins such as PP5 (serine/threonine protein phosphate 5), HOP (Hsp70-Hsp90-organizing protein), or Cyp40 (cyclophilin 40) and the closest homolog of FKBP51, FKBP52.[44,46] It has been postulated that binding of FKBP51 to the GR-Hsp90 complex reduces the responsiveness of GR to cortisol. After hormone binding, FKBP51 is replaced by its counterpart, FKBP52, which in turn binds the complex to the dynein machinery. The dynein-dependent transport machinery then supports the translocation of the receptor complex to the nucleus and thereby may be part of transcriptional regulation.[47,48]

FKBP51 as a drug target

Single nucleotide polymorphism in the *fkbp5* gene showed a correlation with recurrent depressive episodes.[49] The risk alleles appear to lead to increased expression of FKBP51, resulting in higher susceptibility to stress-related diseases due to lower GR responsiveness.[50] FKBP51 knockout mice under chronic social defeat stress (CSDS) had lower adrenal weight and corticosterone levels and showed progressively higher recovery rates and more active stress coping behaviors.[47] The most likely influence on stress coping behavior appears to be the role of FKBP51 in the hypothalamic-pituitary-adrenal (HPA) axis, where the hypothalamus releases CRH (corticotropin releasing hormone) in response to stress. This leads to the production of ACTH (adrenocorticotropin hormone) by the pituitary gland. The secretion of ACTH into the bloodstream results in the release of cortisol by the adrenal gland, which initiates a negative feedback loop by activating the GR receptor. Thereby the release of CRH and ACTH is reduced. FKBP51 appears to regulate the negative feedback loop upon binding by decreasing the affinity of the GR for corticosteroids.[43,48] However, when FKBP51 is overexpressed, the GR becomes resistant to cortisol, leading to a dysregulated HPA axis.[51,52] In addition, FKBP51 knockout mice exhibited a lean phenotype even after a high-fat diet. These mice also exhibited low plasma concentrations of triglycerides and free fatty acids, suggesting that FKBP51 acts directly on adipocytes. In addition, an increase in the expression of enzymes involved in muscle energy expenditure was observed in FKBP51 knockout mice.[43] Maiaru et al. additionally showed that FKBP51 knockout mice exhibit reduced hypersensitivity in various rodent persistent pain models. These studies imply that FKBP51 plays an important role in regulating chronic pain, primarily by modulating glucocorticoid signaling in rodents, which may also be relevant to humans. Spinal FKBP51 also plays a critical role in nociceptive processing, which was emphasized by the reduction of an existing pain state after FKBP51 silencing by a specific inhibitor in a mouse model.[53] Thus, FKBP51 is proposed as a drug target for the treatment of stress-related disorders[54,55], obesity[56,57] and chronic pain[58].

Transient binding pocket of FKBP51

Since FKBP51 has been declared a potential drug target protein, inhibitor development has become of great interest. However, the major challenge in FKBP51 drug development is selectivity towards its closest homolog FKBP52, as the binding pockets of both proteins are very similar (see figure 1.4A), but the biological functions are opposite. In a bump-and-hole study, artificial ligands were designed to bind only the FKBP51/52^{F67V} mutants because the Cα substituent of the ligand was expected to clash with the wild-type F67 residue (see figure 1.4B: the green F67-in side chain clashes with the ligand shown in purple). However, the developed ligand iFit1 surprisingly showed low binding affinity (61000 nM) to wild-type FKBP51 but not to FKBP52. Further developments of the iFit1 ligand resulted in iFit4 with a K_i of 26 nM and finally SAFit1 and SAFit2 with a K_i of 4 nM and 6 nM, respectively.[59] Crystal structure analysis of the FKBP51-iFit complex revealed that the key to selectivity lies in the formation of a subpocket in FKBP51. For FKBP52 this conformational rearrangement is energetically not favorable. The transient protein conformation is characterized by the F67 residue being flipped outward of the binding pocket to accommodate ligands from the iFit series including SAFit1 (see figure 1.4).[38,59,60]

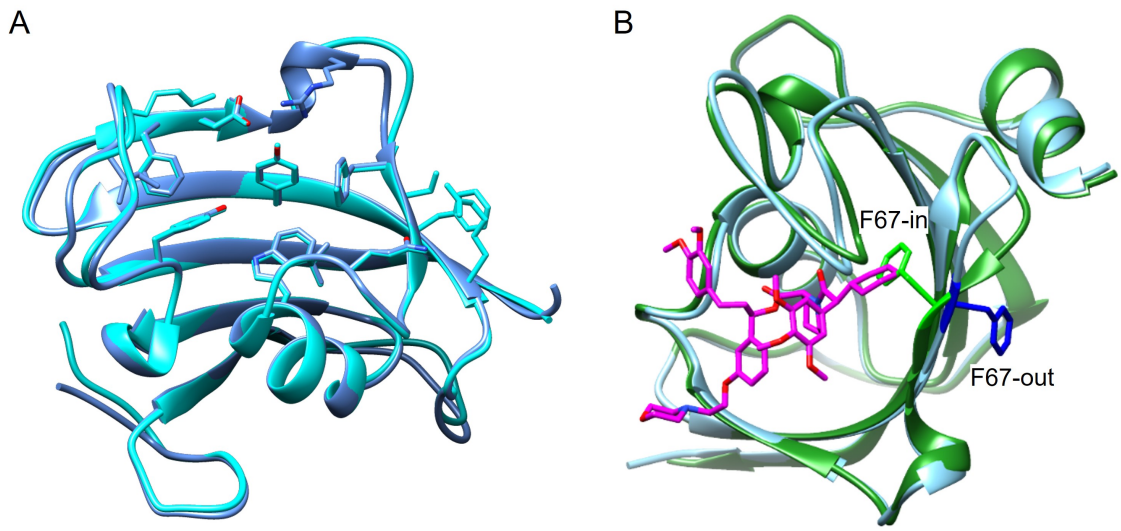


Figure 1.4: A: Overlay of the FK1 domain of FKBP51 (PDB: 4DRI, blue) and FKBP52 (PDB: 4DRJ, cyan). B: Overlay of FKBP51^{FK1} F67-in structure in blue (PDB: 4DRK) and FKBP51^{FK1} F67-out structure in green (PDB: 4TW7) in complex with iFit4 (purple strucutre, PDB: 4TW7) as ligand.

1.1.3 Transient and cryptic binding pockets in drug discovery

Drugs are compounds that are designed to have a therapeutic effect by binding to and modulating the biological (dys)functions of the aimed drug target.[61] In general, there are two main classes of drugs: biologics and small molecules. Biologics are derived from living organisms or biological processes, while small molecules are defined, chemically accessible structures.[62] While biologics have gained prominence in recent years and have made exciting advances in drug development, small molecules continue to have important advantages and play an important role in drug discovery. For example, small molecules have a low molecular weight and a well-defined structure that allows the molecule to be fully characterized. In addition, small molecules are usually stable, orally available and can be produced more (cost-)efficiently, making them more accessible to patients.[62,63] This chapter focuses on small molecule drug discovery and, in particular, how transient binding pockets and cryptic binding sites should be considered in drug discovery. Transient binding pockets are characterized by their ability of conformational changes with or without a ligand present, while cryptic binding sites are hidden binding sites that are only accessible in the presence of a small molecule or other inducers. Thus, targeting cryptic or transient binding sites can expand the number of target molecules that were previously considered undruggable.[64,65]

Druggability and computational methods to identify transient and cryptic binding pockets

The druggable genome introduced by Hopkins and Groom, as a set of about 3000 human genes encoding proteins that can bind drug-like molecules, features fewer than 700 proteins targeted by FDA-approved drugs to date.[6,7,66] To expand the druggable genome, cryptic and transient binding sites should be taken into consideration.[64,65] Various

structure-based methods can be employed to identify hidden and druggable cryptic binding pockets. Computational approaches can be broadly divided into two classes: geometry-based methods and energy-based methods. In geometry-based methods, the protein can be constructed in a three-dimensional grid, and probes are systematically moved along the grid lines to identify accessible or inaccessible regions. Energy-based methods, on the other hand, are based on calculating the interaction energy between the atoms of the protein and a small molecule.[67] To examine, if the cryptic binding pockets are druggable, Volkaner et al.[68] have developed DoGSiteScorer, for assessing computational druggablity of the to be investigated binding site. DoGSiteScorer considers global and local discriptors to generate data for druggability assessment.[69] To identify and track conformational changes in transient binding pockets, the TRAPP (TRansient Pockets in Proteins) web server can be used. It allows users to examine conformational changes in the binding pocket caused by the internal flexibility of the protein and to identify subpockets that may open transiently. The transiently opened binding pockets can be taken into account for ligand design and optimization, so the program also provides physicochemical properties, sequence properties, shape, and dynamic information.[70]

Ligand and fragment screening of transient and cryptic binding pockets

Drug discovery requires the identification of a hit structure through techniques such as combinatorial chemistry and high-throughput screening (HTS), followed by the development of a lead structure and its optimization.[61] In the pharmaceutical industry, high-throughput screening (HTS) is a widely used technique for lead identification, as this method tests large numbers of up to several million compounds against the target drug. However, newer techniques focus on screening drug-like compounds.[71] The term ”drug-like”refers to small molecule compounds in drug discovery that have certain physicochemical properties that increase the likelihood that they will successfully pass through the development processes. Among the important criteria that must be met is an appropriate pharmacokinetic (PK) profile that includes adequate administration, distribution, metabolism, and excretion (ADME) profile.[61,72] In addition, methods such as protein crystallography, electron cryomicroscopy and NMR spectroscopy enable the structure based and rational drug development and optimization.[73] An increasingly popular method that combines the advantages of HTS and structure-based screening is fragment-based drug discovery (FBDD). In FBDD, a library of fragments with a molecular weight of up to 300 Da is screened against the target. Fragment libraries are usually relatively small and therefore need to be structurally diverse and carefully adjusted to the target. Compared to HTS, FBDD requires a screening assay with high sensitivity because the fragments have low binding affinity due to their low molecular weight and thus have fewer heavy atoms that can interact with the protein binding site. However, more specific assays are needed for HTS to avoid false positive hits. Fragment screening holds the advantage that fragments are large enough to allow protein-fragment interactions, yet have low complexity and are small enough to reduce the likelihood of steric clashes and identify hot spots. They are ideal starting points for further development using fragment linking, merging and growing strategies.[74] Several companies demonstrated the utility of fragment-based screening alone or in combination with HTS, showing the potential of FBDD screening to yield more hits against a wider range of protein targets. Abbott, for example, studied 45 targets in FBDD and HTS screenings, both separately and combined. In FBDD screening alone, 25 % of these targets received good hits, compared to only 2 % in HTS screening

alone. 22 % of the targets received optimized lead structures derived from these hits from FBDD screening, of which 16 % of the targets received potent inhibitors (IC_{50} < 100 nM). Lead structure optimization of HTS hits was successful for 13 % of the targets resulting in potent inhibitors for 11 % of targets.[75] As mentioned earlier, consideration of cryptic and transient binding pockets can expand the druggable genome, but also in fragment-based drug design, consideration of these binding pockets can have the advantage of finding ligands with higher selectivity, as discussed by Bartolowits and Davisson.[76] For example, computational methods have revealed the difference in pocket dynamics of human and *Plasmodium falciparum* Hsp90, enabling the targeted development of pathogen-selective inhibitors. This example shows how consideration of even small changes in protein dynamics can be used in the design of selective inhibitors.[77] Another example is the discovery of a cryptic binding site near the redox active site of the disulfide bond-catalyzing protein BpsDsbA (*Burkholderia pseudomallei* disulfide bond-forming protein A), which is a promising target for the treatment of the pathogenic bacterium *Burkholderia pseudomallei*. However, due to its rather flat and hydrophobic binding site, engineering small molecules that inhibit enzyme activity is quite challenging, especially compared to DsbA from E.coli, for which several inhibitors are known. By fragment-based screening with about 1300 fragments, the authors identified two fragments (bromophenoxypropanamide (1) and 4-methoxy-*N*-phenylbenzenesulfonamide (2)) that bind to the open cryptic pocket characterized by the conformational rearrangement of residue Y110 upon fragment binding. NMR and X-ray experiments confirmed the fragment position and weak binding (>2 mM) as well as conformational rearrangement, making these fragments promising starting points for extension and development into potent BpsDsbA inhibitors.[78] Although FBDD has proven to be a powerful tool for drug discovery, there are limitations that need to be overcome. It remains difficult to predict the position of the fragment in the target protein, so structural information is required, but can be difficult to obtain by e.g. crystallography. In addition, more optimization cycles may be required, which increases the time needed to develop the lead structure. Another limitation is the usually low binding affinity between the fragments and the target protein, which can lead to potential hits being overlooked.[79] In this work, the latter hurdle should be overcome by designing and creating a screening tool with a stabilized transient binding pocket to find fragment hits for the protein of interest FKBP51 (for more detail see chapter 2). An example of how protein modification can lead to a promising starting point for lead structure optimization is the work of Nguyen et al.. They studied, tRNA-guanine transglycosylase to fight shegellosis. This protein is functional only in the homodimeric state, so disruption of the homodimer should lead to loss of function. Unfortunately, crystallization of only the monomer was not possible because stable packing into a dimer is preferred during protein crystallization. The reported quasimonomeric state of the protein, achieved by disulfide bridging, leads to a different orientation of the monomeric protein molecules. Here, the relevant loop-helix motif, which is normally part of the homodimeric interface, is exposed, making it possible to obtain structural information about the protein-protein binding interface in its monomeric form. The authors used this variant for fragment soaking experiments and identified 7 fragments that address a novel transient binding pocket within the dimer interface. They state that these results could serve as a starting point for optimizing more potent inhibitors.[80]

1.2 Chemical protein modification

As mentioned above, FBDD is a useful tool to find a starting point for a lead structure. However, potential fragment hits may be overlooked due to low binding affinity. To overcome this hurdle, the targeted cryptic or transient binding pockets can be stabilized, leading to more sensitive and targeted fragment screening. There are several ways to chemically modify a protein to stabilize the targeted cryptic or transient binding pocket. In general, protein modification and bioconjugation is a broad field in which chemical probes are introduced into proteins with a natural or genetically engineered amino acid sequence. Over time, a wide range of bioorthogonal reactions and corresponding probes have been developed, overcoming challenges such as the need to perform the reactions under mild conditions in aqueous solutions and in the presence of several unprotected functional groups. In addition, enzymatic bioconjugation methods have been developed but will not be discussed here. Another important point is that after modification, the structural and functional properties of the proteins under study are preserved. Applications using protein modification and bioconjugation methods include enhancing bioavailability, targeting drug delivery, tracking protein localization using fluorescent labeling, and labeling for a variety of assays, to name a few.[81] A special case of protein modification is crosslinking, in which molecular components such as nucleic acids, drugs, peptides and proteins are covalently linked by crosslinking agents. Three types of crosslinkers can be distinguished: homobifunctional, heterobifunctional, and zero-length crosslinkers. Zero-length crosslinkers activate the functionality of one chemical entity to react with another without introducing extrinsic atoms. In lactam formation, activation reagents such as carbodiimides are an example of a zero-length crosslinker.[82] Other examples include disulfide bond-forming reagents such as 2,2'-bispyridyl disulfide (2-PDS) or 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman reagent or DTNB).[83,84] In contrast, bifunctional crosslinking reagents self-assemble between crosslinked components. Homobifunctional crosslinkers have two identical functional groups that target the same reactive functional moiety, while heterobifunctional crosslinkers contain two different functional groups. Another distinction is that between inter- and intramolecular crosslinking. It is possible that both types occur during protein crosslinking. However, reaction conditions can influence the predominant type of reaction. A high protein concentration favors intermolecular crosslinking, while a low protein concentration can lead to predominantly intramolecular crosslinking. The distinction is important because inter- and intramolecular crosslinking have different applications. When studying protein-protein interactions or labeling proteins for medical applications, intermolecular crosslinking is the tool of choice. On the other hand, intramolecular crosslinking is a useful tool for studying distances within a protein or analyzing surface topology.[82] In the following, different crosslinking techniques with canonical and also unnatural amino acids are discussed.

1.2.1 Protein modification via canonical amino acids

The canonical amino acids Trp, Tyr, Met and His can be considered for protein modification. The advantage of these amino acids is their low-abundance in protein sequences, but the low reactivity is a major drawback. Nevertheless, methods were developed which target these amino acids. For example, metal-mediated reactions with TIPS-EBX (1-[(triisopropylsilyl)ethynyl]-1,2-benziodoxol-3(1H)-one) catalzed by [AuCl(SMe_2)] target Trp with a relatevily high conversion rate.[85] An example of a Tyr-selective modification

is the three-component coupling using rhodium(III) chloride and boronic acid to associate arene complexes with the *ortho* position of the Tyr residue. The authors found that this linkage is reversible by adding H_2O_2 or DTT (dithiothreitol).[86] However, in the above examples, rather harsh conditions are applied, which makes addressing other natural amino acids more compelling. Favored natural amino acids in bioconjugation and protein modification are Lys and Cys because of their nucleophilicity and accessibility. Lys is one of the most abundant amino acids in proteins, which can cause a great heterogeneity when crosslinked with common lysine-directed reagents such as N-hydroxysuccinimide esters (NHS esters) and imido esters.[81] The N-terminal amine, on the other hand, exhibits a different reactivity than the amine in the side chain of a lysine residue. The pKa of the amine in the side chain of the lysine residue has a pKa of approximately 10, while the pKa of the N-terminal amine is around 8, and thus being deprotonated under basic conditions and exhibiting a higher nucleophilicity.[87] The pKa of Cys is similar to that of the N-terminal amine, but compared to Lys, it is less abundant in proteins, which also decreases the probability of side reactions. The relatively low abundance of Cys in proteins, its high nucleophilicity, and the preferential reaction conditions near neutral pH make Cys one of the most favorable modification sites.[81]

Cysteine directed crosslinking

For cysteine-directed crosslinking classical reagents as maleimides, haloacetyl and disulfide forming agents can be used, to name a few (see figure 1.5).[82]

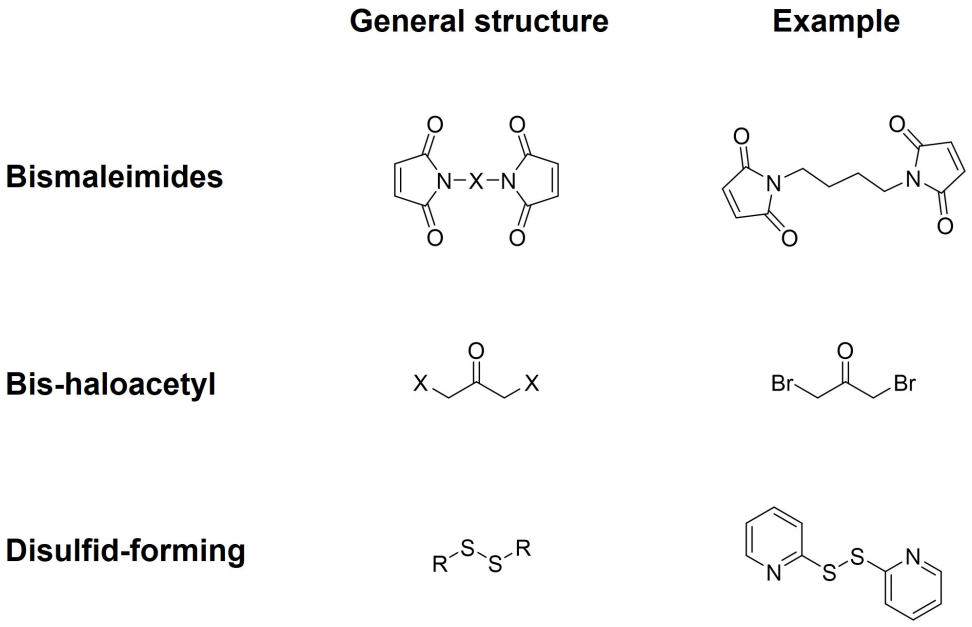


Figure 1.5: Cysteine-directed crosslinker reagents.

Maleimides are widely used cysteine-directed crosslinkers that react prefentially with thiol groups in a pH range of 7.5-8.5 to form an irreversible thioether bond. At higher pH values above 8.5, amines are the preferred reactants, and hydrolysis of maleimides may be observed. Moreover, it is important to avoid external thiol groups, for example in reducing agents such as DTT. If a disulfide bridge needs to be reduced to be available for the maleimide reaction, TCEP (Tris(2-carboxyethyl)phosphine hydrochloride) is a more

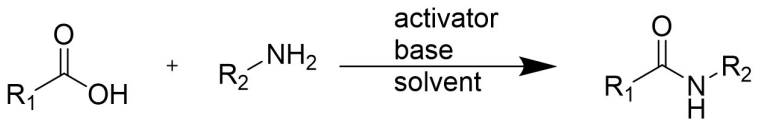
suitable reducing agent.[82] Chemically, maleimides are accessible by the reaction of an amine and maleic anhydride under elevated temperatures.[82,88]

Other cysteine-directed crosslinkers are haloacetyl-containing reagents that form a thioether bond at physiological pH through a nucleophilic substitution reaction. In structural investigations of the active site of ficin and stem-bromelain as well as papain, 1,3-dibromoacetone (DBA) was used as chemical tool, proving its value in cysteine-directed crosslinking.[89] Further bishaloacetyl derivatives were also successfully applied in structural investigations as desribed in the literature.[90–92] However, the potential of crossreacticity with other nucleophils such as in Lys and His residues and the risk of hydrolysis of these crosslinkers, make them less popular nowadays. Nevertheless, a few examples from more recent years are available.[93] For example Mayer et al. generated an artificial metalloenzyme by coupling an olefin metathesis catalyst through a bromoacetamide group to a *Methanocaldococcus jannaschii* protein which was active under acidic conditions.[94]

In case of disulfide-bond formation a disulfide-thiol interchange reaction takes place between cysteine thiolates and an oxidized disulfide reagent such as 2,2'-Bispyridyl disulfide (2-PDS) or 5,5'-dithiobis-(2-nitrobenzoic acid) (Ellman's reagent or DTNB).[83,84] First, the thiolate reacts with the disulfide bond-forming agent, resulting in the formation of a mixed disulfide. Subsequently, the second thiolate in the protein attacks the mixed disulfide, leading to the generation of a disulfide, leaving the disulfide bond-forming agents in their reduced form.[82] In the case of 2-PDS and Ellmans reagent, the reaction can be monitored by photometrically checking the formation of 2-thiopyridone and TNB (5-thio-2-nitrobenzoic acid) at 343 nm and 412 nm, respectively.[83,84] Alternatively, the oxidative reaction conditions need to be investigated, which include pH, reaction temperature, ionic strength, and folding additives. Annis et al. suggested air oxidation for the formation of disulfide bonds in peptides. This is done by diluting the protein or peptide to a low concentration between 1 µM and 1 mM in a slightly alkaline buffer in the presence of atmospheric oxygen, which is introduced into the reaction mixture by simple stirring with an open atomosphere or by blowing oxygen into the reaction mixture.[95,96]

Lactam bridging

The linkage between an amine group and a carboxyl group can form an amide bond, which is a key chemical linkage in peptides and proteins, as well as in synthetic polymers, since amide formation is one of the most commonly carried out reactions in organic chemistry. Generally, amide bonds are generated in the presence of a base and a suitable solvent by activation of the carboxylic acid by an activating group followed by nucleophilic attack of a free amine, resulting in the formation of an amide bond. The development of solid-phase-peptide synthesis and advances in the development of coupling reagents, protecting groups and resins, made it possible to synthesize peptides of moderate size (30-50 amino acids).[97]



Scheme 1.6: Reaction between a carboxyl- and amine-group forming an amide bond in the presence of an activator and a base in a suitable solvent.

When amide formation occurs intramolecularly, e.g., between two side chains of a pep-

tide or protein, it is referred to as lactam formation. Such macrocyclic peptides find wide application in diagnostics, as peptide therapeutics or in drug delivery. Cyclized peptide therapeutics tend to have advantageous properties compared to linear peptides, including increased specificity, affinity, thermostability, and proteolytic stability *in vivo*, as well as potential enhancement of bioavailability and pharmacological potency.[98,99] In addition, large macrocycles have the potential to act as inhibitors of difficult-to-target protein-protein interactions due to their conformational constraint and proteolytic stability.[99,100] For the macrocyclization of peptides, different strategies were developed such as backbone cyclization, sidechain-to-backbone cyclization and sidechain-to-sidechain cyclization.[98] Various strategies can be used for cyclization, including synthetic methods such as native chemical ligation (NCL), Staudinger ligation, the use of orthogonal protecting groups in SPPS (solid-phase peptide synthesis), or chemoenzymatic methods, e.g., using naturally occurring cyclases. In NCL, the peptide to be cyclized must have a C-terminal α-thioester and a Cys residue at the N-terminus to undergo transthioesterification followed by an S-N acyl shift leading to a new amide bond. In contrast, traceless Staudinger ligation requires an azide-functionalized N-terminus and a phosphinothiol modification at the C-terminus. The phosphinothiol and azide form an iminophosphorane intermediate that further reacts with the thioester to form a native amide bond upon H_2O loss.[101] Ha et al. demonstrated the use of traceless Staudinger ligation to generate cyclized medium-sized peptides.[102] Since chemical methods have sequence dependence, chemoenzymatic methods should also be considered if macrocyclization is desired. For example, the transpeptidase enzyme sortase A (SrtA), which is present in Gram-positive bacteria, or intein enzymes have been used for the cyclization of peptides and proteins.[99] However, SrtA has the disadvantage of long reaction times, a long recognition sequence, and is required in high concentrations for the reaction.[99,103] On the other hand, inteins which are autocatalytic protein splicing enzymes, require the fusion of a target protein with an intein domain by genetic manipulation, which can affect the folding or solubility of the resulting protein.[99,104] An alternative is to employ the naturally occurring cyclase butelase 1, which can be extracted from the pods of the tropical plant *Clitoria ternatea*. Cyclases are involved in the biosynthesis of cyclotides, a class of are cysteine-rich cyclic peptides in plants. Nguyen et al. reported how they used Butelase 1 to successfully cyclize peptides and proteins ranging from 26 to over 200 amino acids in length. When butelase 1 recognizes the tripeptide motif N/D-H-V at the C terminus, it mediates backbone cyclization of the N or D residue from the C-terminal recognition sequence after H-V cleavage to the N-terminal residue to form a macrocycle. The reaction proved to be fast, efficient and suitable for a wide range of substrates.[99]

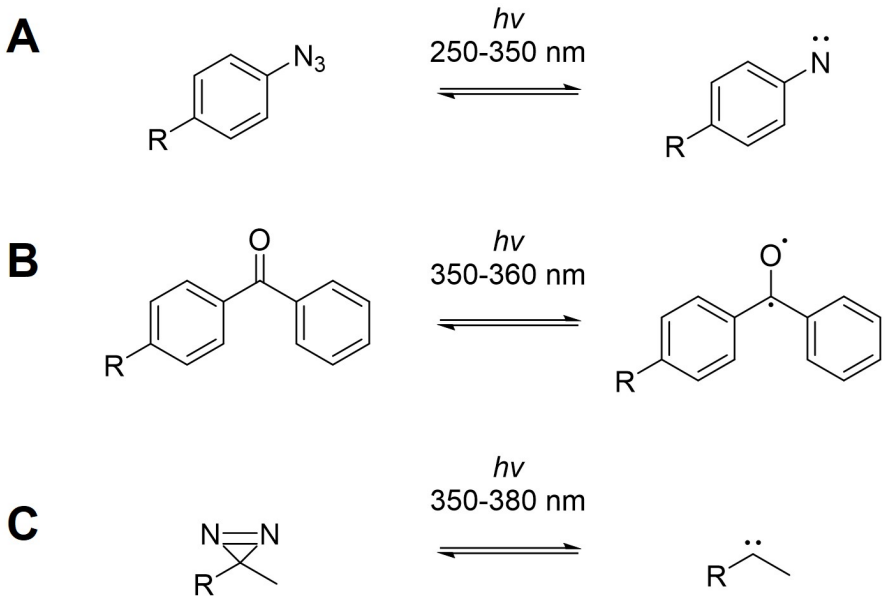
1.2.2 Protein modification via unnatural amino acids

To achieve site-specificity, the most reliable method is to introduce unnatural amino acids with bioorthogonal functional groups. Bioorthogonal reactions are characterized as being fast, selective and bio-compatible reactions which find use in e.g. bioconjugation applications.[105] With the expansion of the genetic code by Schultz and coworkers the introduction of unnatural amino acids in proteins enabled a wide range of new bioconjugation and protein modification techniques.[106] Amber suppression involves the development of orthogonal tRNA/aminoacyl-tRNA synthetase pairs that recognize the amber codon TAG as a sense codon instead of a stop codon, allowing the incorporation of non-canonical amino acids. Such tRNA/aminoacyl-tRNA synthetase pairs are commercially available for

mammalian as well as bacterial protein expression systems, making this method feasbile for protein engineering and research.[107]

Photocrosslinking

Photocrosslinking or photoaffinity labeling is a commonly used technique in chemical biology when, for example, studying protein-protein interactions in living cells and, in particular, detecting transient interactions.[108,109] The methods can also be used in drug discovery for the identification of new targets and to investigate protein-ligand interactions. Photocrosslinking uses a photogroup that forms a reactive intermediate when irradiated with light of a specific wavelength, which subsequently reacts with the target molecule in close proximity.[110] Due to the numerous possibilities of incorporating such photogroups into proteins, they can be widely applied.[110,111] However, photogroups must meet several criteria to be suitable for use. These include criteria such as high stability, low steric hindrance, activation at wavelengths that cause no or only minor damage to biological molecules, formation of highly reactive intermediates, and subsequent stable bond generation.[112] The three prominent photogroups are arylazides, benzophenones and diazirines. Upon UV irradiation, the photogroups form a reactive intermediate as shown in scheme 1.7. All photogroups have their advantages and disadvantages because of differences in size, photoaffinity yield, and potential byproducts, so which photogroup to choose depends on the application.[113]



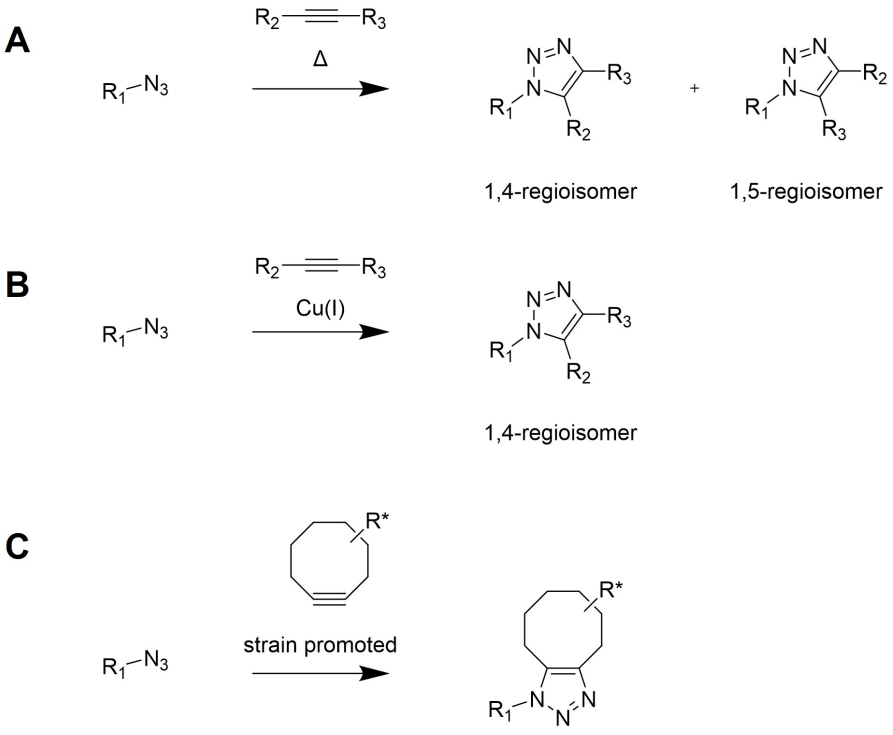
Scheme 1.7: Activation mechanism of three photogroups. A: Arylazide. B: Benzophenone. C: Diazirine. Based on [113]

Upon UV irradiation of aryl azides, a nitrene species is formed with loss of nitrogen. The reactive species reacts preferentially by insertion into a C-H or X-H bond (X=heteroatom) of a wide range of biomolecules in close proximity. However, the shorter wavelengths used to activate arylazides are potentially harmful to biological molecules. Another disadvantage is that nitrenes have a lower photoaffinity yield compared to carbenes. A possible byproduct can be the formation of ketimine if the nitrene does not react during

its lifetime of about 100 µs. Azides can also be reduced to amines, losing the ability to photocrosslink.[110,113] Another class of photogroups are benzophenones, which form a reactive diradical upon UV irradiation. Compared to arylazides, longer wavelengths are sufficient for activation, minimizing the risk of damage to biomolecules. The drawbacks of benzophenone, for example, are the longer reaction time and the rather bulky side chain.[110] Diazirenes, on the other hand, exhibit small moieties and can also be activated with longer wavelengths of 350-380 nm.[113] Upon exposure to UV irradiation, a carbene is generated, which readily undergoes insertion into C-H and X-H (X = heteroatom) bonds present in biomolecules nearby. The reactivity of diazirine can be fine-tuned by functionalization with trifluoromethyl and aromatic groups.[114]

Click chemistry

In 2001, Sharpless, Kolb and Finn defined the term click chemistry as reactions that meet strict criteria, such as high yield, modular application, no or easily separable byproducts, simple purification by non-chromatographic methods, regioselectivity, readily available reactants, insensitivity to oxygen or water, and performance under simple reaction conditions.[115]



Scheme 1.8: Click chemistry reaction types. A: Huisgen 1,3-dipolar cycloaddition reaction. B: Copper-catalyzed-azide-alkyne cycloaddition (CuAAC). C: Strained promoted azide-alkyne cycloaddition (SPAAC).

An important class of click chemistry reactions are 1,3-dipolar cycloadditions, which were introduced by Huisgen in the 1960s and unite two unsaturated reactants to five- and six-membered heterocycles.[115,116] A prominent example is the reaction between an azide acting as a 1,3-dipole and a dipolarophilic alkyne, forming a five-membered ring at elevated temperatures. However, only terminal alkynes with an electron deficit are

suitable for this reaction, and another drawback is the lack of regioselectivity, since both 1,4- and 1,5-regioisomers are formed.[115,117] Further development of the reaction by Fokin, Sharpless[118] and Meldal[119] led to the nowadays widely used copper-catalyzed azide-alkyne cycloaddition (CuAAC), in which copper is used as a catalyst so that the reaction proceeds at lower temperatures, tolerates a variety of functional groups, and yields only the 1,4-regioisomer as a product. Later, it was reported that the 1,5-regioisomer can be obtained by using ruthenium as a catalyst (RuAAC).[120] Since metal-catalyzed reactions can be disadvantageous for biological applications due to their potential toxicity in biological systems, metal-catalyst-free click reactions have been developed. These types of reactions should be bioorthogonal, i.e., they should be stable under biological conditions and not interfere with biological processes. In order to exploit the potential of azide and alkyne coupling in a biological system, since these functional groups do not occur in a biological system, a metal-free cycloaddition of these functional groups was developed.[121] In 2004 Bertozzi and co-workers demonstrated that strain-promoted [3+2] azide-alkyne cycloaddition is a suitable reaction for covalent modification in living systems, allowing a metal-free reaction with ring strain being the alkynes activator.[122] In scheme 1.8 all three reaction types are illustrated. In 2022 Carolyn R. Bertozzi, Morten P. Meldal and Karl Barry Sharpless were awarded with the Noble Prize in chemistry for the development of click chemistry and bioorthogonal chemistry.[123]

1.3 Ligand-based affinity chromatography

Affinity chromatography is a versatile technique that can be used for protein purification or pull-down experiments. In general, the isolation and purification of proteins is obtained by various properties such as size, charge, hydrophobicity or affinity. In the latter case, separation is achieved by selective and reversible interactions with a specific interaction partner, e.g., a ligand in ligand-based affinity chromatography. For this purpose, a ligand is immobilized on a solid support that can selectively capture the target protein from a complex biological mixture (e.g. cell lysate).[124] Immobilization of a ligand on a solid support is not only useful for protein isolation and purification, but can also be applied in a pulldown setup. When the ligand matrix is exposed to a lysate followed by a wash step to remove non-specifically bound molecules, MS-based analysis of the resulting fraction can lead to the identification of off-targets or interacting partners.[125]

1.3.1 Generation of an affinity matrix

For ligand-based affinity chromatography, a suitable matrix must be selected that can be functionalized to immobilize the desired ligand. The ideal matrix would be inexpensive, inert to the complex biological environment to which it will be exposed, modifiable to immobilize a ligand, and would need to allow solutes to access the ligands. However, the ideal matrix material that meets all requirements does not exist. Therefore, the most suitable matrix must be chosen depending on the purpose. One matrix that has been popular since the 1960s is agarose because it is inexpensive, has high stability over a wide pH range, has low nonspecific binding, and has a relatively large pore size.[126] In contrast, agarose is not suitable as support material in high-performance liquid chromatography (HPLC) because of its mechanical instability at high pressures. A material suitable for use in high performance affinity chromatography (HPAC) systems is silica, as it is available in a variety of pore sizes and particle diameters, is stable to mechanical stress, and can be modified for immobilization. A drawback here, is the low tolerance of a broad pH range.[124] Other options include the use of organic polymers, including polystyrene and polymethacrylate. However, it should be noted that polystyrenes, for example, must be hydrophilically coated so that their hydrophobic backbone structure does not cause them to non-specifically bind biomolecules in the sample that are not the target. Polymethacrylates, on the other hand, are more hydrophilic by default, and coatings are possible here as well. These materials are suitable for HPAC applications, and offers the additional advantage of stability across a wide pH range. If a support with a suitable possibility for ligand immobilization is selected, the ligand and especially its immobilization site must be chosen carefully. Immobilization of the ligand could lead to binding disruption due to steric hindrance. To avoid this risk, spacers can be installed between the solid support and the linker, which chemically are often alkyl or polyethylene glycol chains.[127] For targeting FKBP51, SAFit1 and compound MBa377 are suitable tool compounds (see figure 1.9). The development of the SAFit1 ligand is described in chapter 1.1.2. MBa377 is a low-affinity derivative of SAFit1, lacking the Top-group.

The affinity ligand is usually immobilized by a coupling reaction that depends on the active groups (e.g., amino, hydroxyl, aldehyde, thiol, and carboxyl groups) on the matrix and the complementary group in the ligand molecule distant from the target binding site. In this case, an hydroxylated polymethacrylate and amine-functionalized matrix was used to immobilize the ligands shown in figure 1.9 through the free carboxy-groups.

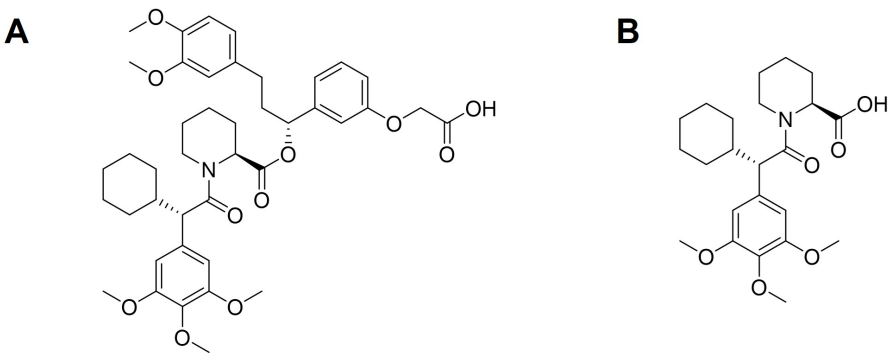
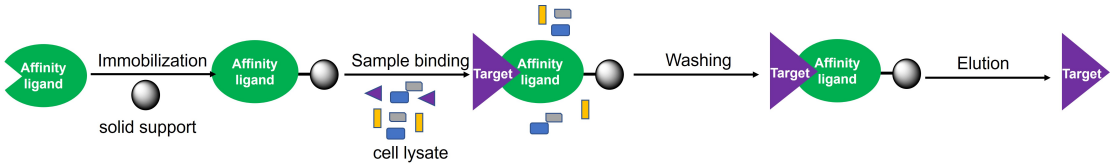


Figure 1.9: Ligands to immobilize for affinity chromatography. A: chemical structure of SAFit1 and B: chemical structure of MBa377.

1.3.2 Principles of affinity chromatography

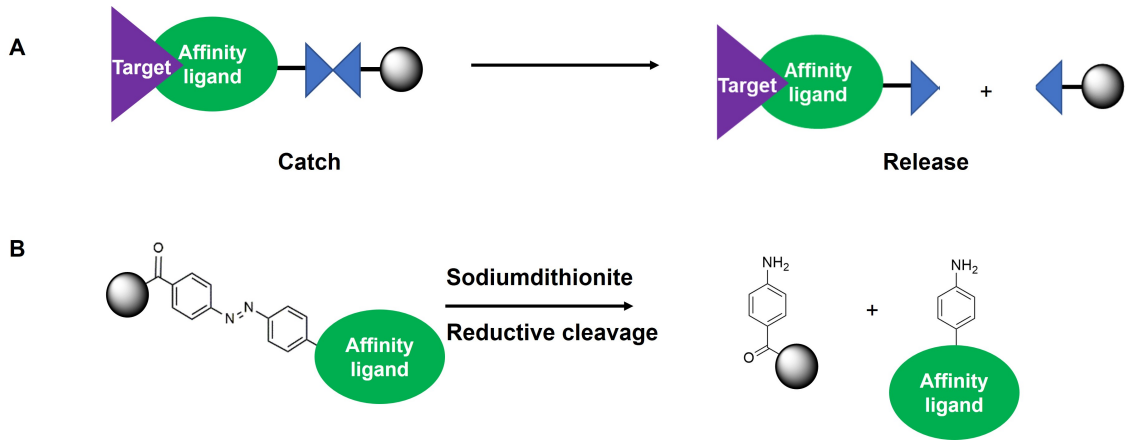
The overall procedure usually consists of three main steps: target binding, washing and elution. In the sample binding step, the target protein from a complex biological mixture binds specifically to the affinity matrix. A washing step is then performed to remove weakly bound and unrelated proteins before the elution step, in which the bound target protein is released from the interaction partner.[127]



Scheme 1.10: Wokrflow in affinity chromatography including the sample binding, washing and elution steps.

Sample binding depends on the affinity ligand chosen for immobilization. In general, there are two categories of ligands: group-specific and highly specific affinity ligands. Group-specific affinity ligands such as protein A or metal-ion chelates bind to a group or family of molecules, as the name implies. Highly specific affinity ligands such as antibodies and specific ligands, on the other hand, bind to only one target molecule and perhaps a few closely related molecules. In both cases, sample binding is followed by a wash step to isolate and purify the target molecule from a complex biological mixture (e.g., cell lysates). The washing step has a major impact on the purification success, and the ideal conditions need to be evaluated. Too harsh washing conditions can lead to dissociation of the target molecule of the immobilized affinity ligand, but if the washing step is not performed thoroughly, the purification effect will suffer. To increase the purity of the sample, additives such as detergents (Tween 20, Triton X-100 or SDS) or specific competitive binders (e.g. imidazole in IMAC) can be used at low concentrations. The ionic strength of the wash buffer must also be evaluated and optimized. Once all non-specific binders have been removed, the next step is to elute the target molecule by disrupting the non-covalent interaction between immobilized ligand and the target molecule. Native and denaturing conditions can be selected for elution. Elution under denaturing conditions is not selective, therefore the degree of purity is lower compared to specific and native elution. Urea or guanidiniumchloride, for example, can be used as denaturants, or organic solvents can be

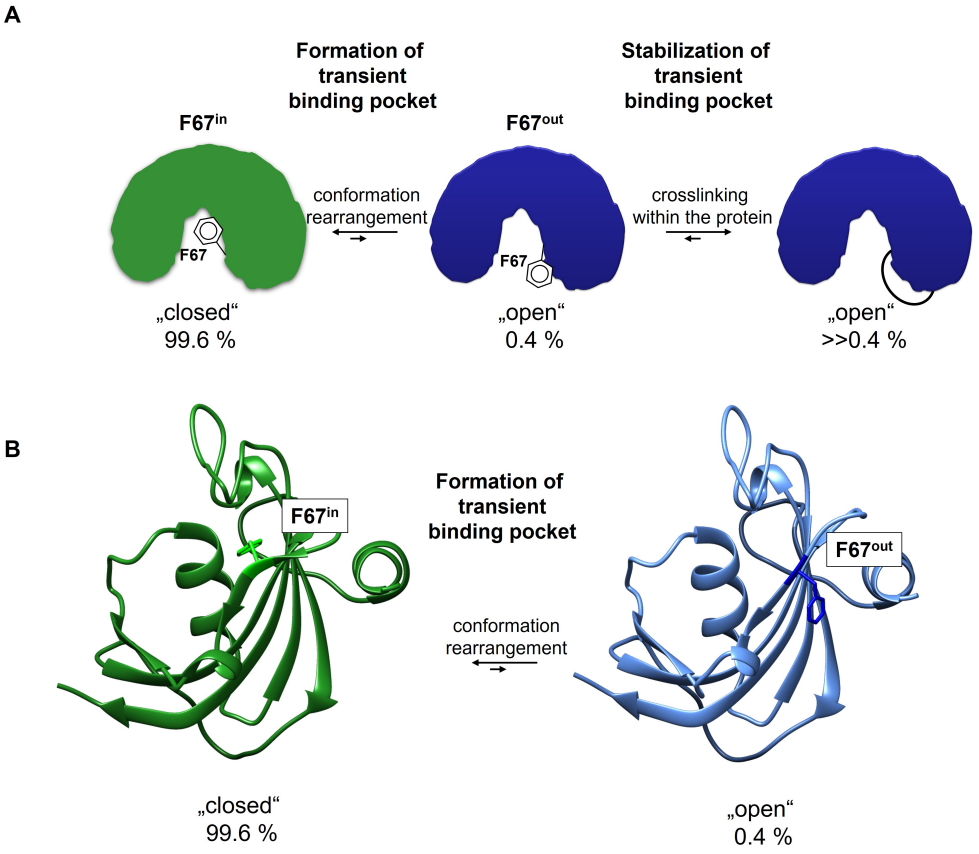
used. For specific elution under native conditions, competing binding substances or ligand homologs can be employed.[127,128] For elution, the catch and release approach is also a valid and useful technique. The incorporation of a so called cleavable linker with a reactive function between the solid support and the affinity ligand enables the site-specific release of the target bound to the ligand in the final step of affinity chromatography. Accordingly, the introduced functional groups and conditions for cleavage must be biocompatible and selective. The cleavable agent should be efficient at a low concentration and the formation of by-products should be minimized. However, it must be mentioned that the immobilized ligand matrix does not regenerate after elution, so other methods must be considered if the amount of affinity ligand is limited. Nevertheless, the method offers high selectivity, so different cleavable linkers have to be considered depending on the purpose.[129,130] They are divided into three main groups: enzymatic, photolytically and chemically cleavable linkers. In enzymatic cleavable linkers, the natural enzyme-substrate specificity is exploited and an enzyme recognizes and selectively cleaves the recognition unit, releasing the ligand-bound target. Moreover, photolabile groups incorporated in a cleavable ligand allow the release of the target by UV irradiation. For example, nitroaryl derivatives are frequently used cleavage moities. But the limits here are the uncontrolled temperature rise and possible molecular damage due to UV light. On the other hand, chemically cleavable linkers, such as oxidation-sensitive (vicinal diolic compounds) or reduction-sensitive (diazobenzenes) units, need a chemical trigger to release the target.[130]



Scheme 1.11: A: catch and release approach. Cleavable unit is cleaved by a trigger for releasing the ligand-bound target. B: example of a reduction-sensitive cleavable linker. Diazobenzene as cleavable unit, release induced by sodiumdithionate.

2 Aim

FKBP51 is a promising drug target for chronic pain[58] , obesity[56,57] and depression[54,55]. Therefore, the development of potent and selective inhibitors is of high interest. The main issue in FKBP51 drug development is the selectivity against its closest homolog FKBP52. Conceptually, selectivity can be achieved by binding to a FKBP51-specific but energetically unfavorable binding pocket characterized by the F67-out conformation.[59]

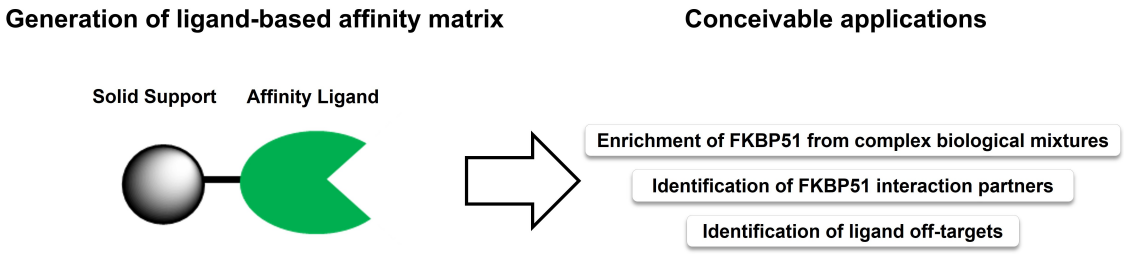


Scheme 2.1: In solution, FKBP51 is mainly present in the F67-in conformation, but a minor population of FKBP51 in the F67-out conformation was detected by NMR-studies. The aim of this work is to stabilize this transient binding pocket by rational protein engineering. A: schematic formation and subsequent stabilization of the transient binding pocket. B: X-ray structures of the FKBP51^{FK1} F67^{in} (PDB: 4DRK, green) and F67^{out} (PDB: 4TW7, blue) conformation.

By serendipity ligands were discovered and optimized that bind selectively the F67-out conformation of FKBP51, but those ligands lack ligand efficiency and physicochemical properties for CNS-directed (central nervous system) drugs, hence new ligand scaffolds

are required. The aim of this project is to generate FKBP51 constructs with a stabilized F67-out-like conformation to reduce the energetic penalty associated with binding to the F67-out state. According to NMR studies[131] a small amount of FKBP51 is already present in solution in the F67-out conformation, and rational protein engineering, including techniques such as photocrosslinking, click chemistry, and cysteine-directed crosslinking, should shift the equilibrium of the conformational rearrangement to the F67-out site by freezing this conformation with a newly formed covalent bond (see scheme 2.1). The goal of this work was to investigate whether an engineered FKBP51 construct with a locked F67-out conformation would lead to an improvement in the binding affinity of conformation-specific ligands. If so, these constructs could in turn allow screening of weak fragments for this desirable conformation, as initial fragment hits might be overlooked in screenings with the wild type due to the usually low binding affinity.

The second project within this work is to generate a ligand-based affinity matrix and establish a sufficient protocol to perform affinity chromatography. This would allow applications such as identification of FKBP51 interaction partners, enrichment of FKBP51 from complex biological samples for further characterization, including identification of post-translational modifications and identification of ligand off-targets.



Scheme 2.2: Aim of the second project of this work: generation of a ligand-based affinity chromatography matrix.

3 Results & Discussion

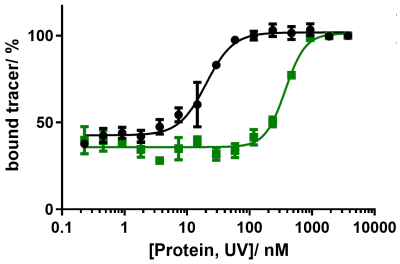
3.1 Stabilization of the FKBP51^{FK1} F67-out conformation

3.1.1 Stabilization by point mutations

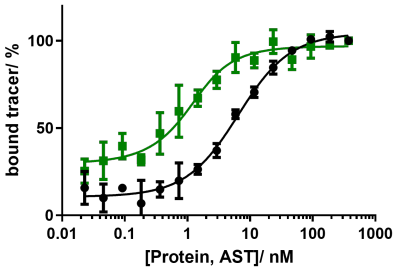
The seemingly simplest solution to generate a stabilized FKBP51^{FK1} F67-out conformation is to replace the phenylalanine in position 67 with a smaller amino acid. Indeed, several F67 mutants have been tested in the past. One of the most significant of these is F67V. In 2015, Gaali et al. designed artificial ligands to specifically bind this mutant and by chance discovered that the ligand scaffold also binds FKBP51^{FK1} WT but not FKBP52^{FK1} WT. This bump and hole approach led to the first selective FKBP51 ligands.[59]

FKBP51^{FK1} F67G

The smallest amino acid, glycine, was tested in position 67. The mutation was introduced by site directed mutagenesis according to protocol 5.1 and the protein was subsequently expressed in *E.coli* (see protocol 5.2). To evaluate whether a variant is suitable for the overall aim of fragment screening to find a new ligand scaffold for new FKBP51 inhibitors, the K_d value of a tracer based on SAFit1 is compared to that of the wild type protein. As described in section 2, a higher binding affinity is expected for the stabilized variants because the F67-in/F67-out equilibrium is shifted toward the F67-out conformation. First, active site titration was performed to determine the concentration of the active site with a fluorescein labeled tracer based on SAFit1. In figure 3.1 the results of the active titration are summarized. It is noteworthy that the AST concentration of FKBP51^{FK1} F67G is significantly lower compared to the UV-determined protein concentration. Approximately 7 % of the active site of the protein could be determined in the assay. This difference between the UV and AST protein concentrations can be explained by the low protein stability, which was evident from the precipitation of the protein in solution. Precipitation issues already occured during the protein purification steps, resulting in a low protein yield of only 1 mg FKBP51^{FK1} F67G from a 1 L *E.coli* culture. The low yield and instability make the F67G mutant an unsuitable candidate for later fragment screening, even though an approximately 9-fold change in K_d was measured by a fluorescence polarization assay compared to the wild type protein (see figure 3.1). The FKBP51^{FK1} WT protein used here was produced and purified by Dr. Christian Meyners.



Active site titration	FKBP51^{FK1} WT	FKBP51^{FK1} F67G
EC_{50}/ nM	20±1.5	373±24
c(Protein, UV)/ µM	350	318
c(Protein, AST)/ µM	563	24



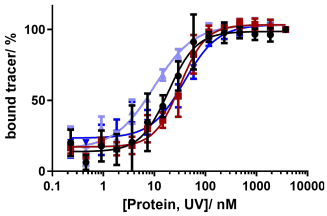
Binding curve	FKBP51^{FK1} WT	FKBP51^{FK1} F67G
K_d/ nM	6±0.7	0.7±0.2

Figure 3.1: Active site titration (AST) and binding curves of FKBP51^{FK1} WT (black curves) and FKBP51^{FK1} F67G (green curves) with a SAFit1 based tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 50 nM for AST and 1 nM for binding curves. Protein concentration range: 230 pM – 3750 nM for AST and 23 pM - 375 nM for binding curves.

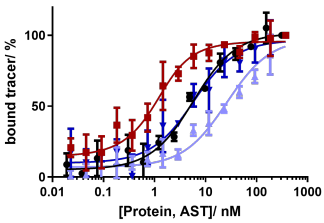
FKBP51^{FK1} F67Y and additional mutations

The FKBP51^{FK1} F67Y variant was produced and purified by Dr. Christian Meyners. The additional hydroxyl group of tyrosine compared to phenylalanine should keep the residue out of the binding pocket, because of steric reasons. The binding curve in figure 3.2 shows, that mutating F67 to a tyrosine, leads to a higher binding affinity comparable to the fold change in K_d of F67G. However, FKBP51^{FK1} F67Y is more stable in solution and thus has the potential of being a screening tool. To see, if the mutant can be further stabilized, the lysine residue in position 58, which is located on the β-sheet below position 67, is replaced by aspartic acid or asparagine to form a hydrogen bond for a improved stabilization. This work was conducted by Marco Pieroni for his bachelor thesis under my supervision. However, the introduction of asparagine at position 58 leads to destabilization, abolishing the positive effect of the F67Y mutant, so that FKBP51^{FK1} F67Y/K58N ends up having the same K_d as the wild type. In case of aspartic acid in position 58, an even lower affinity of FKBP51^{FK1} F67Y/K58D and the SAFit1 tracer is observable. It is also notable, that in the prior AST experiment, more active sites were detected than expected. This could be an indication of partial denaturation of the protein in solution and thus a higher probability of the tracer binding unspecifically to the protein and giving a distorted result. It is not trivial to make assumptions about the behavior of the proteins depending on introduced mutations.

The idea of additional hydrogen bond formation for stabilization, as shown in the modified PDB structure in figure 3.3, cannot be confirmed by the available data. It is possible that the side chains of the amino acids are at an unfavorable distance and geometry



Active site titration	FKBP51^{FK1} WT	FKBP51^{FK1} F67Y	FKBP51^{FK1} F67Y/K58D	FKBP51^{FK1} F67Y/K58N
EC_{50}/ nM	19±3	32±3	10±1	39±5
c(Protein, UV)/ µM	2857	1350	357	11
c(Protein, AST)/ µM	3759	1055	892	7



Binding curve	FKBP51^{FK1} WT	FKBP51^{FK1} F67Y	FKBP51^{FK1} F67Y/K58D	FKBP51^{FK1} F67Y/K58N
K_d/ nM	5±0.9	0.7±0.2	26±5	5±1

Figure 3.2: Active site titration (AST) and binding curves of FKBP51^{FK1} WT (black curves), FKBP51^{FK1} F67Y (red curves) and the variants FKBP51^{FK1} F67Y/K58D or F67Y/K58N, respectively (dark and light blue curves) with a SAFit1 based tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 50 nM for AST and 1 nM for binding curves. Protein concentration range: 230 pM – 3750 nM for AST and 19 pM - 375 nM for binding curves.

to form a hydrogen bond because both amino acids are restricted in their flexibility due to their location in a β-sheet.

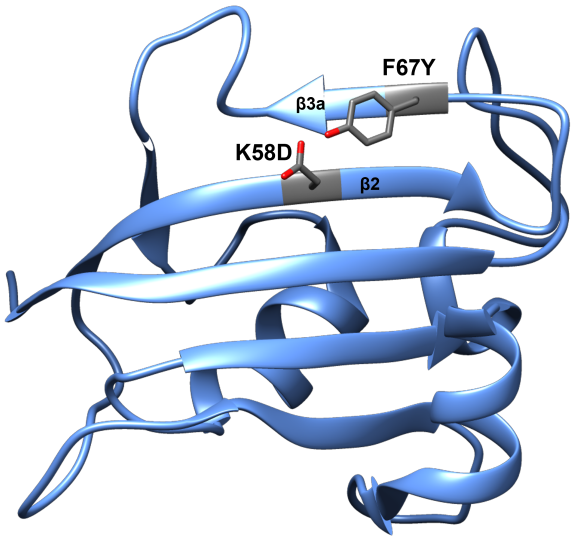


Figure 3.3: By Chimera modified PDB 4TW7. F67 was replaced by tyrosin and K58 by aspartic acid using the structure building tool in Chimera. Both residues are located in β-sheets that are parallel to each other.

FKBP51^{FK1} F67Amb/G64S

To stabilize the β-sheet formation of β3a and β2 while creating a hole in the binding pocket, the phenylalanine residue in position 67 is replaced with the unnatural amino acid derivative 4-aminomethylbenzoic acid (Amb, structure shown in figure 3.4). The advantages of the building block are the lack of a classical amino acid side chain, which allows the creation of a hole in the binding pocket, and the fact that the compound is commercially available. In addition to the F67Amb mutation, G64 is also replaced by serine in this construct. This mutant was discovered by Jorge A. Lerma Romero from the Kolmar Lab. In a high-throughput screening of yeast display libraries, he identified FKBP51 variants with higher binding affinity of FKBP51 and conformation-specific ligands. One of the highest fold changes in K_d was for FKBP51^{FK1} G64S, where the co-crystal structure of the protein revealed that a hydrogen bond formed between lysine 60 and serine at position 64, stabilizes the structure (PDB of G64S in complex with SAFit1: 7R0L).[132] The protein FKBP51$^{FK1,\ 16\text{-}140}$ F67Amb/G64S (based on the shortened 16-140aa FK1 domain sequence from Bracher et al.)[45] was generated by automated flow peptide synthesis. The 128 amino acid long protein was synthesized during my 3 month stay in the Pentelute Lab as a visiting PhD candidate using the Peptidator (see section 5.6) in collaboration with Dr. Satish Gandhesiri. The Peptidator is an in-house developed automated flow peptide synthesizer designed in the Pentelute Lab and optimized for the synthesis of long peptides up to small proteins. In chapter 3.1.4 further examples of synthesized FKBP51^{FK1} proteins are presented.

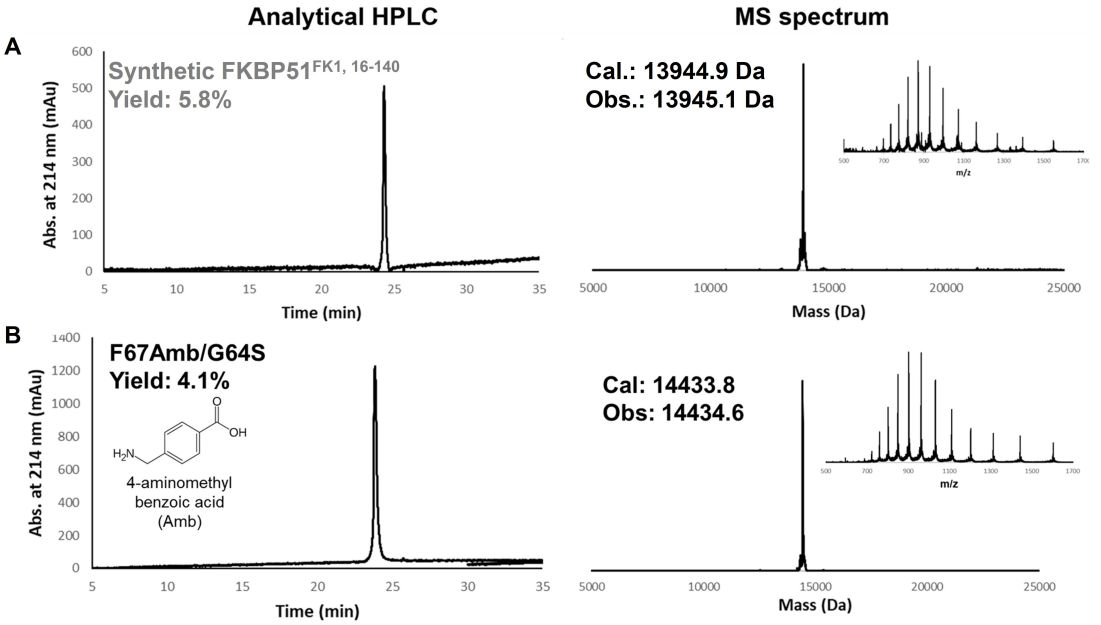


Figure 3.4: Characterization of synthesized protein variants after preperative HPLC purification. Left side: Analytical HPLC chromatograms at an absorbance at 214 nm of synthetic FKBP51$^{FK1,\ 16\text{-}140}$ and F67Amb/G64S. Right side: Deconvoluted MS data with inserts showing the MS data prior to deconvolution. For A and B: Synthetic FKBP51$^{FK1,\ 16\text{-}140}$ (above) and F67Amb/G64S (below).

As a control, FKBP51$^{FK1,\ 16\text{-}140}$ was synthesized. The control variant and F67Amb/G64S differs from the wild-type sequence in the abolishment of the cysteine residues in position

103 and 107, resulting in the C103A and C107I mutations. In addition, the synthetic variants have norleucine incorporated instead of methionine as an oxidation protector (M48Nle, M97Nle, Nle=norleucine). Both purified variants were characterized by analytical HPLC and LC-MS, as shown in figure 3.4.

The observed masses of the synthesized proteins were in agreement with the theoretical ones, and yields of 4.1 % for FKBP51$^{FK1,\ 16\text{-}140}$ F67Amb/G64S and 5.8 % for FKBP51$^{FK1,\ 16\text{-}140}$ were achieved, respectively (see table 3.1).

Table 3.1: Summarized results of LC-MS data and the obtained yields.

	synthetic noCys	**F67Amb/G64S**
Calc. Mass/ Da	13944.9	14433.8
Obs. Mass/ Da	13945.1	14434.6
Mass error/ ppm	14	55
Yield/ %	5.8	4.1
Yield/ mg	17.9	7.0

The lyophilized proteins had to be refolded in the next step. Rapid dilution method was used for refolding the proteins. Different buffers with different pH values, salt concentrations and additives were tested, and the refolding buffer with the following composition was found to be the best one: 50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5 M arginine, 0.4 M sucrose, 0.75 M guanidine HCl.

To confirm the refolding success, size exclusion chromatography-mass spectrometry (SEC-ESI-MS) experiments were conducted by my cooperation partner Thomas Nehls. SEC analysis enables the seperation of different protein conformations or different proteins with various sizes. The refolded proteins and the recombinant FKBP51$^{FK1,\ 16\text{-}140}$ protein were tested under native and denatured conditions. All proteins show a significant change in elution volume when comparing between denatured and native conditions (see figure 3.5). Under native conditions, the proteins elute at a higher volume than at denatured conditions, since the proteins exhibit a smaller hydrodynamic radius at native conditions. This allow the proteins to diffuse into more pores, so a larger volume is required for elution. Since the elution volume of the synthetic proteins is the same as for the recombinant protein, we can assume that the conformation corresponds to the conformation of the native, correctly folded protein. However, for F67Amb/G64S at native conditions several peaks are observable, indicating that in solution, more conformations are present.

Table 3.2: Collision cross section values (CCS) for charge state 8+, measured with traveling wave with N_2 as collision gas. Experiment conducted by Thomas Nehls from the Lermyte Lab.

	recombinant noCys	**synthetic noCys**	**F67Amb/G64S**
$^{TW}\Omega_{N_2}$native/ Å^2	1696±3	1707±4	1784±2
$^{TW}\Omega_{N_2}$denatured/ Å^2	2703±3	2663±9	2641±17

Another technique to confirm the conformation of the refolded proteins is ion mobility spectrometry, which was also conducted by Thomas Nehls.[133] Analytes are separated in the gas phase based on their size and shape, which influences the drift time in the drift tube filled with an inert gas. Since the analyte rotates in the gas phase, only the average cross-sectional area, called the collision cross-section (CCS), can be determined. In case

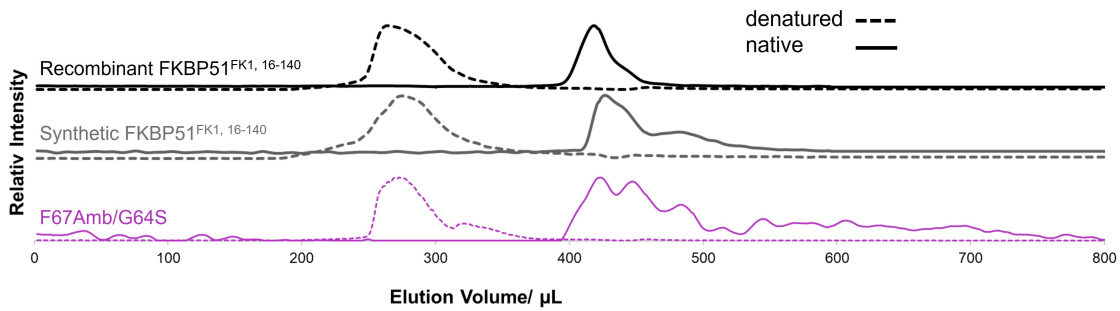
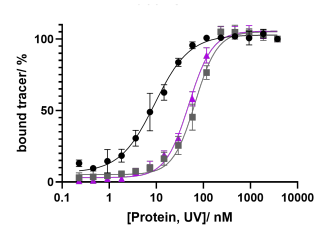


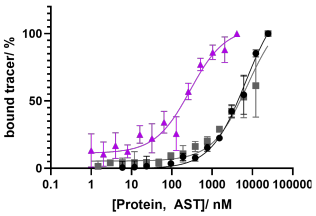
Figure 3.5: Size exclusion chromatography-mass spectrometry of purified and refolded or denatured proteins. Extracted ion chromatograms of synthetic F67Amb/G64S (violet lines) in comparison to the synthetic (gray lines) as well as recombinant (black lines) FKBP51$^{FK1,\ 16\text{-}140}$ variants. Two conditions were tested: native and denatured. The native runs were conducted with 50 mM ammonium acetate buffer pH 7 (continuous lines) and the denatured runs with 0.2 % formic acid in water (dotted lines). This experiment was conducted by Thomas Nehls from the Lermyte Lab.

of travelling wave technology a calibration is required[134,135], which provided suitable regression values for both denatured and native conditions.[136] In table 3.2 CCS values for charge state 8+ at native and denatured conditions are summarized. The recombinant and the synthetic FKBP51$^{FK1,\ 16\text{-}140}$ samples show similar values for the collision cross section, while F67Amb/G64S shows a slightly higher CCS value, but still in the same order of magnitude. Under denatured conditions the CCS values expand by approximately 1000 Å^2.

The efficiency of refolding was also tested by active site titration, and the adjusted protein concentrations were used to determine the K_d value of a SAFit1-based tracer. In this case, a low-affinity tracer characterized by the absence of the top group of the usual SAFit1 scaffold linked to fluorescein was used. The chemical structure is shown in the methods section 5.3.1. The estimated refolding yield is 22 % for synthesized FKBP51$^{FK1,\ 16\text{-}140}$ and 30 % for F67Amb/G64S. Finding efficient refolding conditions is a challenging task, especially when sample amount is limited. Many variables affect the refolding process, e.g., pH, salts and their concentration in the buffer, additives, etc. However, after an active site titration, the protein concentration can be adjusted so that the binding curves can be measured directly. The synthetic and recombinant FKBP51$^{FK1,\ 16\text{-}140}$ variants show similar K_d values for the low affinity tracer, and a 24-fold change in K_d is obtained compared to F67Amb/G64S. Nevertheless, the variant is in the end not suitable as a screening tool, since the sample size is limited and the refolding step lowers the sample yield even further.



Active site titration	FKBP51^{FK1} noCys (recombinant)	FKBP51^{FK1} noCys (synthetic)	FKBP51^{FK1} F67Amb/G64S (synthetic)
EC_{50}/ nM	9±1	67±4	49±2
c(Protein, UV)/ µM	129	143	72
c(Protein, AST)/ µM	215	32	22
Refolding yield/ %	-	22	30

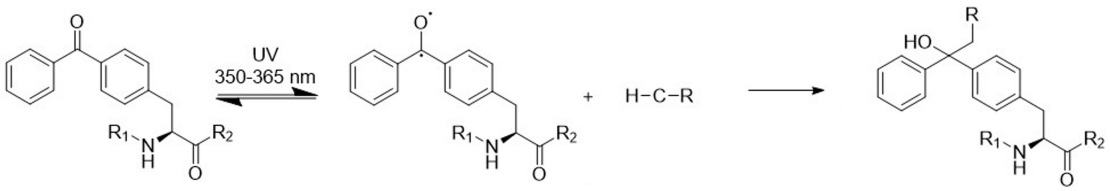


Binding curve	FKBP51^{FK1} noCys (recombinant)	FKBP51^{FK1} noCys (synthetic)	FKBP51^{FK1} F67Amb/G64S (synthetic)
K_d/ nM	6741±873	7443±1850	285±61
Fold change in K_d	1	0.9	24

Figure 3.6: Active site titration (AST) and binding curves of recombinant (black curves) and synthetic (gray curves) FKBP51$^{FK1, 16\text{-}140}$ variants, as well as the F67Amb/G64S (violet curves) with a low affinity binding tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 30 nM for AST and 1 nM for binding curves. Protein concentration range: 230 pM – 3750 nM for AST and 1 nM - 4125 nM for binding curves.

3.1.2 Stabilization by photocrosslinking

To covalently fix the side chain in position 67 to the backside of the binding pocket, a photocrosslinking unnatural amino acid is incorporated in position 67. As mentioned in chapter 1.2.2 before, three types of photocrosslinker are available: benzophenones, diazirines and azides. In this study para-benzoyl-L-phenylalanine (Bpa) was used, since this unnatural amino acid is structurally similar to phenylalannine and can be activated with longer wavelength, minimizing the risk of UV damage. After UV irradiation, a diradical is formed that can react with a C-H or X-H (X=heteroatom) bond as illustrated in scheme 3.7.



Scheme 3.7: Upon UV irradiation para-benzoeylphenylalanine forms a diradical which reacts with a C-H bond.

To incorporate Bpa into the protein, amber suppression in *E.coli* is used. This method is widely spread, hence many various plasmids coding the tRNA/aminoacyl-tRNA synthetase pair are commercially avaible at e.g. Addgene. The pEVOL-pBpF plasmid (depositing Lab: Peter Schultz. [137]) was purchased, and used in an test expression experiment as described in chapter 5.2. After transformation of the plasmid encoding the protein of

interest (FKBP51^{FK1}) and subsequently the pEVOL-pBpF plasmid into *E.coli*, three colonies were picked and analyzed for protein expression via SDS-PAGE. IPTG induces expression of the protein of interest, while L-arabinose induces expression of the tRNA/aminoacyl-tRNA synthetase pair. Another crucial additive for amber suppression is the addition of the unnatural amino acid Bpa to the bacterial growth medium.

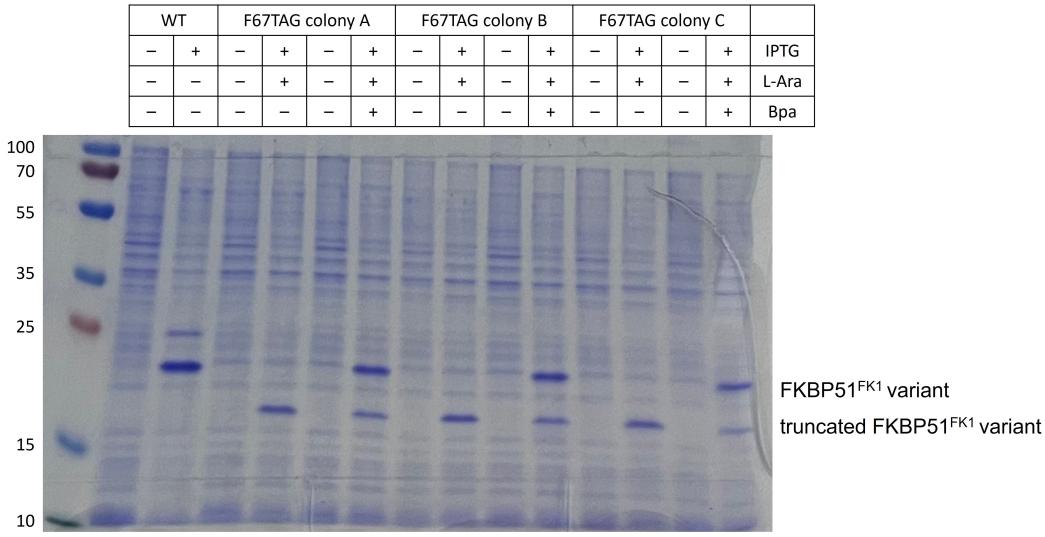


Figure 3.8: SDS-PAGE gel of the testexpression incorparating Bpa into FKBP51^{FK1} by amber suppression.

Indeed, the samples with IPTG, L-arabinose and Bpa show a protein band at the expected height, comparable to the wild-type protein band. Strikingly, the addition of IPTG results in a truncated version of the protein that terminates at amino acid 67, where the stop codon was incorporated. The truncated variant is also observed in the samples treated with all three required additives, suggesting that the incorporation of Bpa into the protein is not successful for all protein chains and a small fraction simply terminates at position 67. However, all three colonies examined show the same behavior (see figure 3.8). Colony A was secured for larger scale protein production by making a glycerol stock.

For a F67-out-like conformation stabilized by photocrosslinking, FKBP51^{FK1} F67Bpa is expressed in *E.coli* (see protocol in chapter 5.2). After IMAC (immobilized metal affinity chromatography), the protein was further purified by SEC (size exclusion chromatography). Figure 3.9 shows the corresponding chromatogram after SEC purification and the SDS-PAGE gel of the obtained fractions. SEC fraction 3 (SEC-Fr3) was selected for further purification because the truncated protein could not be separated. However, after WCX (weak cation exchange chromatography), FKBP51^{FK1} F67Bpa (fraction WCX-Fr1) was obtained relatively pure (see figure 3.9).

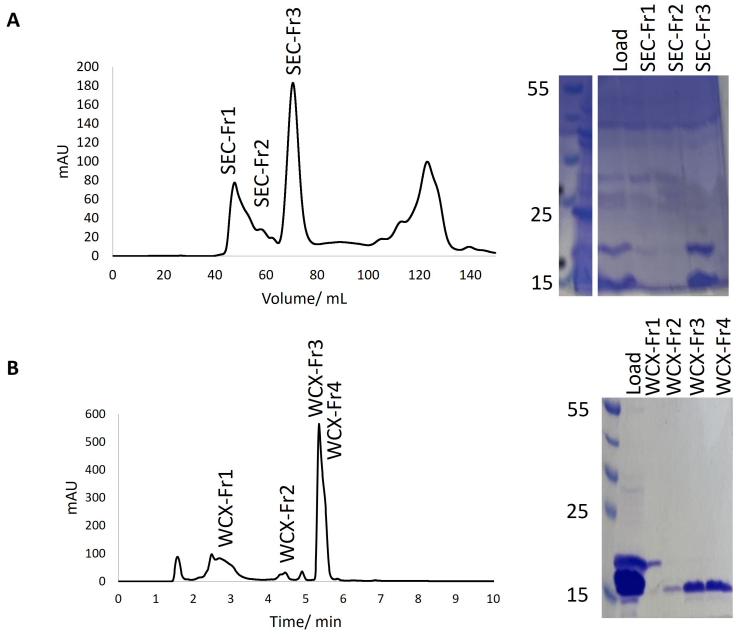


Figure 3.9: A: SEC (size exclusion chromatography) chromatogram and the respective SDS-PAGE analysis. B: WCX (weak cation chromatography) chromatogram with respective fractions analyzed by SDS-PAGE.

The MS analysis performed by Thomas Nehls (see figure 3.10) shows that the incorporation of the unnatural amino acid Bpa was successful, as the mass of the -UV sample corresponds to the theoretical mass. The +UV sample shows no difference in mass distribution in MS analysis. However, the intact mass measurements cannot confirm whether or not covalent bonding was formed upon UV irradiation. To confirm the intramolecular photocrosslinking reaction, electron transfer dissociation mass spectrometry could be performed, or alternatively, the sample could be tryptically digested and then searched for the crosslinked fragments.

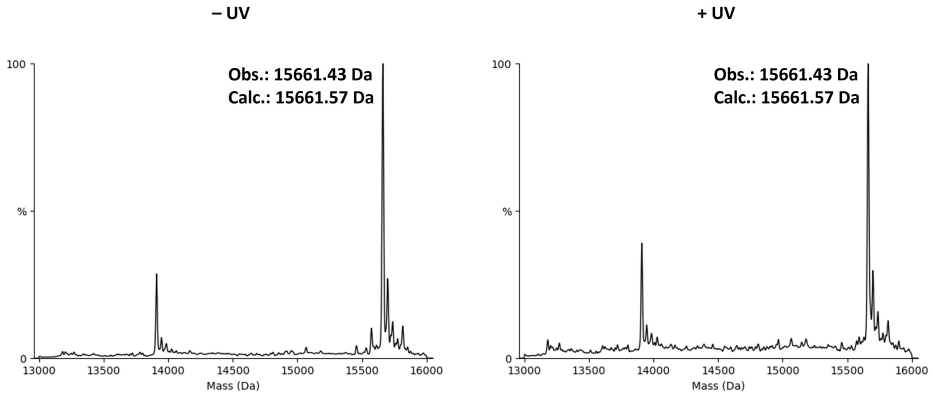
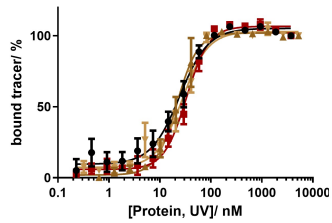
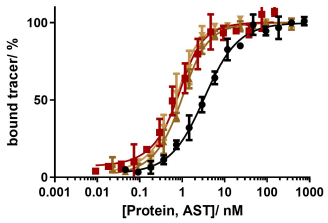


Figure 3.10: Determination of the intact mass by Thomas Nehls. The observed intact mass of the sample without UV irradiation is consistent with the theoretical mass (left side). After UV irradiation, no mass difference can be detected by MS analysis (right side).

Comparing the K_d values of the SAFit-FL tracer and the +UV and -UV samples, no difference is observed (see figure 3.11). Again, it cannot be determined with certainty whether intramolecular photocrosslinking has occurred. Either the photocrosslinking did not work and the +UV sample therefore shows the same K_d value as for the -UV sample, or the photocrosslinking has no further effect on the binding affinity despite the newly formed intramolecular crosslink. However, a fold change in K_d compared to wild type is observed, which is in the same range as the FKBP51^{FK1} F67Y variant. Another problem is that the tracer used is a high affinity tracer, so differences in binding affinity are more difficult to detect once the fluorescence polarization assay reaches its limits. At that time, no conformation-specific low-affinity tracer was available. In addition, producing this variant in large quantities is challenging because protein expression yields are lower when unnatural amino acids are incorporated by amber suppression compared with natural protein production. In addition, photocrosslinking would likely result in different crosslinked protein products that would be difficult to separate. For all these reasons, it was decided to proceed with the other planned variants. However, the established protocol for protein production with unnatural amino acids incorporated in *E.coli* could be adopted by my colleagues Sarah Engel and Maximilian Repity as they investigate the *in vitro* photocrosslinking of FKBP51 with the glucocorticoid receptor or with various yet-to-be-identified binding partners.



Active site titration	FKBP51^{FK1} WT	FKBP51^{FK1} F67Y	FKBP51^{FK1} F67Bpa -UV	FKBP51^{FK1} F67Bpa +UV
EC_{50}/ nM	25±1.5	34±1.5	22±1.0	27±2
c(Protein, UV)/ µM	20	20	64	64
c(Protein, AST)/ µM	20	15	76	59

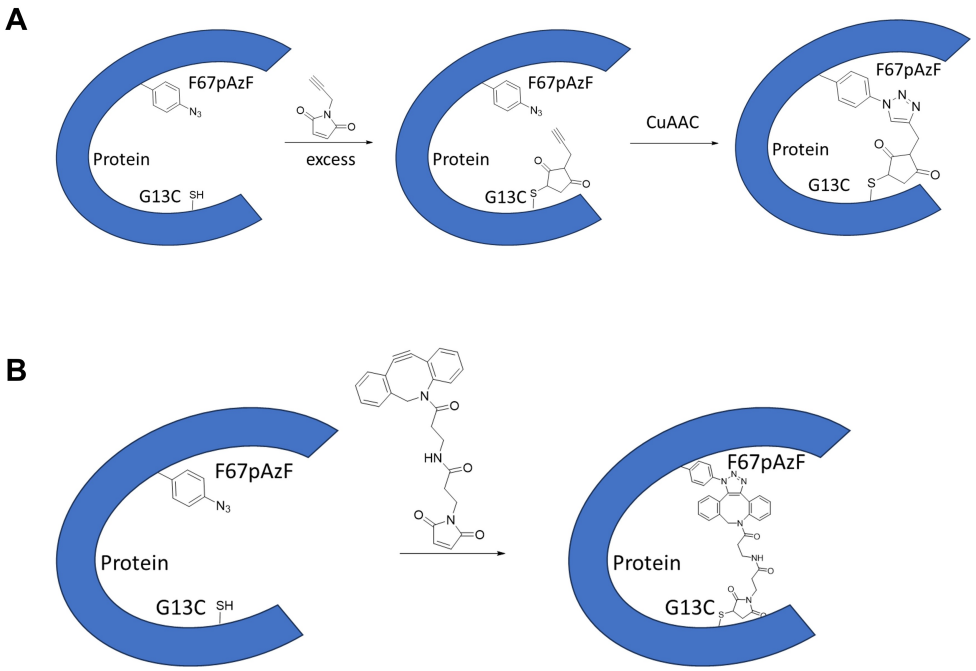


Binding curve	FKBP51^{FK1} WT	FKBP51^{FK1} F67Y	FKBP51^{FK1} F67Bpa -UV	FKBP51^{FK1} F67Bpa +UV
K_d/ nM	2.8±0.2	0.3±0.1	0.5±0.1	0.2±0.1

Figure 3.11: Active site titration (AST) and binding curves of FKBP51^{FK1} WT (black curves), FKBP51^{FK1} F67Y (red curves) and the variants FKBP51^{FK1} F67Bpa with and without UV treatment (dark brown -UV, and light brown curves +UV) with SAFit1-FL tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 50 nM for AST and 1 nM for binding curves. Protein concentration range: 230 pM – 3750 nM for AST and 9 pM - 750 nM for binding curves.

3.1.3 Stabilization by click chemistry

Two approaches of click chemistry were considered for attaching the side chain in position 67 on the backside of the binding pocket, as schematically shown in scheme 3.12. Both approaches require the incorporation of the unnatural amino acid para-azido-phenylalanine (pAzF) at position 67 as a functional group for click chemistry reactions, which was achieved by amber suppression in *E.coli*. Position 13 was chosen here as the anchor point for intramolecular fixation of the side chain at position 67. Since cysteines are suitable sites for chemical protein modifications with a variety of possible cysteine-directed reagents and crosslinkers and controllable selectivity, G13 is replaced by a cysteine. The variant is referred to here as F67pAzF/G13C, which is a monoCys variant since the natural cysteines in the protein were replaced previously (C103A, C107I). The first approach is performed in two steps. First, the protein variant F67pAzF/G13C is coupled to propargylmaleimide to introduce an alkyne moiety into the protein. In the next step, a copper-catalyzed azide-alkyne cycloaddition (CuAAC) can be performed to intramolecularly crosslink the side chains in the protein (see scheme 3.12A). Another option is to use a heterobifunctional crosslinker such as DBCO-MI (dibenzocyclooctinmaleimide). This crosslinker consists of a DBCO moiety that can undergo a copper-free cycloaddition promoted by a structurally restricted alkyne (SPAAC: strain promoted alkyne-azide cycloaddition) and a maleimide moiety that reacts specifically with cysteines at pH 7-8.[82] The reaction can be performed in one step, as shown in figure 3.12B.



Scheme 3.12: Intramolecular crosslinking of FKBP51^{FK1} F67pAzF/G13C to stabilize the F67-out-like conformation. A: Two-step intramolecular crosslinking within the protein by copper-catalyzed azide-alkyne cycloaddition (CuAAC) after coupling propargylmaleimide to the cysteine at position 13. B: DBCO-MI (dibenzocyclooctinmaleimide) is used as an heterobifunctional crosslinker to generate an intramolecular crosslink within the protein by strain-promoted alkyne-azide cycloaddition (SPAAC).

Crosslinking F67pAzF/G13C by CuAAC

Optimal CuAAC (copper-catalyzed azide-alkyne cycloaddition) conditions were explored using the monoCys variant $FKBP51^{FK1}$ E140C with a readily accessible cysteine residue at the C-terminus previously coupled with propargylmaleimide to introduce an alkyne moiety into the protein. The protein with the introduced alkyne group, which was confirmed by SDS-PAGE and MS analysis, was to react with a 10 kDa PEG polymer with a terminal azide group in the presence of copper ions and sodium ascorbate to reduce the copper ion to the active oxidation state Cu(I). In addition, various additives such as *tert*-butanol and tris(3- hydroxypropyltriazolylmethyl)amine (THPTA) as well as denaturing buffer and native buffer were tested at different concentrations and ratios. This experiment was performed by the practical student Paula Henke under my supervision.

A shift in protein band is observed in this experiment only at a minimum 40-fold excess of azide-PEG and only in the presence of *tert*-butanol, as shown in figure 3.13. The amounts of copper reagent and THPTA tested at a constant ratio of 1:5 and the buffer conditions (denatured vs. native) do not appear to affect the success of the cycloaddition.

A conceivable problem with the presented experimental setup could be the low concentration of the alkyne, since concentrations of both reaction partners of more than 10 µM are recommended in the literature. However, this can be compensated by increasing the concentration of at least one reaction partner so that it is present in a molar excess. If the reaction conditions were to be transferred to the variant F67pAzF/G13C, which is actually to be functionalized with an alkyne and crosslinked intramolecularly, the problem arises that a high protein concentration would be insufficient, since the generation of the variant is eloborate and a high concentration would promote undesirable intermolecular crosslinking.[138] In addition, the use of sodium ascorbate in biological systems can be detrimental. Copper ions are able to oxidize ascorbate by molecular oxygen, leading to the formation of peroxide and radical anions, which in turn can oxidize side chains in the protein such as Cys, Met and His. To prevent this, ligands such as THPTA should be added, or additives such as aminoguanidine or pyridoxamine.[139] THPTA is a water-soluble chelating ligand, a so called accelerator, which serves two functions. First, it protects the essential but unstable Cu(I) species from oxidation and ensures that a sufficiently high concentration is maintained. Second, it acts as a radical scavenger for radicals generated by O_2/Cu(II)/ascorbate as described before.[138] In addition, *tert*-butanol proved to be a sufficient and necessary co-solvent, probably increasing the solubility of the reactants. Native and denatured conditions were tested to examine the effect of protein and PEG polymer folding on reactivity and accessibility. SDS-PAGE was used as the readout method because successful azide-PEGylation of the protein would result in a shift of the protein band. A total of 60 conditions were tested, with the publication of Hong et al. serving as a reference example.[138,139]

Although the screening provides a reasonable starting point for the optimization of the CuAAC reaction conditions, this approach was discarded at this stage as the second approach is more feasible. The use of the DBCO-MI heterobifunctional crosslinker allows a shorter workflow and does not require additives and the associated troublesome separation of them.

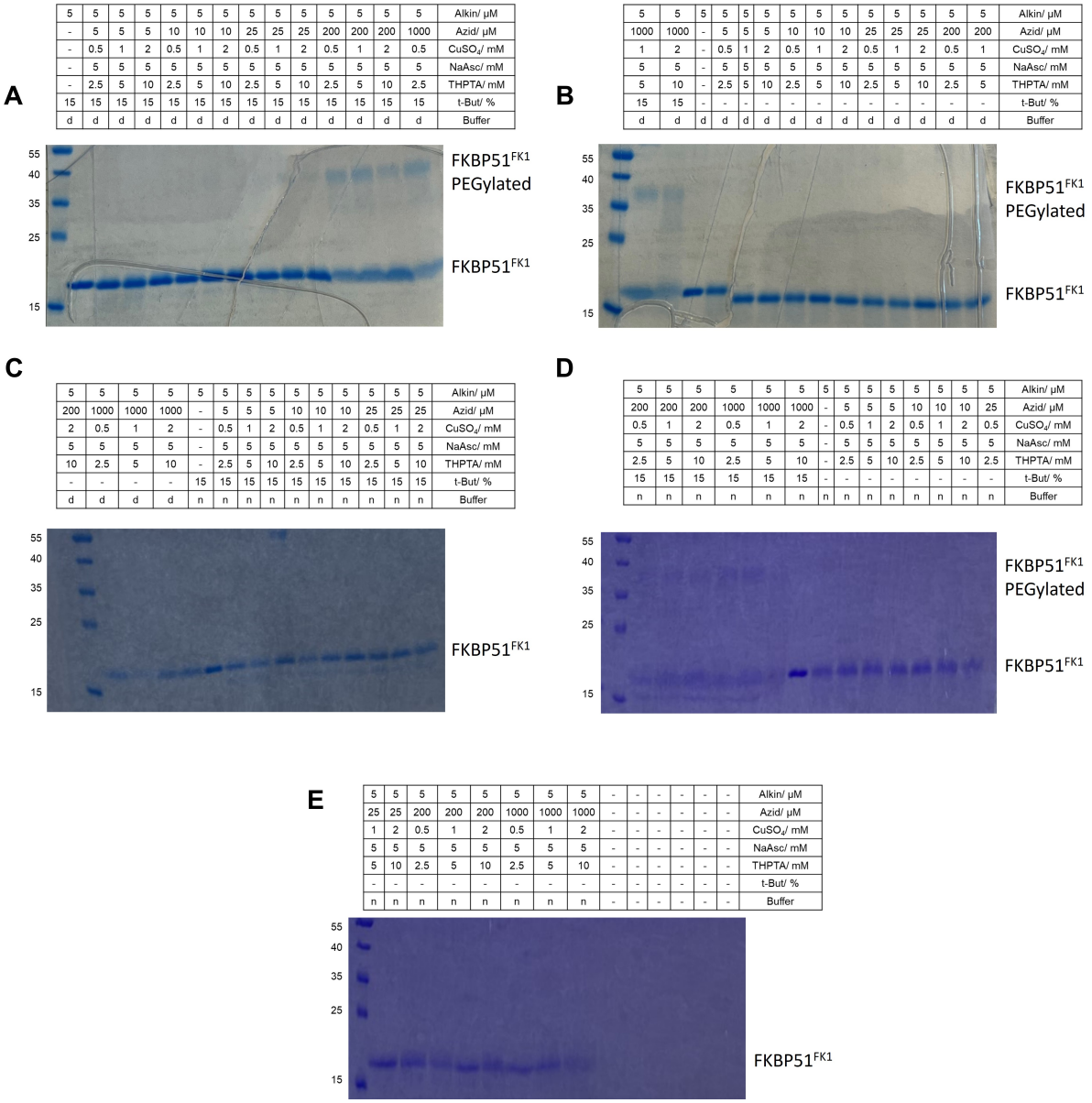


Figure 3.13: SDS-PAGE analysis of the copper-catalyzed azide-alkyne cycloaddition (CuAAC) screen to find optimal conditions for the reaction. The protein with the introduced alkyne moiety (Alkin) is used at a concentration of 5 µM. The azide used here is azide-PEG (10kDa), so a shift should be observed in the SDS-PAGE gel upon cycloaddition. The azide was tested in an equimolar ratio and in 2-, 5-, 40- and 200-fold excess. The copper concentration was also varied with THPTA (tris-hydroxypropyltriazolylmethylamine) at a constant ratio of 1:5. The sodium ascorbate concentration is kept constant at 5 mM. Also the influence of tert-butanol was tested and the protein was used in native (n) and denaturing (d) buffer. Performed by practical student Paula Henke.

Crosslinking F67pAzF/G13C by SPAAC

For the pilot experiment to investigate the reactivity of the crosslinker DBCO-MI, the monoCys protein FKBP51^{FK1} E140C was used again. This experiment was performed by the practical student Paula Henke under my supervision. To confirm the accessibility and reactivity of the cysteine in the protein, the protein was mixed with 10 kDa MI-PEG (see right side in the figure 3.14), which resulted in an almost quantitative shift of the protein band. MI-PEGylation can also be used for indirect determination of free cysteine groups. If the maleimide moiety of the DBCO-MI crosslinker reacts completely with the cysteine residue of the protein, subsequent addition of MI-PEG should not lead to a shift in the protein band, since there are no sulfhydryl groups left to react with. Indeed, lane two in figure 3.14 shows no shift in the SDS-PAGE gel after DBCO-MI followed by MI-PEG treatment. However, when DBCO-MI reacts with the protein at the cysteine site, a DBCO moiety is free to react with, for example, azide-PEG. In lane 1 in figure 3.14, a protein band shift is observed, confirming both binding to the protein and reactivity of the functional groups of the linker.

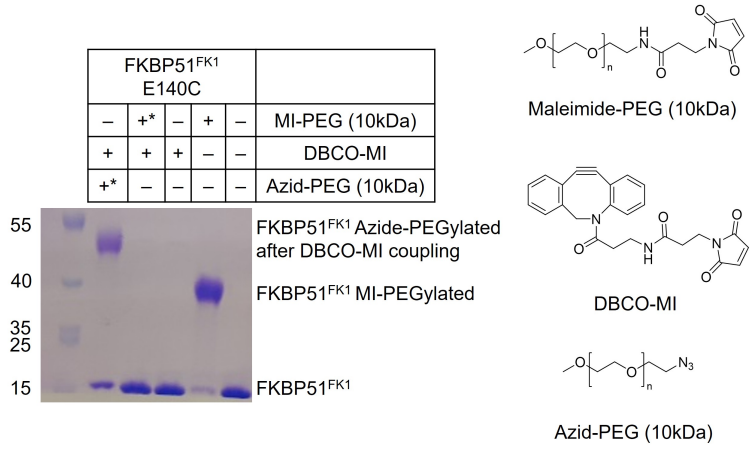


Figure 3.14: Test reaction to investigate the reactivity of DBCO-MI (dibenzocyclooctinmaleimide) using FKBP51^{FK1} E140C as a model protein. The success of DBCO-MI coupling to the protein could be determined indirectly by treatment with azide PEG (10 kDa) or by detection of free cysteines with MI-PEG (10 kDa). After reaction of the PEG reagents with the protein, a shift in the SDS-PAGE gel is observed. The * indicates that the reagent used was added after DBCO-MI coupling. The chemical structures of the reagents used are displayed on the right.

Since DBCO-MI has emerged as a suitable crosslinker for generating a stabilized F67-out-like conformation the corresponding protein variant F67pAzF/G13C had to be produced and purified. Once again, amber suppression in *E.coli* is the method used, since the required plasmid coding for a respective tRNA/aminoacyl-tRNA synthetase pair is commercially avaible at Addgene. The pEVOL-pAzF plasmid (depositing Lab: Peter Schultz [111]) was purchased, and used in a test expression experiment. After subsequential transformation of the plasmid encoding the protein of interest with the corresponding TAG mutation and the plasmid pEVOL-pAzF, a colony was picked and analyzed for protein expression by SDS-PAGE. The plasmid encoding the protein of interest is induced by IPTG, whereas expression of the tRNA/aminoacyl-tRNA synthetase pair is induced by l-arabinose. In

addition, it is essential to add the unnatural amino acid to the bacterial growth medium. Figure 3.15 shows the expression pattern of the F67pAzF/G13C protein variant and the wild-type FKBP51^{FK1} variant for reference. After addition of IPTG, a strong protein band is observed at approximately 25 kDa for the wild-type protein. At the same level, a protein band can be observed for the F67pAzF/G13C variant when all three required additives (IPTG, L-Ara, pAzF) are added, while the absence of the unnatural amino acid leads to the production of the truncated protein, stopping the biosynthesis of the protein chain at position 67. Since the test expression of the incorporation of pAzF was successful, the protein was produced on a larger scale and mixed with the crosslinker to form an intramolecular crosslink between position 67 and 13.

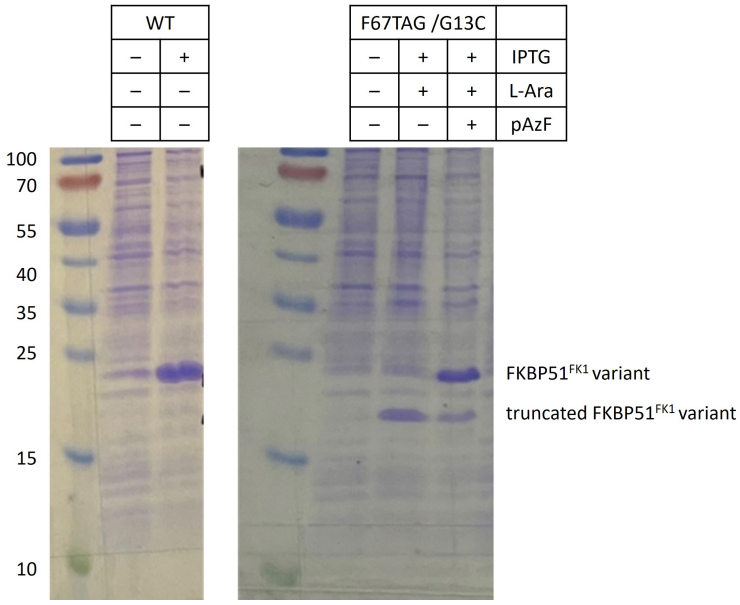


Figure 3.15: SDS-PAGE analysis of the testexpression incorporating pAzF (para-azido-l-phenylalanine) into FKBP51^{FK1} by amber suppression. The test expression was conducted by the practical student Paula Henke.

The protein sample crosslinked with DBCO-MI was analyzed by MS. This showed that not only did the desired intramolecular crosslinking occur, but also that a protein variant that reacted twice with the crosslinker was obtained. In addition, the azide moiety was partially reduced to an amine, resulting in a protein variant that was no longer amenable to intramolecular crosslinking with DBCO-MI. As a control variant, the monoCys variant G13C was produced and coupled to DBCO-MI as well. The intact MS anaylsis confirmed the desired product to be present as shown in figure 3.16A&B. Both samples were also analyzed in an IM-MS (ion mobility mass spectrometry) and CIU (collision induced unfolding) experiment. In an IM-MS experiment, proteins in the gas phase pass through an electric field and a controlled gas flow. During this process, the analyte, in this case, the protein, is propelled by a travelling wave through a drift tube filled with nitrogen molecules serving as drift gas. Due to soft collisions, the gas molecules exhibit a resistance that depends on the cross-sectional area of the analyte. Consequently, smaller or more compact analytes exhibit faster drift times compared to analytes with a larger collision cross-section.

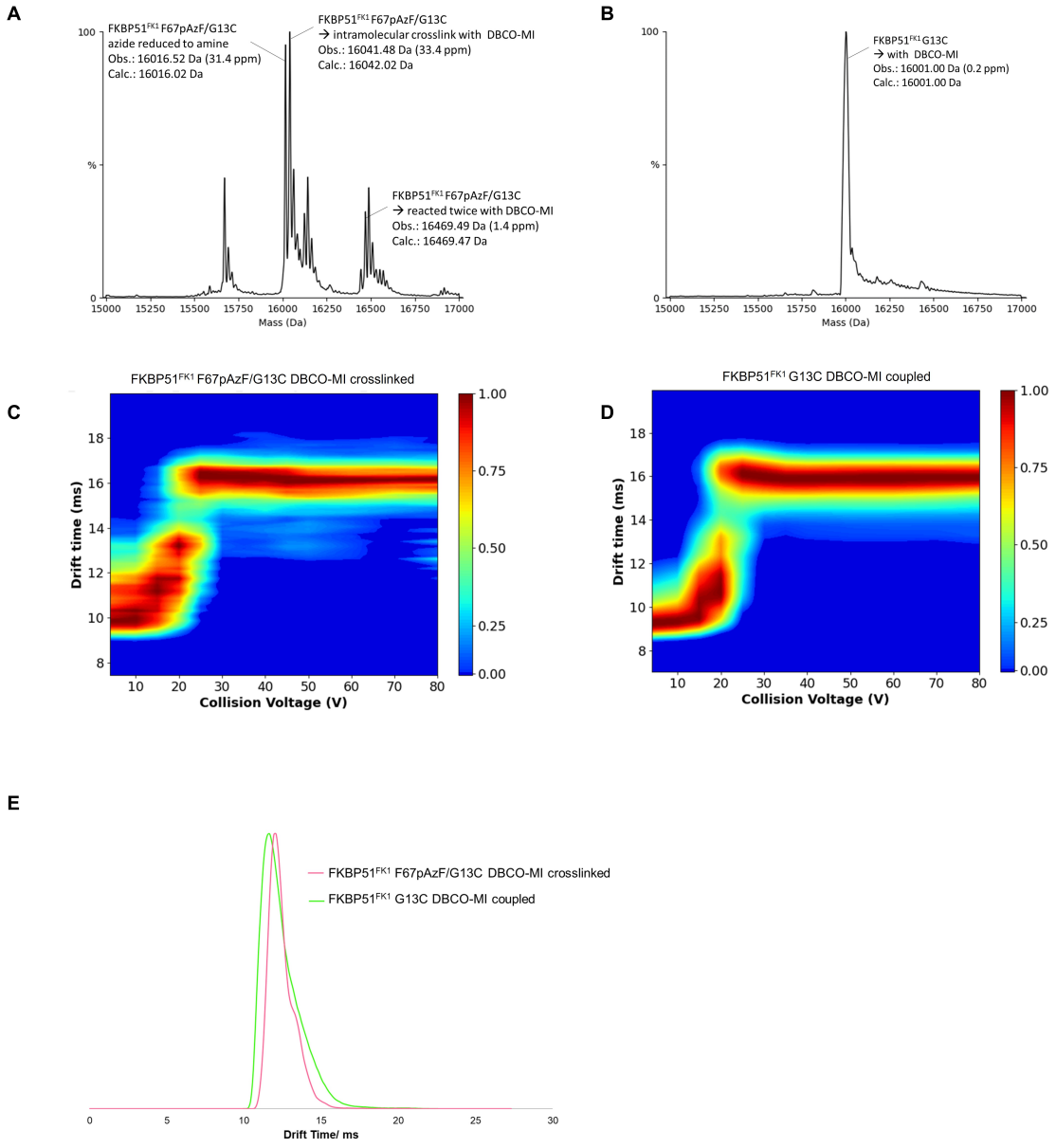
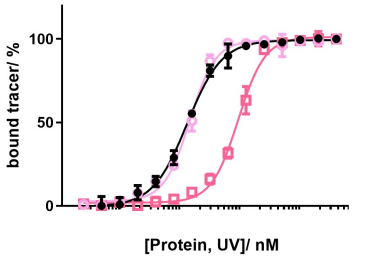


Figure 3.16: Intact MS measurements for F67pAzF/G13C crosslinked with DBCO-MI (A) and as control G13C coupled with DBCO-MI (B). In addition, CIU measurements were performed for both variants (C: CIU fingerprint of F67pAzF/G13C cross-linked with DBCO-MI and D: CIU fingerprint of G13C coupled with DBCO-MI). E: Mobilograms of the two variants from IMS measurements. The measurements were performed by Thomas Nehls.

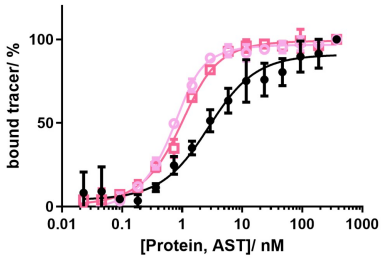
Comparing the two analyzed variants F67pAzF/G13C with DBCO-MI and G13C with DBCO-MI, a small shift in drift time can be observed (see figure 3.16E). Indeed, the more compact cross-linked F67pAzF/G13C variant exhibits a faster drift time than the DBCO-MI-coupled G13C protein variant. Interestingly, a second signal with an even faster drift time is observed for F67pAzF/G13C, indicating the heterogeneity of the sample as already established by intact MS measurements. In addition, the presence of other

species in the sample can be detected by the CIU experiment. In a CIU experiment, the anaylte is unfolded in the gas phase with increasing collision voltage, which causes a change in drift time. In the CIU fingerprints in figure 3.16C&D, the change in drift time as a function of collision voltage is shown. The warmer the color, the more intense the signal. For F67pAzF/G13C crosslinked with DBCO-MI, two species can be observed up to a collision voltage of about 20 V. In contrast, only one species can be detected for G13C coupled with DBCO-MI. The MS measurements including the intact MS, IM-MS and CIU measurements (see figure 3.16) were all performed by Thomas Nehls.

Active site titration revealed that only a part of the protein binds the fluorescein-tagged SAFit1 tracer. Comparing the K_d values of the wild-type FKBP51^{FK1} with the F67pAzF/G13C variants with and without DBCO-MI crosslinking, it is apparent that the replacement of phenylalanine in position 67 alone exhibits a similar improvement in K_d fold change as the DBCO-MI crosslinked variant.



Active site titration	FKBP51^{FK1} WT	FKBP51^{FK1} F67pAzF/G13C	FKBP51^{FK1} F67pAzF/G13C + DBCO-MI
EC_{50}/ nM	12±0.4	14±0.5	87±2.9
c(Protein, UV)/ µM	93	25	212
c(Protein, AST)/ µM	116	27	37



Binding curve	FKBP51^{FK1} WT	FKBP51^{FK1} F67pAzF/G13C	FKBP51^{FK1} F67pAzF/G13C + DBCO-MI
K_d/ nM	2.3±0.4	0.2±0.1	0.5±0.1

Figure 3.17: Active site titration (AST) and binding curves of FKBP51^{FK1} WT (black curves), FKBP51^{FK1} F67pAzF/G13C with (dark pink curves) and without (light pink curves) DBCO-MI (dibenzocyclooctinmaleimide) coupled. SAFit1-FL (fluoresceine labeled SAFit1 derivative) was the used tracer. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 30 nM for AST and 1 nM for binding curves. Protein concentration range: 230 pM – 3750 nM for AST and 23 pM - 375 nM for binding curves.

This fold change is also similar to the F67Y variant (see 3.1.1 section), leading to the assumption that the additional azide group in the phenylalanine residue might have the same structural impact on the protein structure as the additional hydroxy group in the F67Y mutant. In addition, the crosslinked sample was heterogeneous and isolation of the desired intramolecular protein variant would require additional purification steps, so a lower yield is expected. For these reasons, other options were explored, as will be shown in the next sections.

3.1.4 Stabilization by lactam bridge formation

As shown in the previous sections, the introduction of non-canonical amino acids with different chemical properties due to amber codon suppression in *E.coli* is a widespread possibility. To expand the genetic code a unique codon as well as a corresponding tRNA/aminoacyl-tRNA synthetase pair is needed.[140] But, incorporating more than one residue is still challenging. Another option is the cell-free unnatural protein synthesis which faces as well challenges such as multiple-site incorporation and low efficiency.[141] For this reason, the strategy in this section is to deploy chemical protein synthesis using the recently developed automated flow synthesizer in the research group of Prof. B. Pentelute.[142] The fully automated peptide synthesizer incorporates each residue within 40 seconds and the purity of the isolated peptide correlates with those for standard solid phase peptide synthesis.[143] Macrolactam formation is the chosen reaction type to covalently fixate the residue side chain in position 67 to the backside of the binding pocket in order to stabilize the F67-out conformation of FKBP51$^{FK1,\ 16\text{-}140}$.

Synthesis of lactam bridged FKBP51$^{FK1,\ 16\text{-}140}$ variants

To develop FKBP51 variants with a stabilized F67-out conformation, AFPS (Automated Peptide Flow Synthesis) is used to enable site-specific incorporation of unnatural amino acids and building blocks with orthogonal protecting groups. This approach enables the subsequent on-bead orthogonal deprotection and lactam formation, after the coupling of the 128 amino acids via AFPS. Four variants were designed and synthesized to investigate the influence of a lactam cyclization within the protein. The positions K58 and K60 are thereby the most promising anchor points for lactamization to position 67. Residue 67 was expected to be trapped in a conformation resembling the F67-out-like form in stabilizing the β2-β3a-loop by generating a macrocycle of seven (i, i+7) or nine (i, i+9) amino acids. In addition, for position 60, the size of the macrocycle was adjusted by incorporating smaller lysine analogs such as ornithine (Orn) or diaminobutyric acid (Dab) (see figure 5.2 in section 5.6). As a control, the corresponding wildtype FKBP51$^{FK1,\ 16\text{-}140}$ domain[45] was chemically synthesized as well.

In all constructs the native cysteines (C103A and C107I) and methionines (M48Nle and M97Nle) were abolished and replaced to protect the protein from oxidation. The synthesized variants, analyzed by LC-MS and analytical HPLC, show the expected masses and satisfying purity after preperative HPLC as demonstrated in figure 3.18.

ETD (electron transfer dissociation) as a top-down mass spectrometry fragmention technique was used to confirm the desired lactam formation at the correct site.[144] Therefore, a radical is transferred from a radical anion to the analyte, which leads to fragmentation at the backbone amide in case of proteins. The resulting fragments correspond to the c- and z-fragments according to Roepstorff, Fohlman and Biemann.[145,146] The obtained residue cleavage coverage of 47 to 66 % is in the common range for a small protein.[147] The ETD fragment maps are shown in figure 3.19. The measurements were performed by Thomas Nehls. For the synthetic FKBP51$^{FK1,\ 16\text{-}140}$ variant, large fragments containing the sequence of F67/K60 or F67/K58 were observed. This is also true for the recombinant FKBP51$^{FK1,\ 16\text{-}140}$ variant. In contrast, for the lactam-bridged variants, fragments resulting from cleavage between F67E/K58, F67E/K60Orn, and F67E/K60Dab are undetectable because the covalent bond formed protects the protein from dissociation after fragmentation at this site. F67E/K60 could not be measured due to limited sample size. However, the intact masses of all cylized variants show the expected water loss in mass, indicating

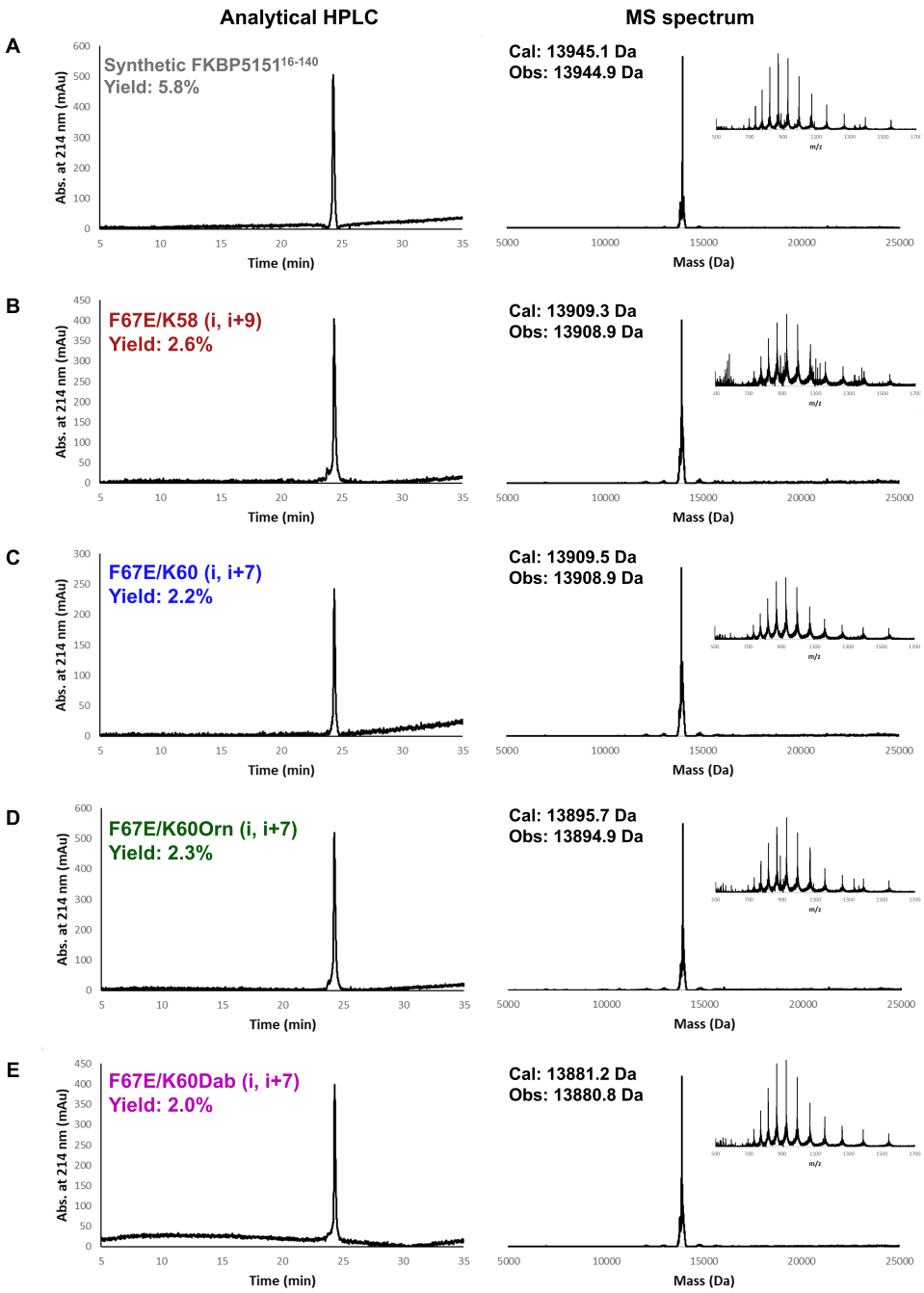


Figure 3.18: Characterization of chemically synthesized FKBP51$^{FK1,\ 16\text{-}140}$ variants after preperative HPLC. Left side: Analytical HPLC chromatograms. Right side: Deconvoluted MS data, measured by LC-MS. Inserts show the MS data prior to deconvolution. A: synthetic FKBP51$^{FK1,\ 16\text{-}140}$, B: cyclized F67E/K58 (i, i+9), C: cyclized F67E/K60 (i, i+7), D: cyclized F67E/K60Orn (i, i+7) and E: cyclized F67E/K60Dab (i, i+7).

the success of lactamization. Additionally top-down ETD measurements confirmed both the primary sequence and the modifications in the synthetic proteins.

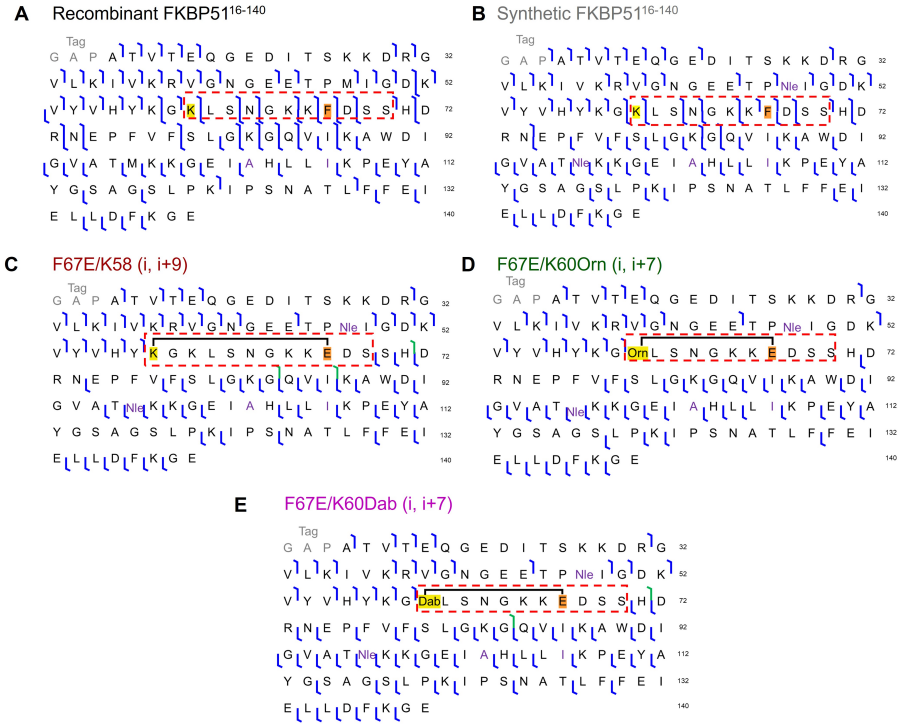


Figure 3.19: ETD fragmentation map from top-down measurements of recombinant FKBP51$^{FK1,\ 16\text{-}140}$ (A), synthetic FKBP51$^{FK1,\ 16\text{-}140}$ (B), F67E/K58 (i, i+9) (C), F67E/K60Orn (i, i+7) (D) and F67E/K60Dab (i, i+7) (E). The c- or z-fragments are indicated in blue. In green are the c-fragments with a H_2O-loss. The FKBP51 sequence starts with the alanine residue in position 16. The purple amino acids indicate the modifications C103A and C107I as well as the exchange of methionine in norleucin (Nle) in positions 48 and 97. Position 60 and 58 are highlighted in yellow, position 67 in orange and the black line is indicating the formed lactam-brigde. The numbers on the right side correspond to the FKBP51$^{FK1,\ 16\text{-}140}$ sequence.

Refolding of synthesized FKBP51$^{FK1,\ 16\text{-}140}$ variants

In the next step, the obtained lyophilized proteins needed to be refolded to obtain proteins with native 3D conformation and biological activity in solution. The refolding method used was the rapid dilution method. The refolding buffer has a great influence on the refolding success and a high number of variations can be tested. Four buffers were chosen from the Quickfold Refolding Kit with slight modifications, resulting in the following compositions of the to be tested refolding buffers:

Refolding Buffer 1

50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5 M arginine, 0.4 M sucrose, 0.75 M guanidine HCl

Refolding Buffer 2

50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 1 mM EDTA, 0.5 % Triton X-100

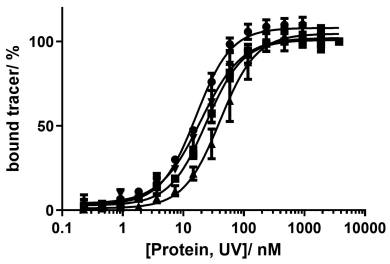
Refolding Buffer 3

50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5 M arginine, 0.4 M sucrose, 0.75 M guanidine HCl, 0.5 % Triton X-100, 0.05 % polyethylene glycol 3,550

Refolding Buffer 4

50 mM Tris-Cl pH 8.5, 240 mM NaCl, 10 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.4 M sucrose

For the refolding tests, the synthetic FKBP51$^{FK1,\ 16\text{-}140}$ was used according to the refolding protocol described in chapter 5.6. Via active site titration[148] the refolding yield was estimated, showing refolding buffer 1 to be the optimal refolding buffer (see figure 3.20).



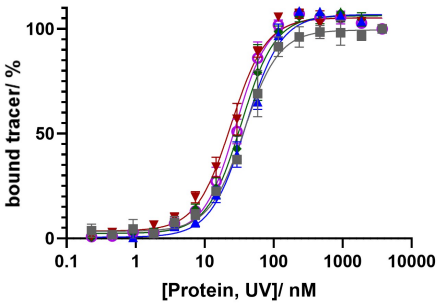
Active site titration	Refolding Buffer 1	Refolding Buffer 1	Refolding Buffer 3	Refolding Buffer 4
EC_{50}/ nM	16±1	24±1	40±3	20±1
c(Protein, UV)/ µM	65	107	129	93
c(Protein, AST)/ µM	60	67	48	70
Refolding yield/ %	92	62	37	75

Figure 3.20: Active site titration (AST) of synthetic FKBP51$^{FK1,\ 16\text{-}140}$ (black curves) under four different refolding conditions. The used tracer is SAFit1-FL. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 30 nM. Protein concentration range: 230 pM – 3750 nM.

The most important step in refolding was the slow addition of the dissolved protein solution (in denaturing buffer containing 6 M guanidine HCl) with stirring to avoid local concentration peaks. However, when the established refolding protocol was adapted for all synthetic variants, the refolding yield decreased for the synthetic FKBP51$^{FK1,\ 16\text{-}140}$ variant. The refolding efficiency also appears to be dependent on the variant, ranging from 22-62 % in this experiment (see figure 3.21). This suggests that the refolding protocol is susceptible to minor changes, such as using a new buffer batch, temperature variations, and dilution rate.

The obtained protein solutions were further investigated via conformation-sensitive MS with the help of my cooperation partner Thomas Nehls. He conducted a size exclusion chromatography electron spray mass spectrometry (SEC-ESI-MS) experiment. SEC analysis enables the separation of different conformations or different proteins, so that the influence of native and denatured conditions on the recombinant and refolded protein variants can be studied. All proteins show a significant change in elution volume when compared between denatured and native conditions (see figure 3.22A). Under denatured conditions, a protein has a higher hydrodynamic radius and therefore elutes with a smaller volume than in its native, folded form. Since the elution volume of the refolded, synthetic proteins is the same as for the recombinant protein, we can assume that the conformation corresponds to the conformation of the native, correctly folded protein. However, the signal for the native samples is broad and exhibits several peaks for same variants (e.g. F67E/K58 (i, i+9), red curve).

Another technique to confirm the 3D structure of the refolded proteins, is ion mobility spectrometry.[133] The analytes in the gas phase pass through an electric field and a



Active site titration	synthetic FKBP51$^{FK1,16-140}$	F67E/K58 (i, i+9)	F67E/K60 (i, i+7)	F67E/K60Orn (i, i+7)	F67E/K60Dab (i, i+7)
EC_{50}/ nM	67±4	39±2	24±1	34±2	28±1
c(Protein, UV)/ µM	143	44	47	39	57
c(Protein, AST)/ µM	32	16	29	17	31
Refolding yield/ %	22	36	62	44	54

Figure 3.21: Active site titration (AST) of synthetic FKBP51$^{FK1, 16-140}$ (gray curve), F67E/K58 (i, i+9) (red curve), F67E/K60 (i, i+7) (blue curve), F67E/K60Orn (i, i+7) (green curve) and F67E/K60Dab (i, i+7) (magenta curve) after refolding with Refolding Buffer 1 (50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5 M arginine, 0.4 M sucrose, 0.75 M guanidine HCl.). The used tracer is SAFit1-FL. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 30 nM. Protein concentration range: 230 pM – 3750 nM.

controlled gas flow. Thus, due to soft collisions, the gas molecules exhibit a resistance that depends on the cross-sectional area of the analyte. Since the analyte in the gas phase rotates, it is only possible to determine the average cross-sectional area, the so called collision cross-section (CCS). The proteins studied were measured in the IMS experiment under native conditions by Thomas Nehls. All proteins under native conditions exhibit CCS values in the same range, indicating a similar conformation of the refolded synthesized proteins compared to the recombinantly produced protein, as shown in the mobilograms in figure 3.22B.

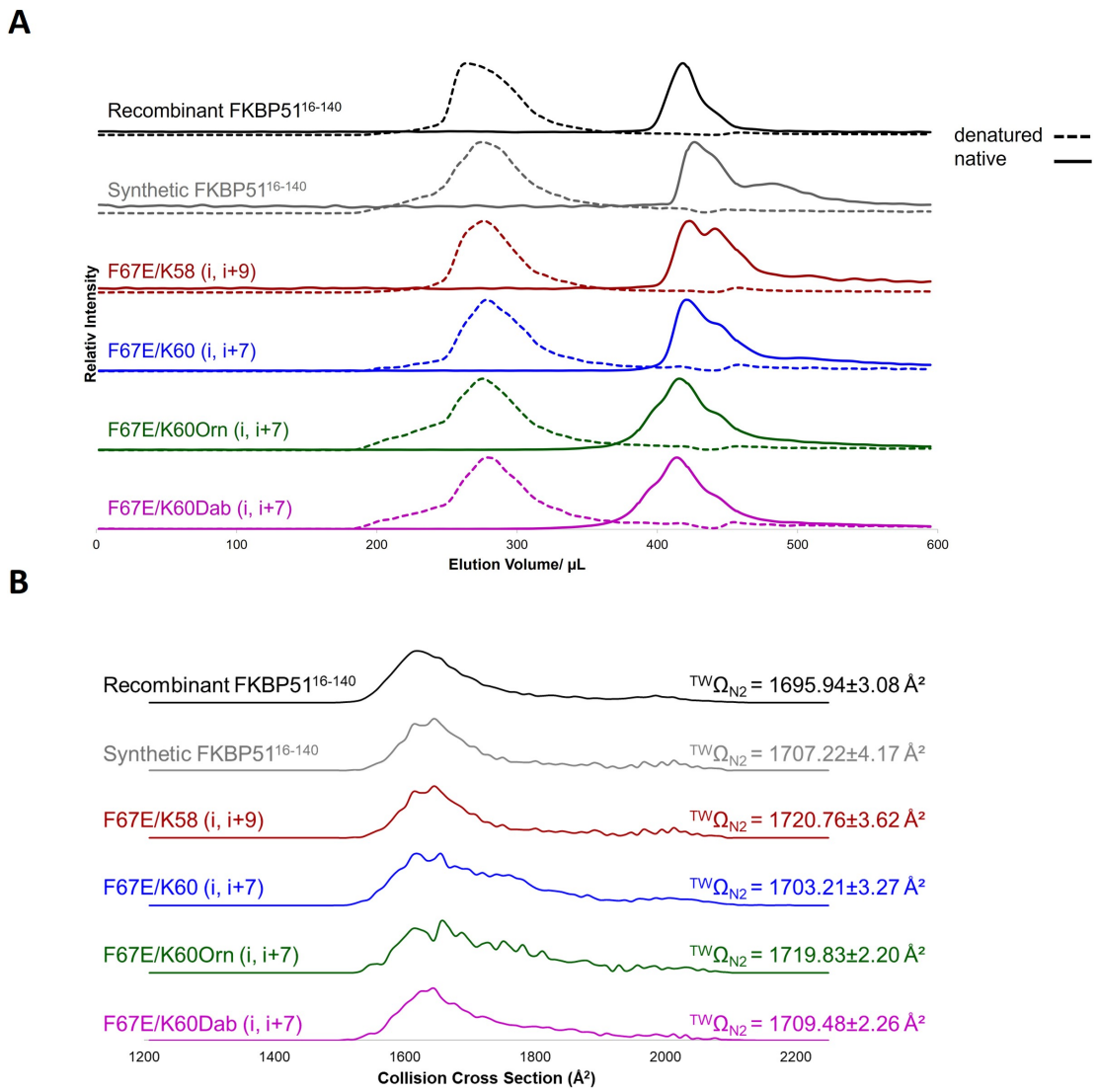


Figure 3.22: A: Extracted ion chromatograms (native: z=8+, denatured: z=16+) from SEC-ESI-MS measurements of lactam-bridged variants in comparison to the synthetic and recombinant FKBP51$^{FK1,\ 16\text{-}140}$ variants. Dotted lines indicate denaturing conditions (0.2 % formic acid in water) and continous lines nativ conditions (50 mM ammonium acetate buffer pH 7) during the run. B: Mobilograms at charge state 8+ of lactam-bridged variants as well as synthetic and recombinant FKBP51$^{FK1,\ 16\text{-}140}$ variants. The top values of the collision cross section values (CCS) for charge state 8+ are displayed on the right. Measured with traveling wave with N_2 as collision gas.

Protein crystallography confirmed the 3D structure of the synthesized and refolded F67E/K58 (i, i+9) and F67E/K60Orn (i, i+7) proteins. The structures of the stabilized F67-out-like conformation can be accurately overlaid with the wild-type FKBP51$^{FK1,\ 16\text{-}140}$ protein, as shown in figure 3.23, verifying the correct folding of the synthesized proteins. The binding mode of ligand SAFit1 is also identical. The electron density maps of the lactam bridges additionally confirm the success of the cyclization in the protein.

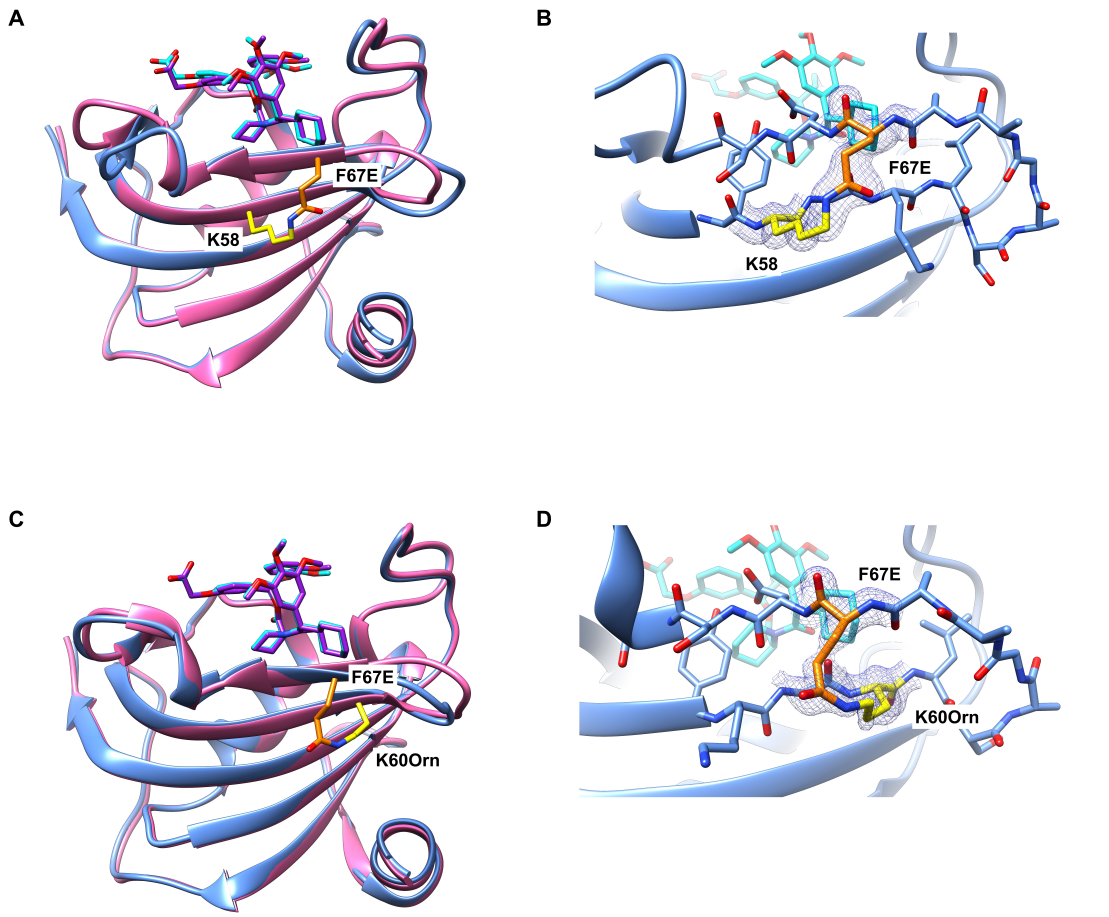
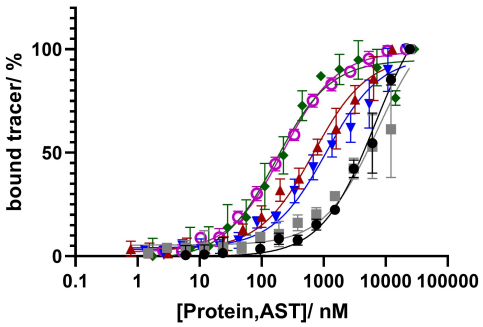


Figure 3.23: A, C: Overlay of the F67-out-like conformation in blue (A: FKBP51$^{FK1,\ 16\text{-}140}$ F67E/K58 (i, i+9) PDB: 8PJA and B: FKBP51$^{FK1,\ 16\text{-}140}$ F67E/K60Orn (i, i+7) PDB:8PJ8) with the wildtype FKBP51$^{FK1,\ 16\text{-}140}$ pink structure (PDB: 8CCA) in complex with SAFit1 (cyan sticks for PDB:8PJA&8PJ8 and purple sticks for PDB: 8CCA). B, D: View on the backside of the binding pocket with the electron density map for the lactam bridge. The structures were generated by Dr. Christian Meyners.

Fluorescence polarization assays with synthesized FKBP51$^{FK1,\ 16\text{-}140}$ variants

Lactamization within the protein leads to a shift in the conformational rearrangement equilibrium towards the F67-out-like conformation. This theoretically results in a higher binding affinity of conformation-specific ligands, as the conversion from F67-in to F67-out is omitted. Indeed, when comparing the K_d values of recombinant FKBP51$^{FK1,\ 16\text{-}140}$ with the synthesized cyclized variants, a 6-10-fold stronger binding of the low affinity tracer for lysine/glutamic acid cyclization is observed. For shorter lysine derivatives, an up to 39-fold improvement is achieved when lysine is replaced by ornithine.

Four stabilized F67-out-like variants were successfully synthesized, cyclized, and re-folded. Thus, the residue in position 67 was trapped in an F67-out-like conformation by a macrocycle of seven (i, i+7) or nine (i, i+9) amino acids. The stabilized variants showed a 6-39-fold improvement in K_d for the low binding affinity tracer, with F67E/K60Dab (i,



Binding curve	recombinant FKBP51$^{FK1,16-140}$	synthetic FKBP51$^{FK1,16-140}$	F67E/K58 (i, i+9)	F67E/K60 (i, i+7)	F67E/K60Orn (i, i+7)	F67E/K60Dab (i, i+7)K
K_d/ nM	6741±874	7443±1850	685±88	1103±151	172±27	220±12
Fold change in K_d	1	0.9	10	6	39	31

Figure 3.24: Binding curves of recombinant FKBP51$^{FK1, 16-140}$ (black curve), synthetic FKBP51$^{FK1, 16-140}$ (gray curve) and the syntehtic, lactam-bridged variants with a low affinity tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 1 nM. Protein concentration range: 778 nM - 288750 nM for binding curves.

i+7) showing the highest improvement. Although this theoretically makes the variant a suitable candidate for screening fragment libraries for new ligand scaffolds, due to limited access to the automated flow peptide synthesizer in the Pentelute Lab, new variants were still sought. Another disadvantage is the high effort required to produce the variants and the small sample size.

3.1.5 Stabilization by cysteine-directed crosslinking

The final approach to stabilize the F67-out conformation of FKBP51^{FK1} was cysteine-directed crosslinking. The advantage of this approach is the application of the rapid and simple method of site-directed mutagenesis. The SPRINP (single primer in parallel) method introduces site-specific cysteine mutations into the protein, as described in more detail in section 5.1. Therefore, double cysteine variants (referred to here as doubleCys) were generated and purified to site-specifically crosslink the cysteine at position 67 to the 5 different selected anchor sites 13, 58, 60, 65 and 69. Figure 3.25A shows the structure of FKBP51^{FK1}, with the sites to be crosslinked highlighted in gray, and the SDS-PAGE showing the successful production and purification of the 5 doubleCys variants. The maleimide-based homobifunctional crosslinker BMB (1,4-bismaleimidobutane) was used for long-distance crosslinking in F67C/G13C, whereas the short-distance crosslinking was performed by either DBA (dibromoacetone) crosslinking or simple disulfide bridging. The chemical structures of the homobifunctional linkers are displayed in figure 3.25B.

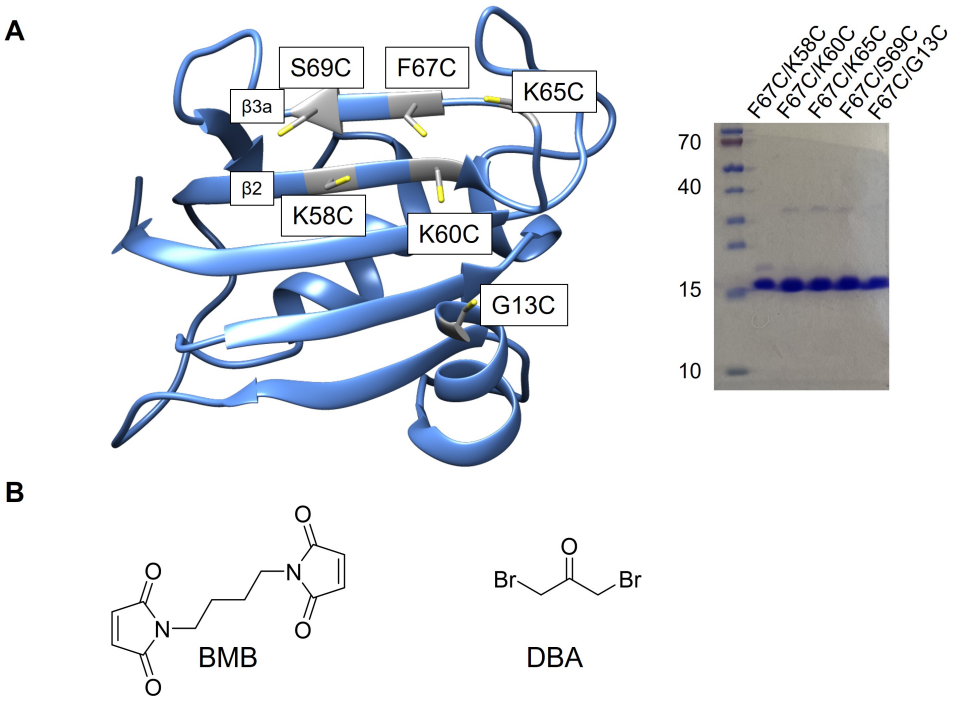


Figure 3.25: SDS-PAGE of the doubleCys variants used for intramolecular crosslinking in this chapter. B: chemical structure of the used crosslinker BMB and DBA.

It is noticeable that the use of a homobifunctional linker can lead not only to the desired intramolecular crosslinking product, but also to by-products including dimerized protein, coupling of the crosslinker with only one free thiol group, coupling of two crosslinker molecules once with each thiol group in the doubleCys variants, or no reaction of the protein at all, leaving the protein unmodified (see figure 3.26).

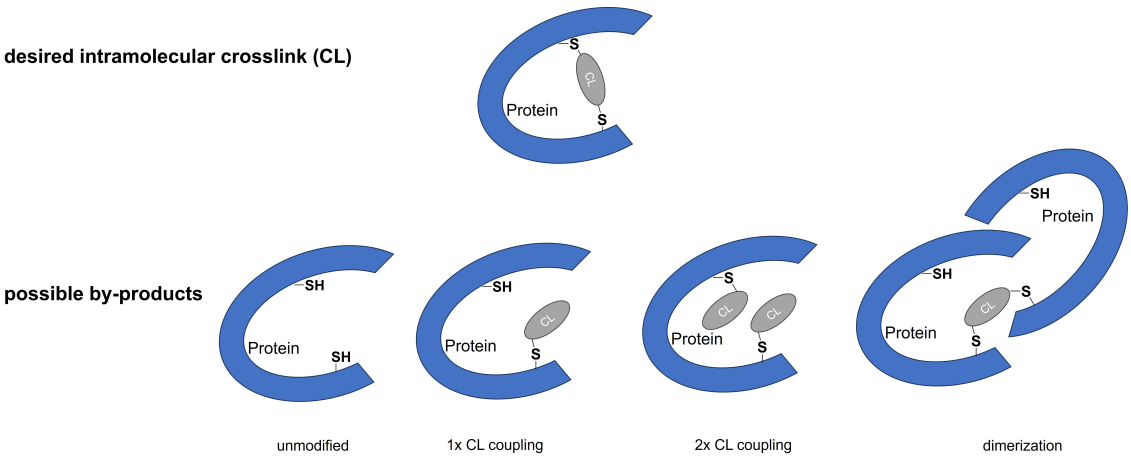


Figure 3.26: Schematic represantation of the desired intramolecular crosslinking product and examples of possible by-products during crosslinker (CL) coupling.

Crosslinking by 1,4-bismaleimidobutane (BMB)

Crosslinking of F67C/G13C with BMB was performed as described in 5.5.5 with the corresponding control proteins. noCys is the negative control because the protein variant does not have free thiol groups and therefore would not react with the crosslinker and MI-PEG, which is used as an indirect detection tool for free thiol groups. Indeed, the SDS-PAGE gel in figure 3.27A shows no shift in the protein band after MI-PEGylation. E140C, on the other hand, has an easily accessible thiol group at the C-terminus and serves as a positive control. MI-PEGylation of the protein results in an almost quantitative shift of the protein band, confirming the accessibility and reactivity of the thiol group under these conditions. However, BMB crosslinking and subsequent MI-PEGylation show that only a small fraction reacted with the crosslinker and another fraction gave dimerized protein, as indicated by the protein bands at about 40-55 kDa. The same is true for the monoCys variant F67C, but with a less intense dimerization band, which can be explained by the position of the cysteine group (C-terminus vs. within the protein). Interestingly, the doubleCys variant, which is supposed to produce intramolecular crosslinking to stabilize the F67-out conformation, shows a second protein band that appears just below the protein band around 25 kDa after BMB treatment. The intensity of the lower protein band did not change after subsequent MI-PEGylation, suggesting that it may be the actual desired intramolecular cross-linked protein product that is not accessible during MI-PEGylation because the thiol groups form a thiol ether with the crosslinker. In SDS-PAGE, the proteins move through a polyacrylamineinde gel denatured depending on their hydrodynamic radius. The intramolecular crosslinked protein theoretically has a smaller hydrodynamic radius due to the covalent linkage between two residues within the protein, resulting in a loop structure, compared to the unmodified protein, which explains the different drift times in the gel. The lower formation of dimers is also noteworthy. The intact MS measurements (see figure 3.27B), performed by Thomas Nehls, confirm the results of the SDS-PAGE characterization: after BMB crosslinking, a peak with the corresponding mass appears. However, neither intact MS measurements nor SDS-PAGE can unequivocally verify whether the crosslinker reacted with the protein only once or whether intramolecular crosslinking actually occurred. The reason for this is that maleimides react with thiols in an addition reaction (Micheal addition), which means that the mass does not change as a

result of the reaction.

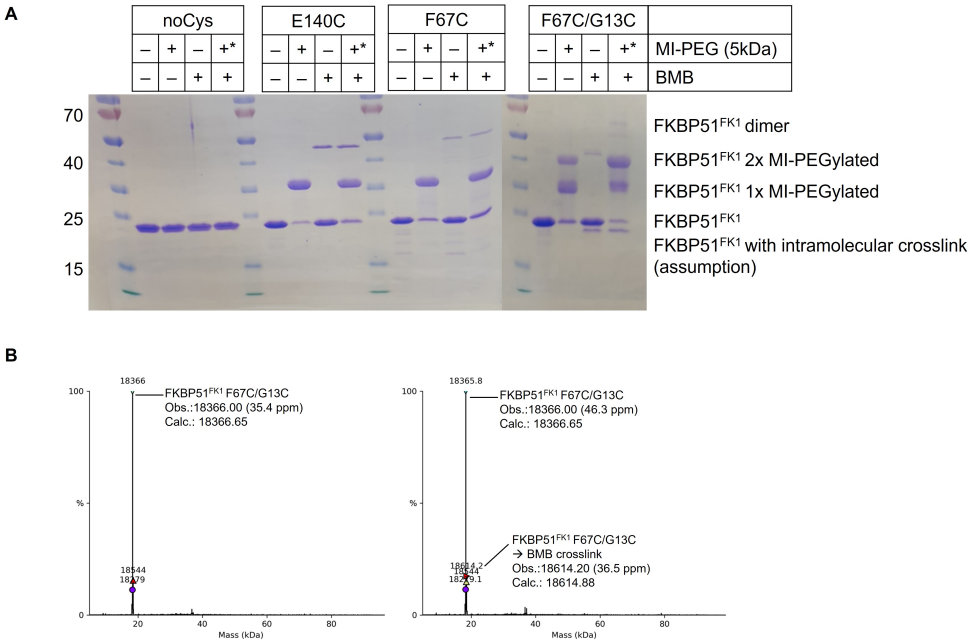


Figure 3.27: Crosslinking with BMB. A: SDS-PAGE of BMB crosslinking with F67C/G13C and the corresponding controls. * indicates, that MI-PEGylation was performed after BMB crosslinking. B: intact MS measurements before (left) and after (right) BMB crosslinking.

To investigate the protein band appearing after BMB treatment, the coupling was repeated with a 5× higher protein concentration (see figure 3.28A). In addition, MI-PEGylation was performed with a 10 kDa PEG chain to increase the change in molecular weight after MI-PEGylation to separate the crosslinked protein from the non-crosslinked protein. Figure 3.28B shows the corresponding SEC chromatogram and the SDS-PAGE showing the fractions analyzed. SEC successfully separated the MI-PEGylated protein from the BMB-crosslinked products, but the double band resulting from BMB treatment is not resolved. Fraction 4 was then analyzed by intact MS measurements (performed by Thomas Nehls), which showed the mass of the doubleCys variant F67C/G13C coupled to BMB and the corresponding sodium adduct (see figure 3.28C). However, as mentioned above, intact MS measurements are not suited to confirm intramolecular crosslinking. Therefore, the fraction was reloaded onto an SDS-PAGE gel and both bands were cut out separately to perform in-gel digestion. The samples were then analyzed by Thomas Nehls. Unfortunately, no peptides to a sufficient extent were detected, likely due to low efficiency of the in-gel digestion.

Crosslinking by dibromoacetone (DBA)

Since intramolecular BMB crosslinking could not be confirmed by MS measurements, DBA was used as an alternative crosslinker. Here, it is possible to analzye an intramolecular crosslinking reaction by intact MS measurements, since a substitution reaction occurs during the reaction with DBA, resulting in a molecular weight loss of 2 Br (159.808 Da). First, the double Cysvariants F67C/K60C, F67C/K65C, and F67C/S69C were tested, and

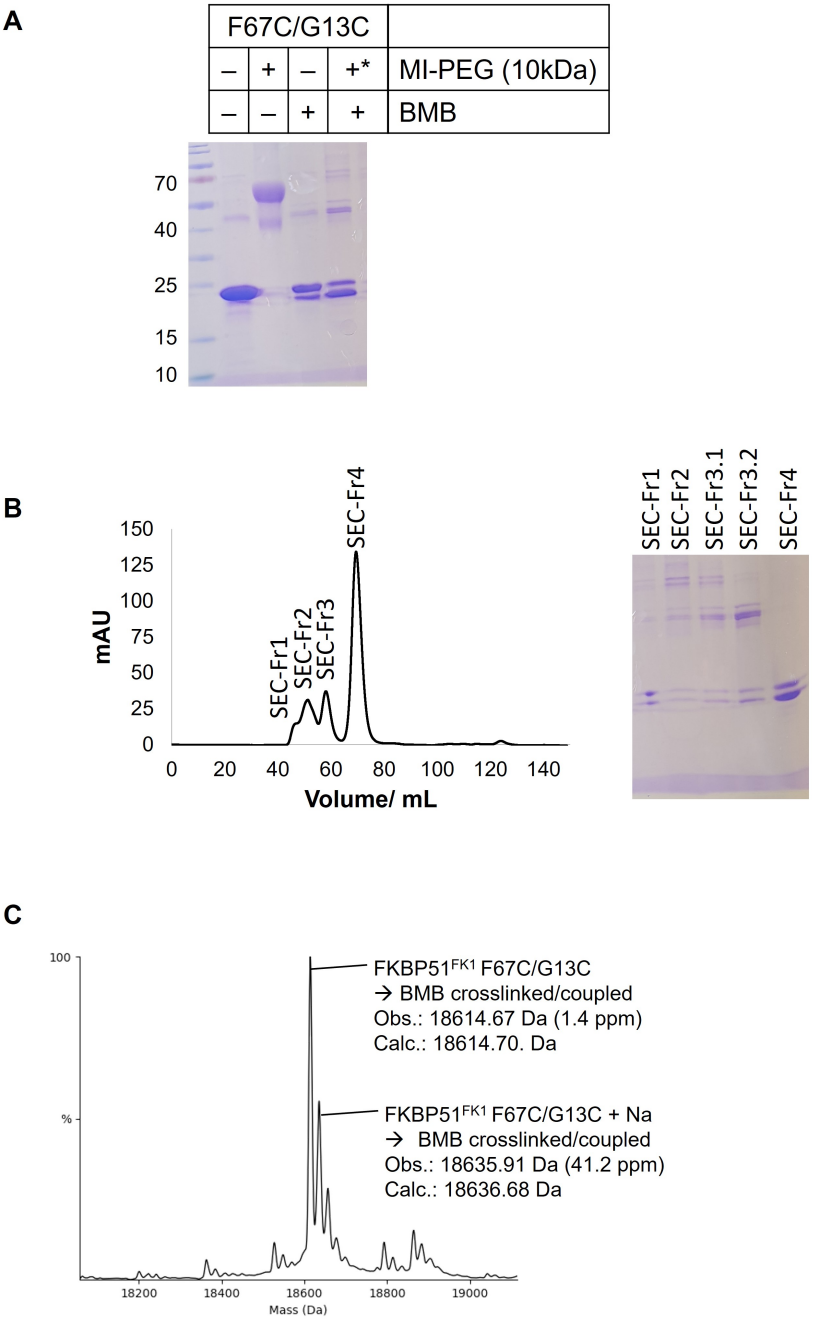


Figure 3.28: Scaled-up batch of F67C/G13C coupled with BMB. A: SDS-PAGE of BMB crosslinking with F67C/G13C and the corresponding MI-PEGylation samples. B: chromatogram of SEC run of the protein crosslinked with BMB and subsequently MI-PEGylated. C: deconvoluted MS data of the purified sample.

MI-PEG (10 kDa) was again used as a tool to detect free thiol groups. The SDS-PAGE gel in figure 3.29 shows that after DBA crosslinking, subsequent MI-PEGylation results in only a small shift of the band in the gel, indicating a successful reaction. This is true for all three doubleCys variants tested.

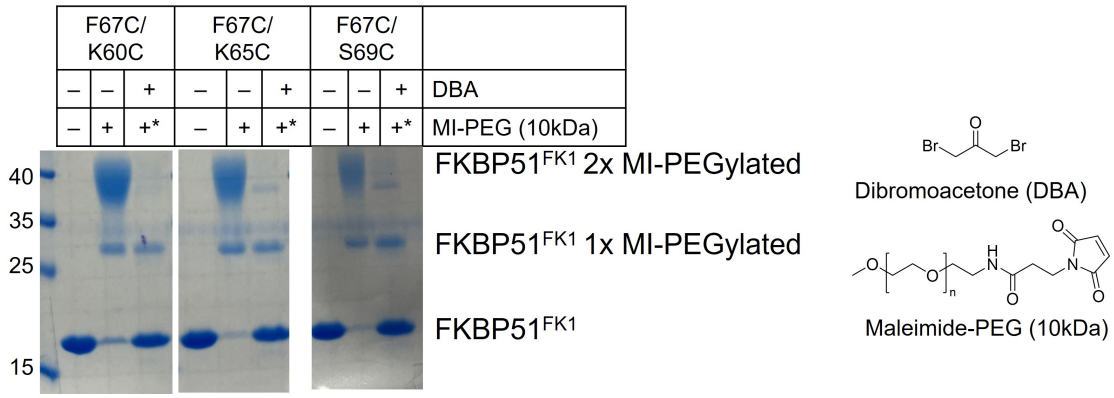


Figure 3.29: SDS-PAGE of the crosslinking of doublCys variants F67C/K60C, F67C/K65C and F67C/S69C with DBA. For each variant applies the following: first lane shows the unmodified protein, the second lane the MI-PEGylation of the protein variant and the last lane shows the MI-PEGylation after DBA treatment, which is indicated by the symbol *. On the right side, the chemical structures of DBA and MI-PEG are displayed.

To obtain a homogeneous DBA-crosslinked protein product, MI-PEGylation was used to increase the molecular weight of protein species with free thiol groups after DBA treatment and then separate them by SEC. The corresponding SEC chromatograms for all three variants and the SDS-PAGE gels of the obtained fractions are shown in figure 3.30. The chromatograms of the three doubleCys variants tested show a very similar pattern. Comparing the loading (first lane in the gels, see figure 3.30) with the fractions obtained by SEC, it is clear that the separation of the higher molecular weight proteins was successful. However, the loss of product seems to be unavoidable, since at least a small amount of protein with a molecular weight of about 15 kDa is found in all fractions. Fraction 4 with an elution volume of approximately 15 mL showed a pure protein band and was further analyzed by Thomas Nehls using MS.

The MS measurements in figure 3.31 confirm that the samples contain the desired intramolecular product. However, it is clear that most of the protein is present as a disulfide, so that the thiol groups cannot be modified by either DBA or MI-PEG. This is particularly true for the F67C/K60C and F67C/S69C variants, with the most intense signal being the disulfide-bridged variants. In contrast, for the F67C/K65C variant, the most intense signal is the potassium adduct of the intramolecularly crosslinked protein, making this variant the most successful under the conditions tested. However, signals are present in all variants that cannot be assigned because they are most likely contaminated.

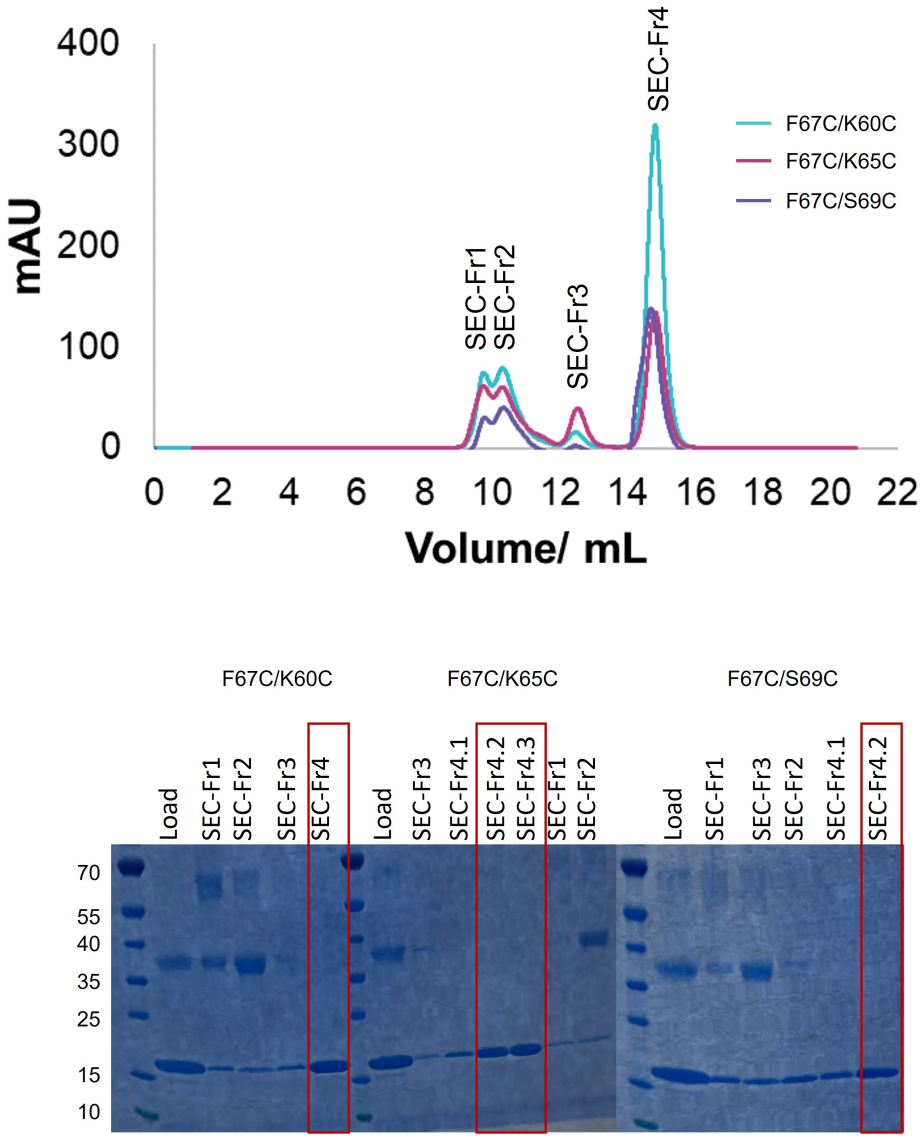


Figure 3.30: SEC chromatograms of crosslinked doubleCys variants after DBA crosslinking and MI-PEGylation. The obtained fraction were analyzed by SDS-PAGE. The red boxes highlight the recovered fractions for the next steps.

Since the disulfides seemed to form naturally and without external factors, I decided to investigate the disulfide variants in more detail as well. This will be the topic of the next chapter. Nevertheless, it was interesting to form a covalent bond within the protein by crosslinkers, since the size of the crosslink and thus the flexibility can be fine-tuned. Hence, the crosslinking conditions had to be optimized to avoid the formation of disulfides. The 18 conditions tested are summarized in table 3.3. TCEP was used as the reducing agent because, unlike the commonly used reducing agents DTT (dithiothreitol) or β-ME (β-mercaptoethanol), TCEP has no free thiol groups. Concentrations ranging from 40 µM to 500 µM were tested, corresponding to a 2- to 25-fold excess compared to the F67C/K60C protein (20 µM). The concentration of the crosslinker was also varied and ranged from 500-2000 µM. Additionally, the influence of the buffer substance was tested (HEPES vs.

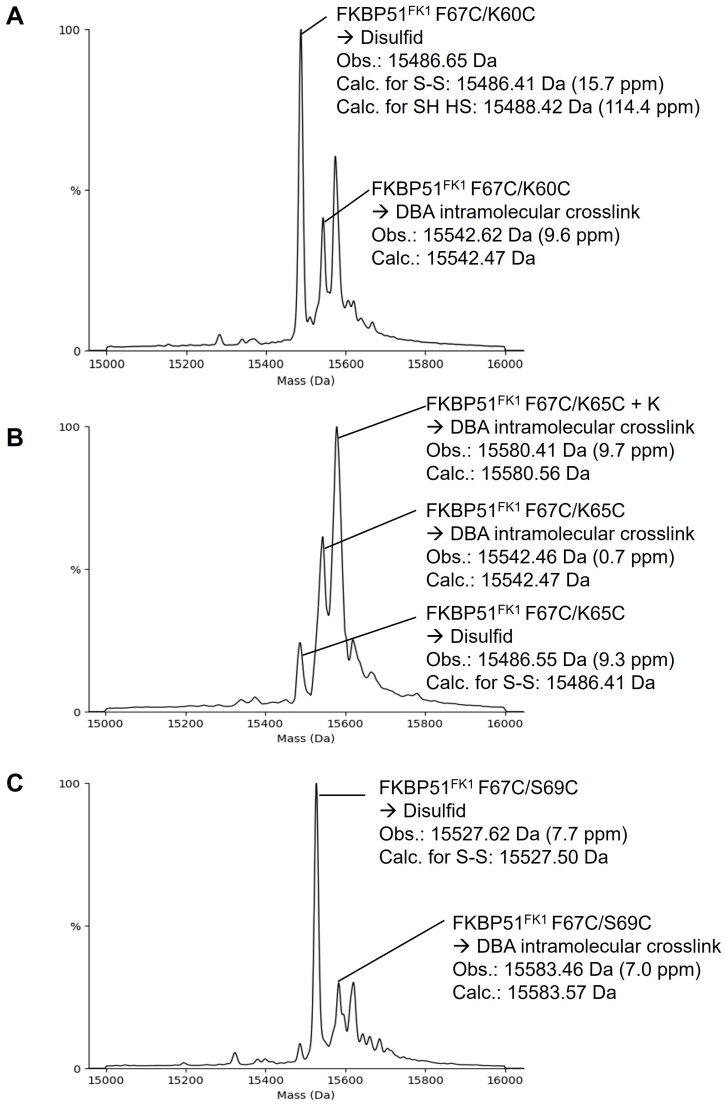


Figure 3.31: Deconvoluted MS data of DBA crosslinked and SEC purified doubleCys variants F67C/K60C (A), F67C/K65C (B) and F67C/S69C (C).

Tris). The samples were all analyzed by MS to determine the relative intensities of the crosslinked compared to the unmodified protein, as shown summarized in the table 3.3. The MS measurements were conducted by Thomas Nehls.

The best results were achieved with condition 9 and 18, showing that a high excess of TCEP and DBA favors the intramolecular crosslinking. Even though TCEP has a number of advantages compared to the thiol containing reducing agents (e.g. DTT) including being a crystaline and odorless solid. TCEP is also non-volatile and water-soluble, but drawbacks are the capability of TCEP to react with typical reagents used in cysteine-directed bioconjugation such as maleimides and iodoacetamide.[149] Here, an excess of both reagents was tested to overcome this hurdle. Indeed, increasing the concentration of TCEP in Tris buffer at constant DBA concentration (e.g. conditions 13-15) leads to an increase in the relative intensity of the desired product in MS measurements. When the DBA concentration is increased from 1000 µM to 2000 µM, the relative intensities of the

product increase even further, as shown for conditions 16-18.

Table 3.3: DBA crosslinking of doubleCys variant F67C/K60C under various conditions. The samples were analyzed by MS comparing the relative intensities of the unmodified species and the desired, intramolecular crosslinked species.

#	Protein	TCEP	DBA	Buffer	Rel. Intensity unmodified	Rel. Intensity intramolecular CL
1	20 µM	40 µM	500 µM	HEPES	100	47
2	20 µM	100 µM	500 µM	HEPES	100	59
3	20 µM	500 µM	500 µM	HEPES	100	30
4	20 µM	40 µM	1000 µM	HEPES	100	45
5	20 µM	100 µM	1000 µM	HEPES	100	70
6	20 µM	500 µM	1000 µM	HEPES	100	32
7	20 µM	40 µM	2000 µM	HEPES	100	25
8	20 µM	100 µM	2000 µM	HEPES	100	20
9	20 µM	500 µM	2000 µM	HEPES	100	94
10	20 µM	40 µM	500 µM	Tris	100	41
11	20 µM	100 µM	500 µM	Tris	100	43
12	20 µM	500 µM	500 µM	Tris	100	28
13	20 µM	40 µM	1000 µM	Tris	100	50
14	20 µM	100 µM	1000 µM	Tris	100	71
15	20 µM	500 µM	1000 µM	Tris	100	80
16	20 µM	40 µM	2000 µM	Tris	100	50
17	20 µM	100 µM	2000 µM	Tris	100	60
18	20 µM	500 µM	2000 µM	Tris	67	100

Figure 3.32 shows the MS data for conditions 5, 9, 14, and 18, visually demonstrating the intensification of the desired product signal at high TCEP and DBA concentrations (conditions 9 and 18). When comparing conditions 9 and 18, it seems as if Tris is the more suitable buffer substance. Finally, condition 18 was used in a scaled-up experiment with F67C/K60C and F67C/K58C as proteins for DBA crosslinking.

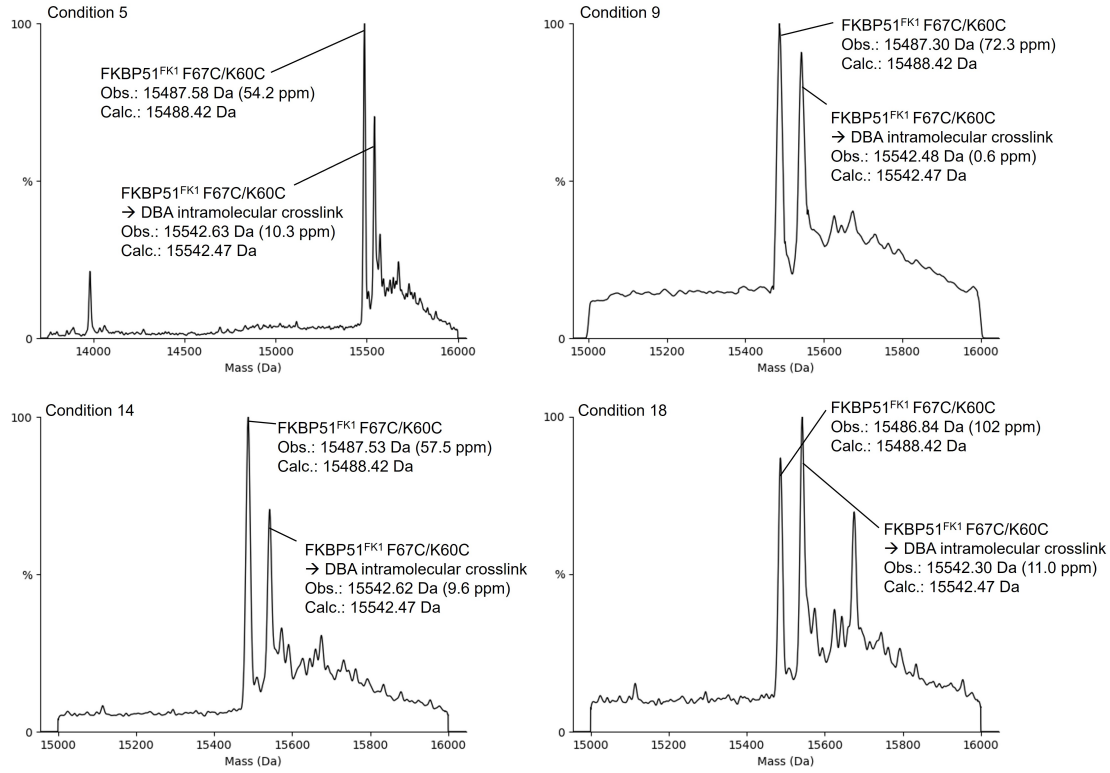


Figure 3.32: Deconvoluted MS data of DBA crosslinked doubleCys variant F67C/K60C under various conditions.

After each step, the samples are further characterized by MS measurements by Thomas Nehls. The workflow is shown in figure 3.33A and the corresponding MS data of each sample are shown in 3.33B. The MS data after DBA crosslinking reveal that the desired intramolecular crosslinking has formed for both proteins (F67C/K58C and F67C/K60C) and even proves to be the most intense signal for F67C/K60C. In the case of F67C/K58C, the most intensive signal has an observed mass of 15574.05, which can not be assigned to a conceivable by-product, leaving the identification unclear. In addition, disulfide-bridged variants are still present in both protein samples. After DBA crosslinking, samples were incubated with immobilized maleimide (MI-resin), which resulted in intensification of the desired intramolecular crosslinking product, whereas incubation with SH-resin showed no significant improvement in purity. Samples were then tested in an FP assay, determining the K_d value of the low affinity tracer compared to the WT and the F67Y variant of $FKBP51^{FK1}$. In addition, the unmodified and reduced doubleCys variants were also tested, and these variants were found to have a similar K_d value to the WT variant. In contrast, the crosslinked DBA variants showed an 11-fold (F67C/K60C, DBA CL comparable to F67Y) and 39-fold (F67C/K58C, DBA CL) improvement in binding affinity, respectively (see figure 3.34).

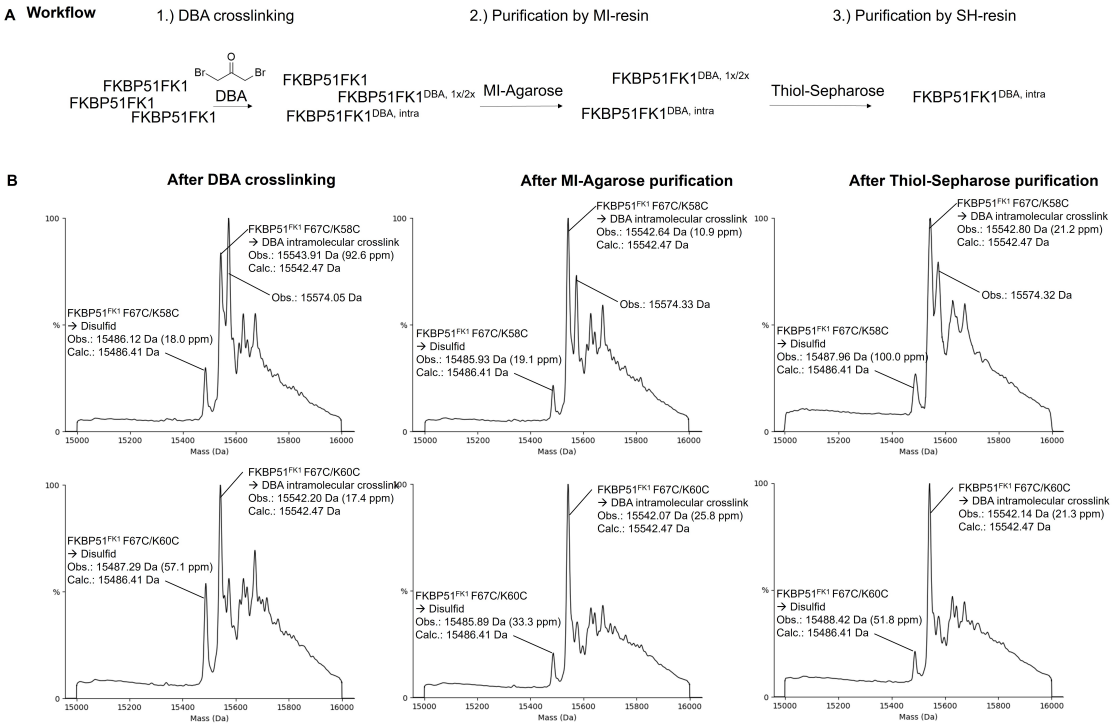
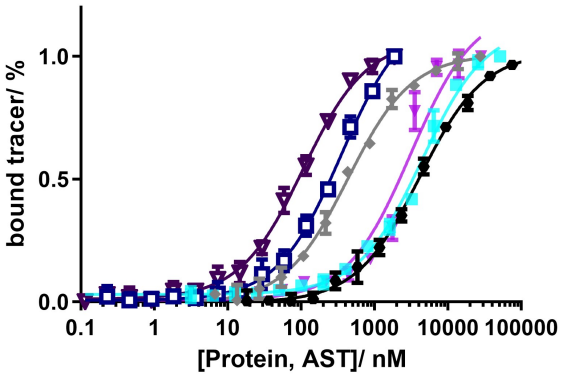


Figure 3.33: A: workflow of the DBA crosslinking of F67C/K58C and F67C/K60C and subsequent purification by MI-resin, and followed by SH-resin. B: Deconvoluted MS data of F67C/K58C or F67C/K60C after DBA crosslinking (left), after MI-resin purification (middle) and SH-resin purification (right).

This high improvement in binding affinity makes these variants and especially F67C/K58, DBA CL, a suitable candidate as a screening tool in the search for new ligand scaffolds for FKBP51. But the downside of these variants is the multistep workflow, which leads to an inefficient process. Moreover, the samples are still heterogeneous after the purification steps, which makes protein crystallization difficult but is an important tool for the subsequent characterization of new ligand scaffolds.



Binding curve	$FKBP51^{FK1}$ WT	$FKBP51^{FK1}$ F67Y	$FKBP51^{FK1}$ F67C/K60C	$FKBP51^{FK1}$ F67C/K60C, DBA CL	$FKBP51^{FK1}$ F67C/K58C	$FKBP51^{FK1}$ F67C/K58C, DBA CL
K_d/ nM	4122±209	433±23	4437±303	350±26	3107±439	106±6

Figure 3.34: Binding curves of $FKBP51^{FK1}$ (black curve), F67Y (gray curve) and the doubleCys variants in their reduced form as well as crosslinekd with DBA, with a low affinity tracer labeled with fluorescein. Each data point is indicated as the mean of three technical replicats. Final concentration of the tracer: 1 nM. Protein concentration range: 114 pM - 150000 nM for binding curves.

Crosslinking by disulfid bridging

As mentioned above, the disulfide bridge forms naturally and even after DBA crosslinking with the reducing agent TCEP present, a part of the protein is found with a disulfide bridge. Thus, doubleCys variants F67C/K58C and F67C/K60C were now purified without addition of a reducing agent and additionally incubated overnight at room temperature after purification. The samples were tested in MS measurements with and without TCEP. The MS data is shown in figure 3.35A. The MS measurements were performed by Thomas Nehls. The protein samples show high purity and quantitative yield of disulfide-bridged protein without TCEP, which is true for both protein variants. As expected, the addition of TCEP leads to a partial reduction of the disulfide bridges. As further proof for the formation of a disulfide bridge in the proteins, ETD measurements were performed by Thomas Nehls. The ETD fragment maps are shown in figure 3.35B. In the samples with TCEP, large fragments reflecting breaks between F67C/K58C or F67C/K60C were observed. In contrast, in the disulfide-bridged variants, without TCEP, fragments resulting from cleavage between F67C/K58C and F67C/K60C are not detectable because the covalent bond formed protects the protein from dissociation after fragmentation at this site. Also in this case, top-down ETD measurements confirmed both the primary sequence and the modifications in the proteins due to disulfide bridge formation.

Samples were also analyzed by FP assays to determine the K_d of the low affinity tracer compared to the wild type (see figure 3.36A). It is apparent that the disulfide bridge leads to a significant change in the K_d value, which is 25-fold for $F67C/K60C^{ox}$ and about 6-fold for $F67C/K58C^{ox}$. However, the reduced samples show higher affinity than the reduced samples in the last section (see figure 3.34), indicating that the proteins were only partially reduced here. An improvement in binding affinity for both variants was also observed in the competitive FP assay with FKN22 (kindly provided by Dr. Fabian Knaup),

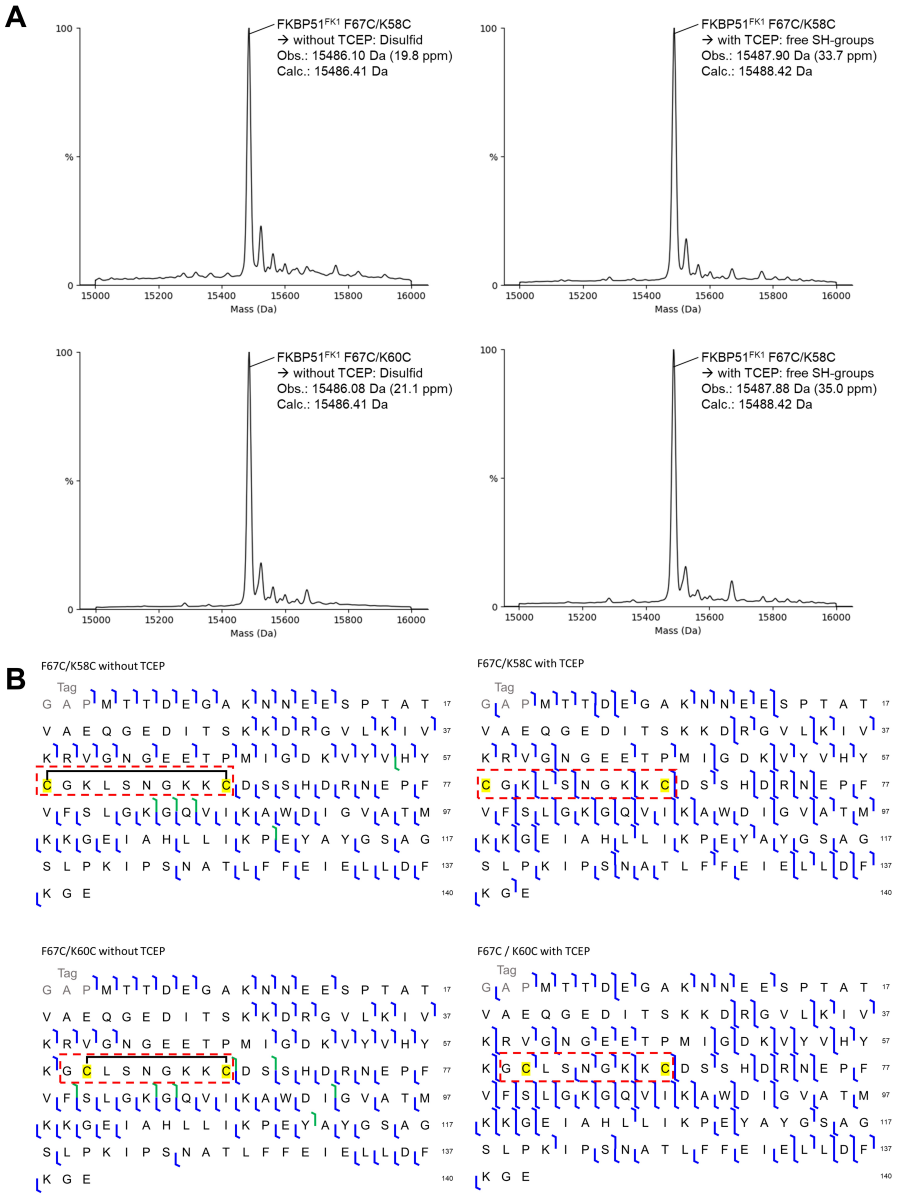
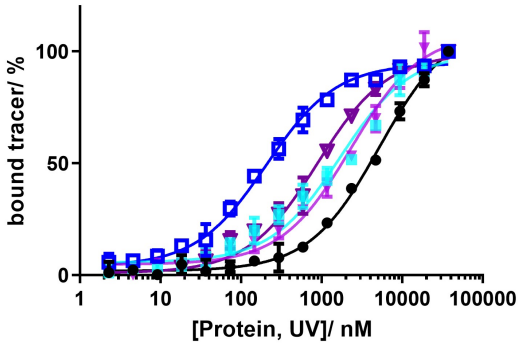


Figure 3.35: A: deconvulated MS spectra of F67C/K58C and F67C/K60C with and without TCEP. B: ETD fragmentation map from top-down measurements of F67C/K58C and F67C/K60C with and without TCEP. The c- or z-fragments are indicated in blue. In green are the c-fragments with a double H-loss due to disulfid formation. Position 67 and 60 or 58 are highlighted in yellow, and the black line is indicating the formed disulfide brigde. The numbers on the right side correspond to the FKBP51$^{FK1,\ 1\text{-}140}$ sequence, starting with methionine.

a low affinity SAFit1 analog as shown in figure 3.36B.

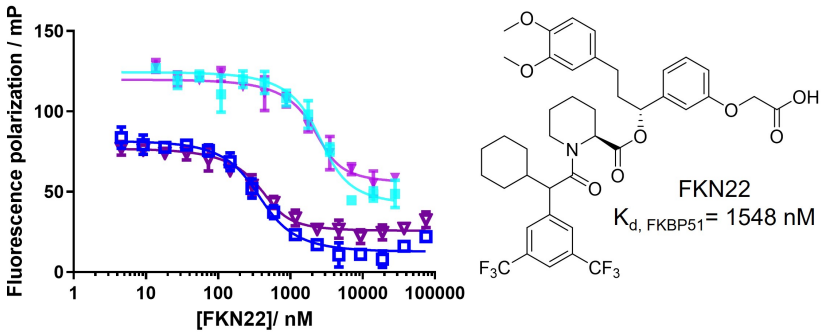
Overall, F67C/K60C in its oxidized form (with a disulfide bridge), is the winning variant to be further used as a screening tool in the search for new ligand scaffolds for FKBP51.

A



Binding curve	FKBP51^{FK1} WT	FKBP51^{FK1} F67C/K60C^{ox}	FKBP51^{FK1} F67C/K60C^{red}	FKBP51^{FK1} F67C/K58C^{ox}	FKBP51^{FK1} F67C/K58C^{red}
K_d/ nM	5181±338	206±14	1695±202	890±73	2299±202

B



Competitive mode	FKBP51^{FK1} F67C/K60C^{ox}	FKBP51^{FK1} F67C/K60C^{red}	FKBP51^{FK1} F67C/K58C^{ox}	FKBP51^{FK1} F67C/K58C^{red}
$K_{d, FKN22}$/ nM	64±12	317±218	56±21	242±108

Figure 3.36: A: binding curves of FKBP51^{FK1} (black curve), F67C/K58C (ox.: dark violet and red.: light violet curves) and F67C/K60C (ox.: dark blue and red.: light blue curves), with a low affinity tracer labeled with fluorescein. B: Competitive FP-Assay with low affinity tracer and FKN22 (Structure shown on the right side). FKN22 starting concentration for reduced proteins was 225 µM and for oxidized proteins 75 µM. The end concentration of the tracer was 1 nM and the end concentrations of the proteins were 300 nM (F67C/K60C^{ox}), 1 µM (F67C/K58C^{ox}) and 25 µM (F67C/K58C^{red} and F67C/K60C^{red}). Each data point is indicated as the mean of three technical replicats.

Three main reasons lead to this decision: First, the production of this variant is feasible and the protein can be generated in high quantity and purity. Second, the variant is stable in solution, even at high concentrations and at room temperature. While F67C/K58C^{ox} precipitated at high protein concentrations and after several minutes at room temperature, F67C/K60C^{ox} did not. Third, the variant can be crystallized in its apo form as shown in figure 3.37.

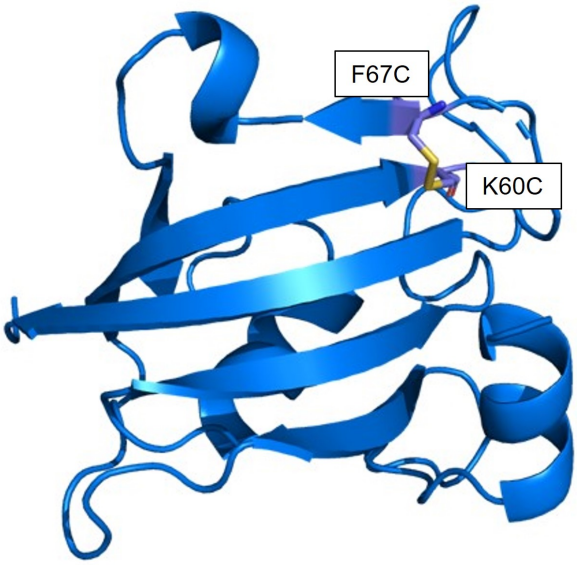


Figure 3.37: Unpublished protein structure of oxidized F67C/K60C. Disulfide bridge is highligted in purple. Structure generated by Dr. Christian Meyners.

3.1.6 Fragment library screening with a stabilized FKBP51^{FK1} F67-out variant

As described in the introduction, fragment-based drug discovery is a promising tool for finding new hits for the identification of novel lead structures. If the fragment library is carefully selected and adapted to the particular case, the probability of a successful result increases. However, multiple orthogonal methods must be selected to identify promising hits, and the binding mode must also be elucidated. Here we used the so called TRABITA (transient binding pocket) fragment library provided by the research group of Prof. Eugen Proschak from Frankfurt. The generation of this focused fragment library for transient binding pockets was performed by Laura Isigkeit during her master thesis, which was supervised by Dr. Steffen Brunst. The library is based on fragments resulting from *in silico* fragmentation of known ligands, which was subsequently converted to fragments that are commercially available. Through a virtual screening with docking approaches of the fragments with transient binding pockets, we finally received 71 fragments for experimental validation. For the first experimental screening of the fragment library, the thermal shift assay was chosen. In this assay, the protein is mixed with the fragment and heated stepwise (20-95 °C), measuring changes in the protein's intrinsic fluorescence. In this way, the stability of the protein can be determined, especially in the presence of a ligand or fragment. The FKBP51^{FK1} WT was tested in comparison to the F67Y and F67C/K60C^{ox} variants. The DMSO control was compared here with the other samples containing the SAFit1 ligand or fragments. Screening with the winning variant F67C/K60C^{ox} identified eleven hits, which are summarized in table 7.2.

Table 3.4: Determined melting temperatures (T_m) of FKBP51^{FK1} WT, F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM). ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	WT		F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	53.7	-	48.1	-	50.4	-
SAFit1	58.5	4.8	64.3	16.2	64.3	13.9
G2	52.8	-0.9	48.1	0.0	52.4	2.0
H2	53.0	-0.7	-	-	52.6	2.2
A3	53.1	-0.6	47.1	-1.0	53.6	3.2
C3	52.8	-0.9	50.3	2.2	53.0	2.6
H3	55.9	2.2	48.4	0.3	52.5	2.1
G4	52.7	-1.0	47.6	-0.5	53.3	2.9
G5	-	-	48.7	0.6	53.3	2.9
F7	52.8	-0.9	47.7	-0.4	53.0	2.6
C8	52.9	-0.8	47.5	-0.6	52.8	2.4
H8	56.2	2.5	49.5	1.4	52.8	2.4
F9	53.0	-0.7	47.7	-0.4	53.2	2.8

The wild-type protein has the highest melting point of 53.7 °C, while F67Y and F67C/K60C^{ox} have melting temperatures of 48.1 °C and 50.4 °C, respectively. However, the ligand SAFit1 shows a higher stabilization effect on the variants than on the wild-type protein, which is also shown in a graph in figure 3.38.

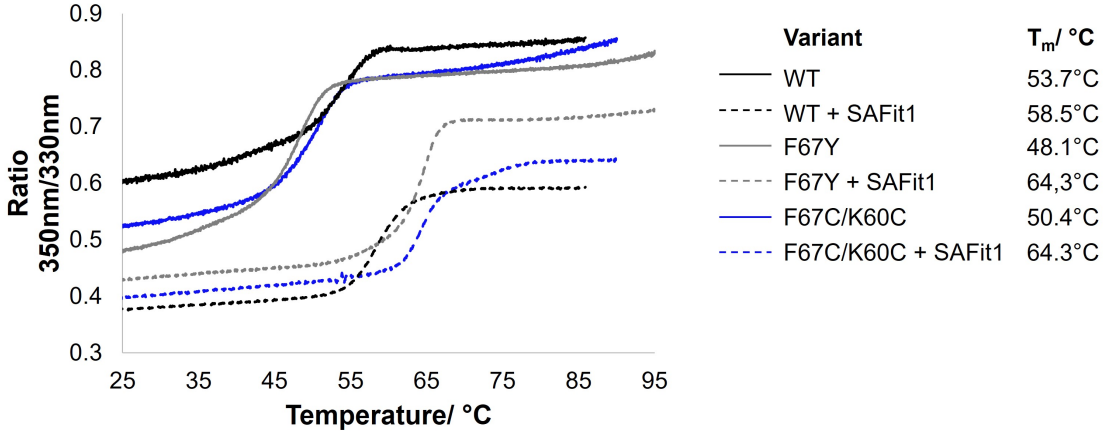


Figure 3.38: Melting curves of FKBP51^{FK1} WT, F67Y and F67C/K60C^{ox} with and without SAFit1.

Even though the difference in melting temperature for the FKBP51^{FK1} variants with SAFit1 is above 10 °C, such a change in T_m is not expected for the fragments as they contribute a lower stabilization energy. When screening the fragments, it is noticeable that the eleven hits were identified only with the high-affinity protein variant F67C/K60C^{ox}. A hit was considered a positive hit if the difference in melting temperature compared to the

DMSO control sample was greater than 1.5 °C. With this chosen threshold, only fragments H3 and H8 would have been positive hits when screening with the wild-type protein, and only C3 when screening with F67Y. However, the initial thermal shift assay is not sufficient to determine whether the fragment truly binds in the binding pocket or even less what the binding mode is. To assess whether the fragments could displace the tracer with known binding affinity used in the previous experiments, competitive FP assays were performed with SAFit1 as a control ligand. Six of eleven hits from the thermal shift assay actually show displacement of the low affinity tracer, as indicated by the decrease in fluorescence polarization with increasing fragment concentration as shown in figure 3.39.

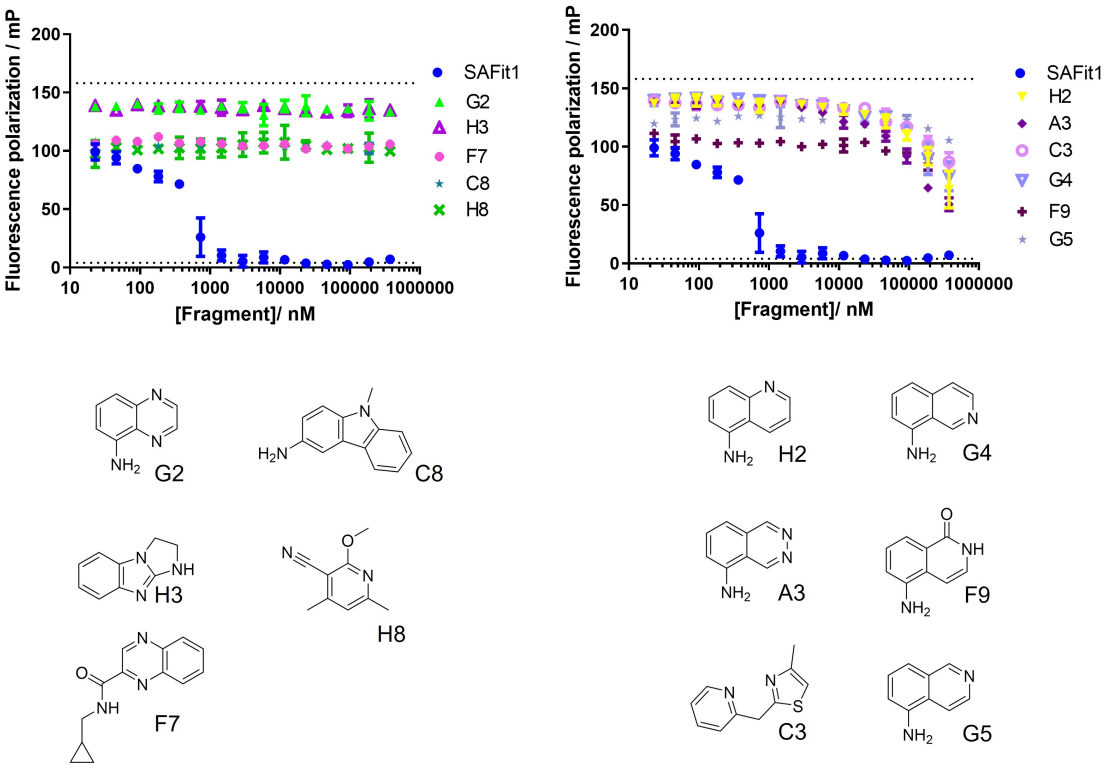


Figure 3.39: Competitive FP-assay. On the left side, no displacement of the tracer with low binding affinity is seen, while on the right side the hits are shown. Below the chemical structures of the fragments are displayed. The fragments were used with a starting concentration of 375 µM. The final protein concentration was 975 nM and the tracer concentration was 1 nM.

The validated positive hits H2, A3, G4, G5 and even F9 show a great similarity in structure as they are amine substituted quinoline and isoquinoline derivatives. Figure 3.40A shows the fals positive hits compared to the validated true positive hits with similar structure. It is apparent that the amine-substituted naphthalene core structure has preferred positions for the nitrogen, namely positions 5, 6 and 7. Based on these findings, new fragments were selected for testing that were commercially available. The 16 new fragments are shown in figure 3.40B. In a competitive FP assay, only the fragment 1,6-naphthyridin-5-amine showed the ability to display the tracer with low affinity as shown in figure 3.40C.

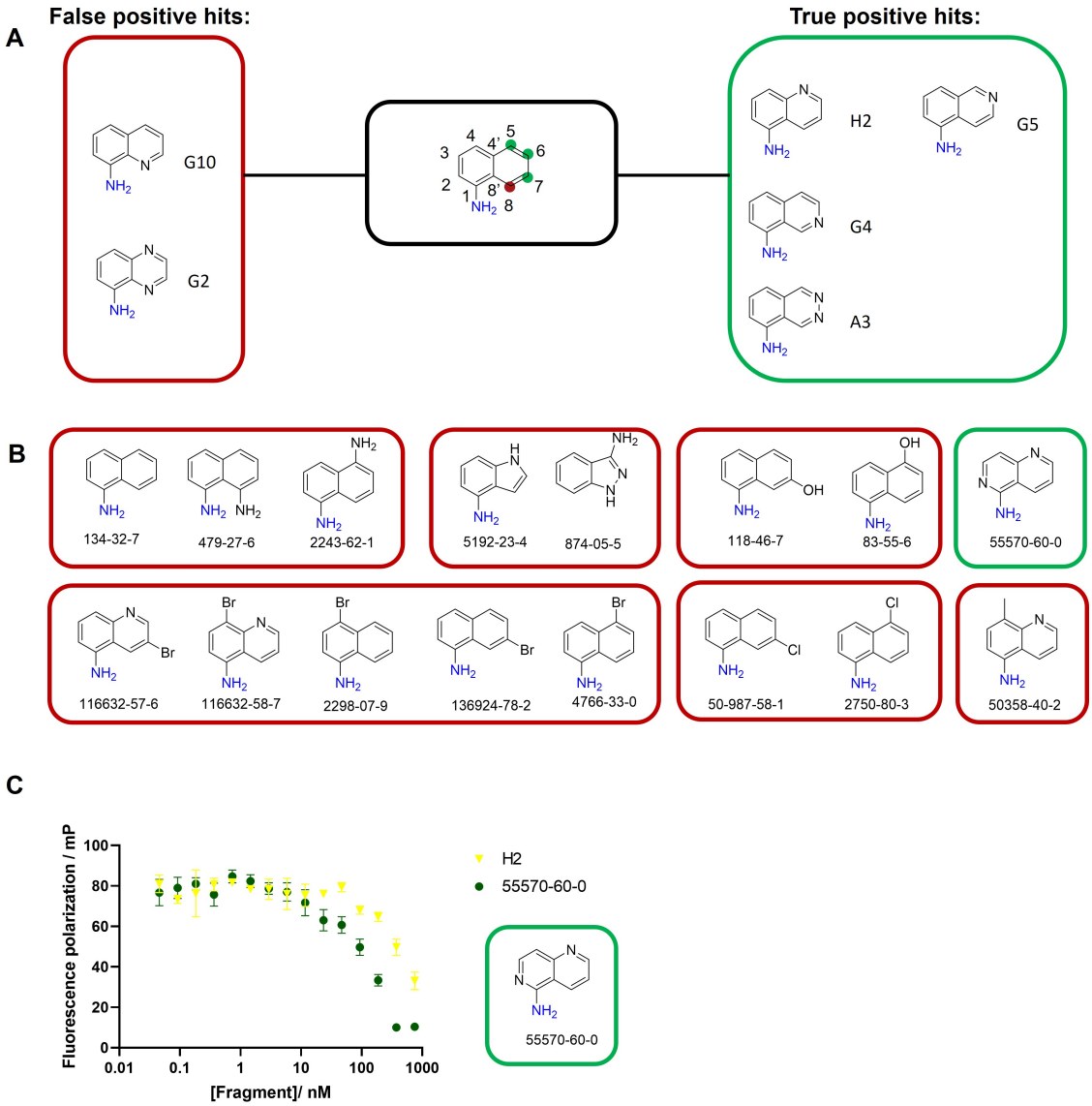


Figure 3.40: A: analysis of false and true positive hits to identify favorable positions in the identified amine-substituted naphthalene structure. B: structures of the new fragments to be tested. C: competitive FP assay of the new positive hit 1,6-naphthyridin-5-amine. The fragments were used with a starting concentration of 375 µM. The final protein concentration was 975 nM and the tracer concentration was 1 nM.

In addition, the company Atomwise was commissioned to create and provide a compound library for screening. The Atomwise library, consisting of 96 compounds, was also tested using the thermoshift assay, as shown in table 3.5.

Compounds with a melting temperature difference of more than 0.9 °C for the variant F67C/K60C^{ox} are listed here. When the threshold for melting temperature difference is set again to 1.5 °C, seven hits are identified. The compounds were tested in a competitive FP assay, resulting in two hits: A9 and C1 (see figure 3.41). Although the melting temperature difference of compound D2 is only 1.1 °C, the compound shows displacement of the low

Table 3.5: Determined melting temperatures (T_m) of FKBP51^{FK1} WT, F67Y and F67C/K60C^{ox} (each 60 µM) with compounds (500 µM) from the Atomwise library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding compound.

Compound	WT		F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	53.7	-	47.6	-	50.1	-
SAFit1	58.5	4.8	64.0	16.4	62.8	12.7
C1	52.9	-0.8	47.6	0.0	53.2	3.1
D2	-	-	47.2	-0.4	51.2	1.1
A3	53.1	-0.6	47.8	0.2	51.0	0.9
H3	52.5	-1.2	47.2	-0.4	53.1	3.0
G5	-	-	47.3	-0.3	51.6	1.5
C6	53.3	-0.4	47.5	-0.1	52.9	2.8
B7	55.1	1.4	47.6	0.0	52.4	2.3
H7	53.0	-0.7	47.7	0.1	51.3	1.2
A8	52.5	-1.2	48.8	1.2	51.0	0.9
A9	53.9	0.2	48.7	1.1	51.6	1.5
F9	52.7	-1.0	48.1	0.5	51.9	1.8
H12	52.5	-1.2	47.8	0.2	51.2	1.1

affinity tracer in the competitive FP assay, which eventually resulted in 3 hits from the Atomwise library. It is noticeable that some compounds precipitate at high concentrations, as indicated by the increase in fluorescence polarization at high concentrations of compound in some samples. The final hits from the Atomwise library have a thiazole (A9 and C1) or thiophene (D2) moiety, indicating the importance of these types of motifs for conformation-specific FKBP51 ligands. For example, the thiophene motif has already been studied by Dr. Fabian Knaup in his recent publication[150], confirming that thiophenes are an efficient replacement for the cyclohexyl moiety in the SAFit structure responsible for flipping the residue F67 out of the binding pocket. However, a co-crystal structure of D2 and the protein is not yet available, so only assumptions about the orientation of the hit in the binding pocket are possible. Interestingly, Dr. Christian Meyners succeeded in co-crystallizing compound A9 with FKBP51^{FK1} F67C/K60C^{ox}, showing that the thioazole on the other side forms a hydrogen bond with the amine of the backbone of residue I67. The extended space due to the F67-out-like conformation is filled with the 2-methylbenzofuran moiety of compound A9.

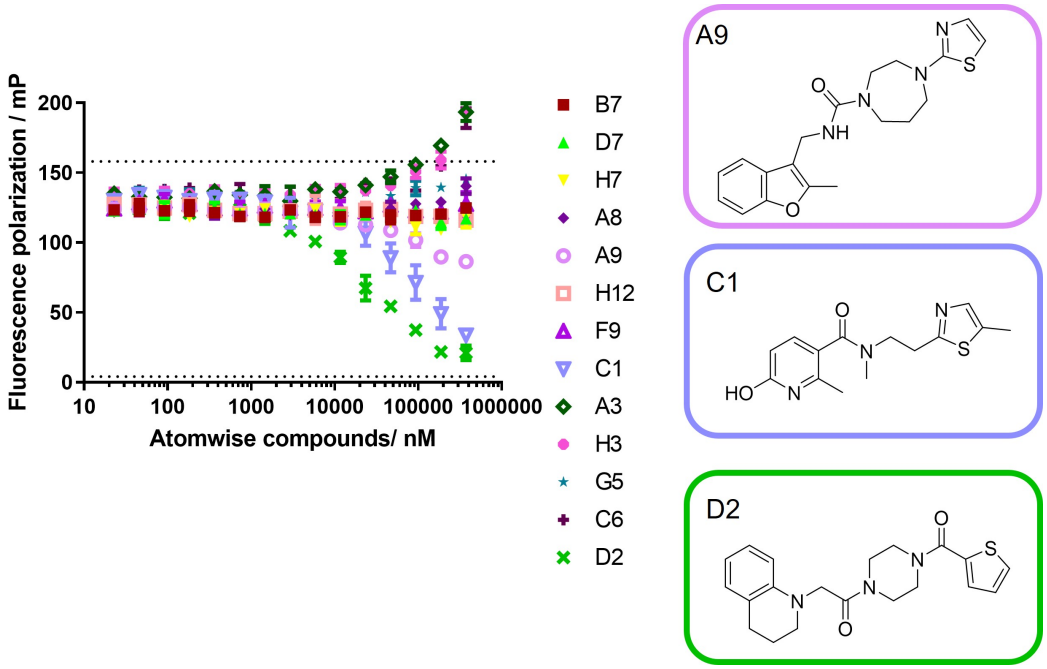


Figure 3.41: Competitive FP-assay with a low affinity tracer (1 nM) and the prescreened compounds (375 µM) from the Atomwise library with F67C/K60C^{ox} protein (975 nM). Hit compounds A9, C1 and D2 are displayed on the right side.

3.2 Ligand based affinity chromatography

Ligand-based affinity chromatography has the potential to enrich the protein of interest and identify interaction partners or off-targets.

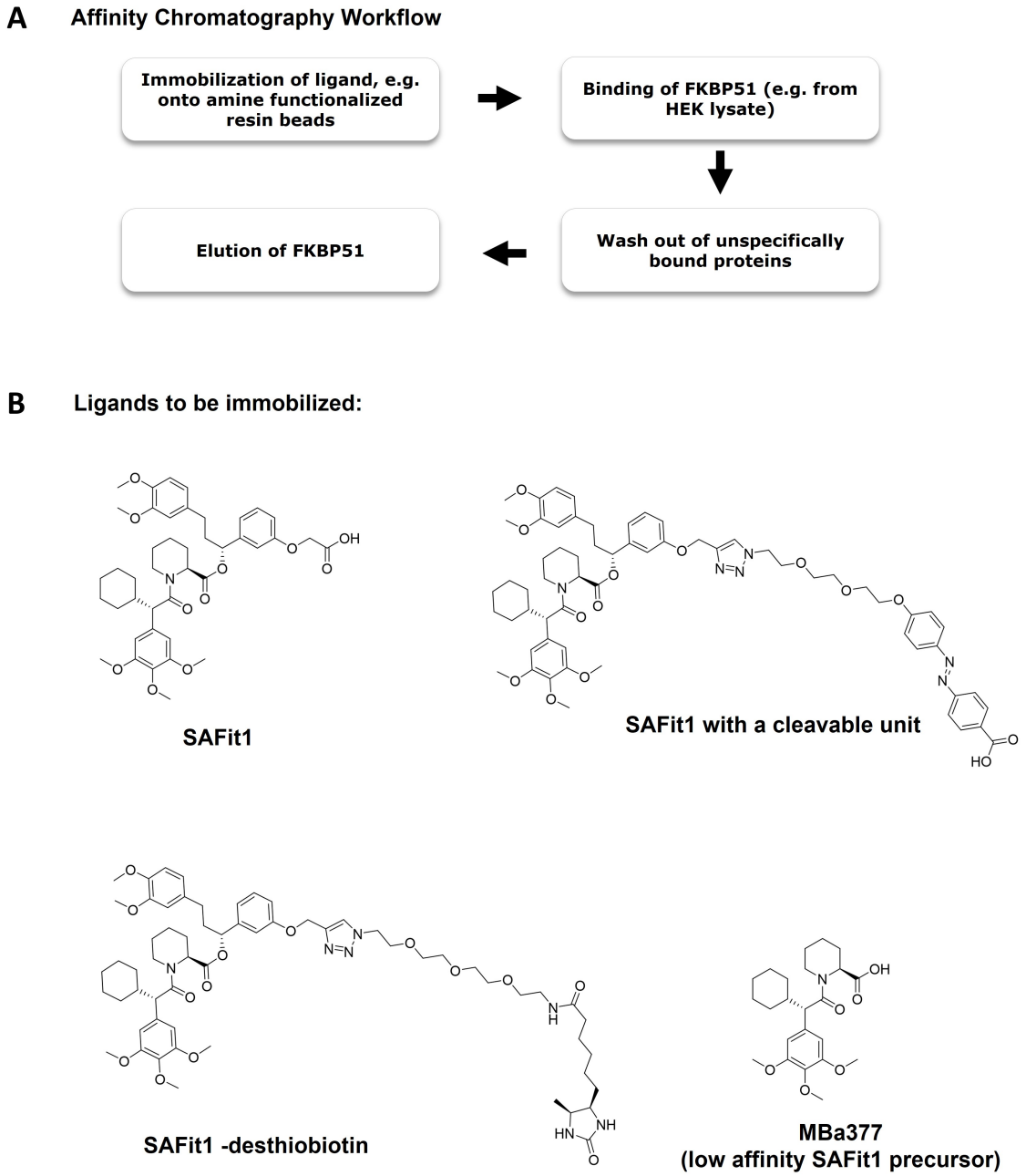


Figure 3.42: A: 4-step workflow for ligand-based affinity chromatography. B: chemical structures of four ligands immobilized for affinity chromatography in this work.

Therefore, in this chapter, SAFit1 and analogs are immobilized and used in different experimental setups depending on the research question. The workflow generally consists of a step of ligand immobilization on a solid support, followed by loading with a sample,

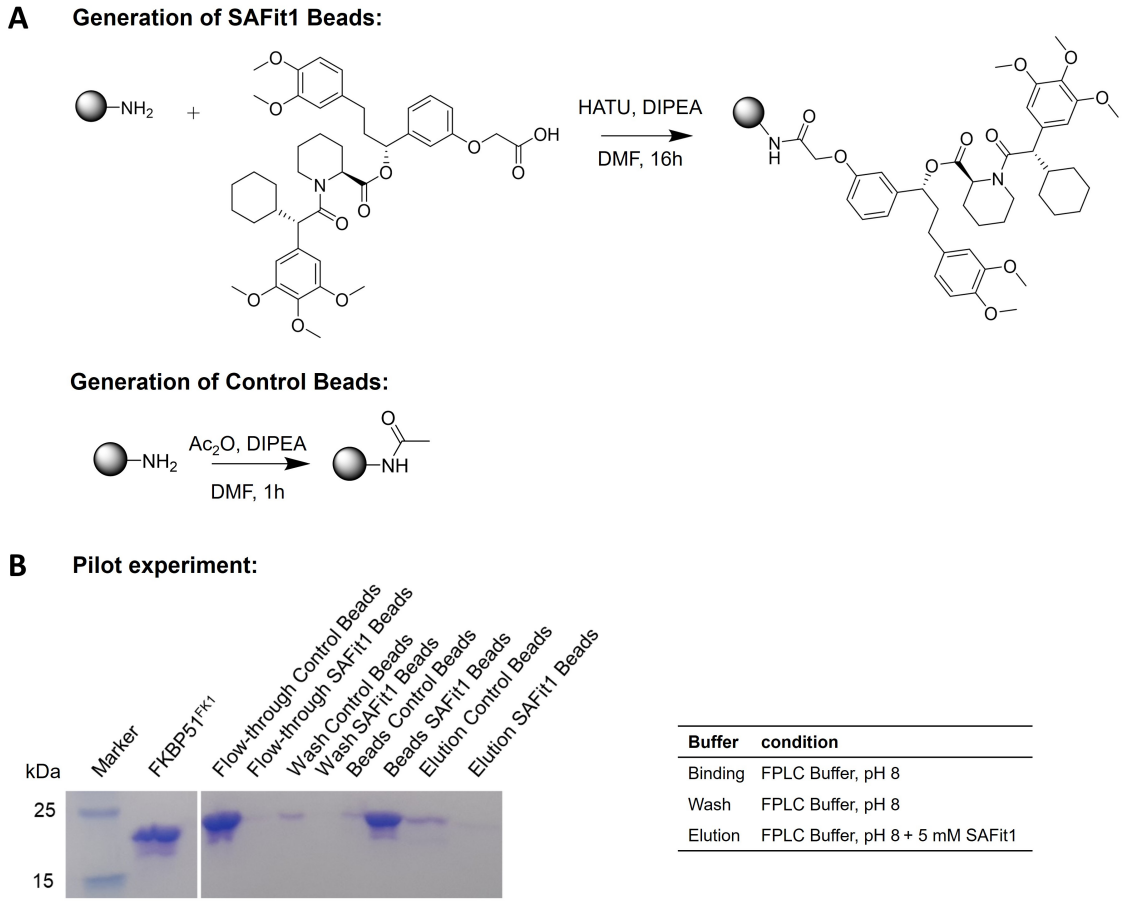
e.g. cell lysate, washing of the solid support to remove unbound protein, and in the last step elution of the target protein (see figure 3.42A).

3.2.1 Generation and examination of ligand-based affinity matrices

In this work, four ligands were tested. Their chemical structures are shown in figure 3.42B and discussed in the following sections.

Immobilization of SAFit1 on amine-functionalized beads

First, the ligand was immobilized on a suitable solid support. In this case, polymethacrylate beads with a terminal amino group are used to immobilize SAFit1 via the formation of an amide bond (see figure 3.43A). As a control experiment, the same amine containing resin was fully acetylated. A Kaiser test ensured complete capping of the amine groups on the beads.



Buffer	condition
Binding	FPLC Buffer, pH 8
Wash	FPLC Buffer, pH 8
Elution	FPLC Buffer, pH 8 + 5 mM SAFit1

Figure 3.43: A: Immobilization of SAFit1 on amine-functionalized beads by amide bond formation and generation of control beads by acetylation. B: SDS-PAGE gel of the analyzed affinity chromatography samples of the pilot experiment with pure FKBP51^{FK1} and competitive SAFit1 elution.

To test whether the generated SAFit1 beads are indeed capable of capturing FKBP51, recombinantly produced and purified FKBP51^{FK1} was used. After loading the sample, the

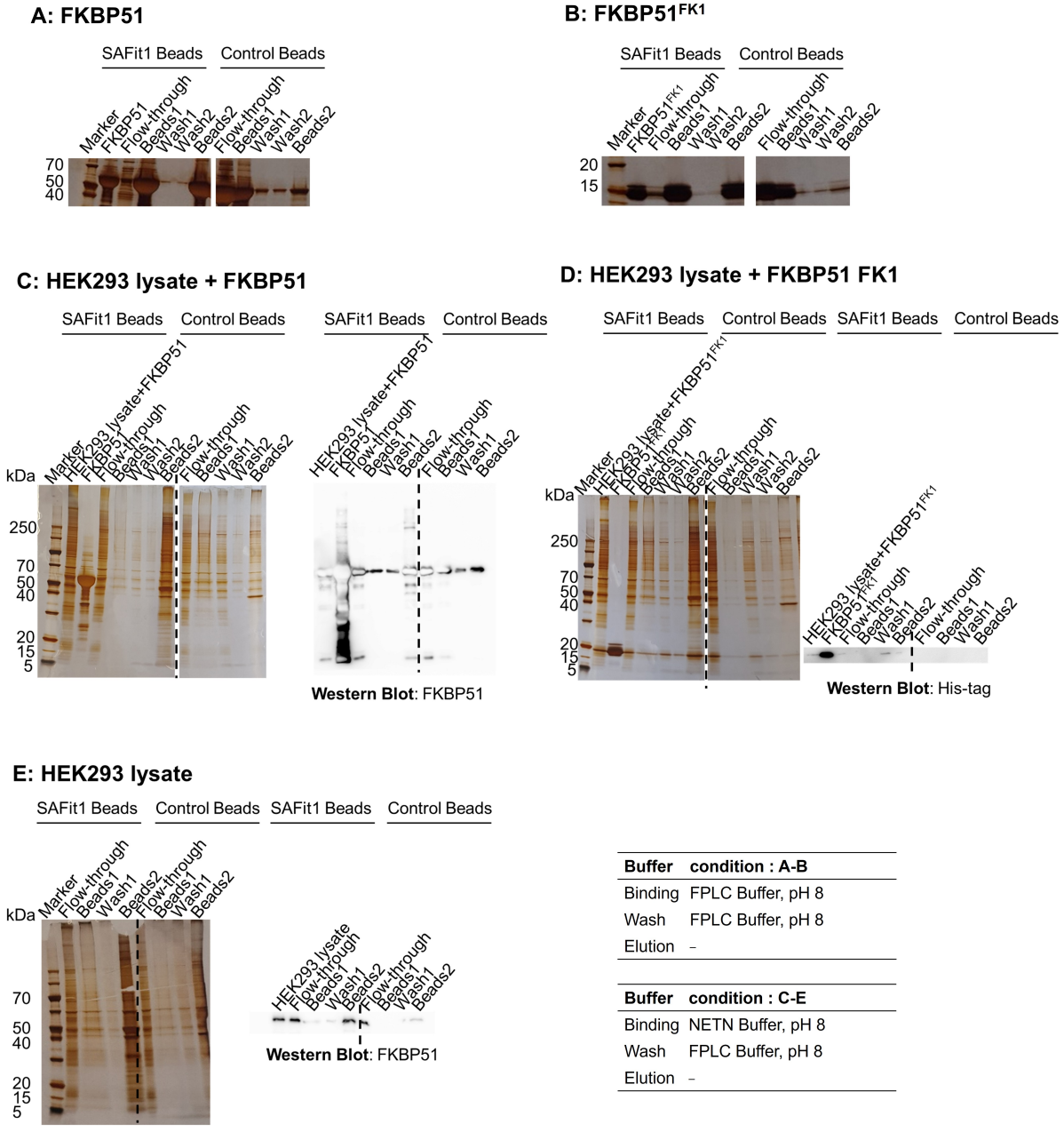
flow-through was collected and analyzed by SDS-PAGE. As expected, the flow-through fraction of the Control Beads shows an FKBP51^{FK1} band, while in the experiment with the SAFit1 Beads no FKBP51^{FK1} band was observed. This indicates successful capture of the protein on the beads, which is further confirmed by the so called Beads samples of the Control Beads compared to the SAFit1 Beads. Here, the Beads samples are obtained by boiling a small fraction (usually 10 µL) of the beads in Laemmli buffer to determine the loading of the beads. No or little signal was observed in the Wash samples of Control Beads and SAFit1 Beads because the sample loaded onto the beads was a purified protein. However, it is noticeable that competitive elution with SAFit1 (immobilized SAFit1 vs. SAFit1 in solution) was unsuccessful. Therefore, the capture of FKBP51 was examined in more detail.

Purified FKBP51 was tested again. This time the recombinantly obtained full length FKBP51 protein (figure 3.44A) as well as the FK1 domain (figure 3.44B) was examined. The full length FKBP51 protein was kindly provided by Thomas Geiger and the FK1 domain by Dr. Stephanie Merz and Johanna Bartmuß. The full length FKBP51 has still impurities, as seen in the first lane of the SDS-PAGE gel (silver stained) in figure 3.44A. Nevertheless, it is clear that the recombinant FKBP51 sample was captured by the SAFit1 Beads when comparing the flow-through and Beads2 samples of both conditions (Control Beads vs. SAFit1 Beads). The flow-through sample from the SAFit1 Beads shows only a weak protein band compared to the flow-through sample from the Control Beads. However, it is clear that non-specific binding to the Control Beads is occurring, as the Beads1 sample (before washing) shows a strong protein band under both conditions. For the Control Beads, most of the protein is washed away, leaving only a weak protein band in the Beads2 sample for the Control Beads. The same is true for the FK1 domain of FKBP51 as shown in the SDS-PAGE gel in figure 3.44B, indicating that binding of the binding domain as well as the full-length protein is not affected by immobilization of the ligand.

To investigate the capturing of FKBP51 in complex biological mixtures such as cell lysates, HEK293 lysate was prepared and used in the next affinity chromatography experiments. Three conditions were tested: HEK293 lysate spiked with purified FKBP51 or the FK1 domain, and unspiked HEK293 lysate. In addition to SDS-PAGE used to analyze the samples of the experiments, Western Blots were performed with anti FKBP51 and anti His-tag antibodies, respectively. In figure 3.44C the corresponding silver stained SDS-PAGE gels and Western Blots are shown. The FKBP51 protein band is detectable in all samples here, regardless of the sample (Control Beads or SAFit1 Beads). For example, the flow-through samples for the Control Beads and the SAFit1 Beads show approximately the same protein band intensity, indicating unsuccessful or at least nonspecific capture of FKBP51 from cell lysate when the Beads2 samples are also compared. For the lysate spiked with the FK1 domain, on the other hand, almost no protein bands are detected in the Western Blot, except for the pure FKBP51^{FK1} sample (see figure 3.44D). The spiked HEK293 lysate sample in lane 1 already shows a very low intensity of the protein and the FK1 domain can only be visibly detected in the Flow-through, Wash1 and Beads2 fractions for the SAFit1 Beads. The plain HEK293 lysate (see figure 3.44E) also shows insufficient capturing of FKBP51, as the flow-through samples for the Control Beads and SAFit1 Beads show similar protein band intensities.

In summary, the initial experiments using immobilized SAFit1 as an affinity chromatography matrix raise two issues. First, capture of pure full-length FKBP51 and the FK1 domain undoubtedly works, but capture of the protein from complex mixtures does not appear to be specific nor efficient. In addition, the pilot experiment showed that elution

with 5 mM SAFit1 was not sufficient for competitive elution despite the excess of compound in the elution buffer.



Buffer	condition : A-B
Binding	FPLC Buffer, pH 8
Wash	FPLC Buffer, pH 8
Elution	–

Buffer	condition : C-E
Binding	NETN Buffer, pH 8
Wash	FPLC Buffer, pH 8
Elution	–

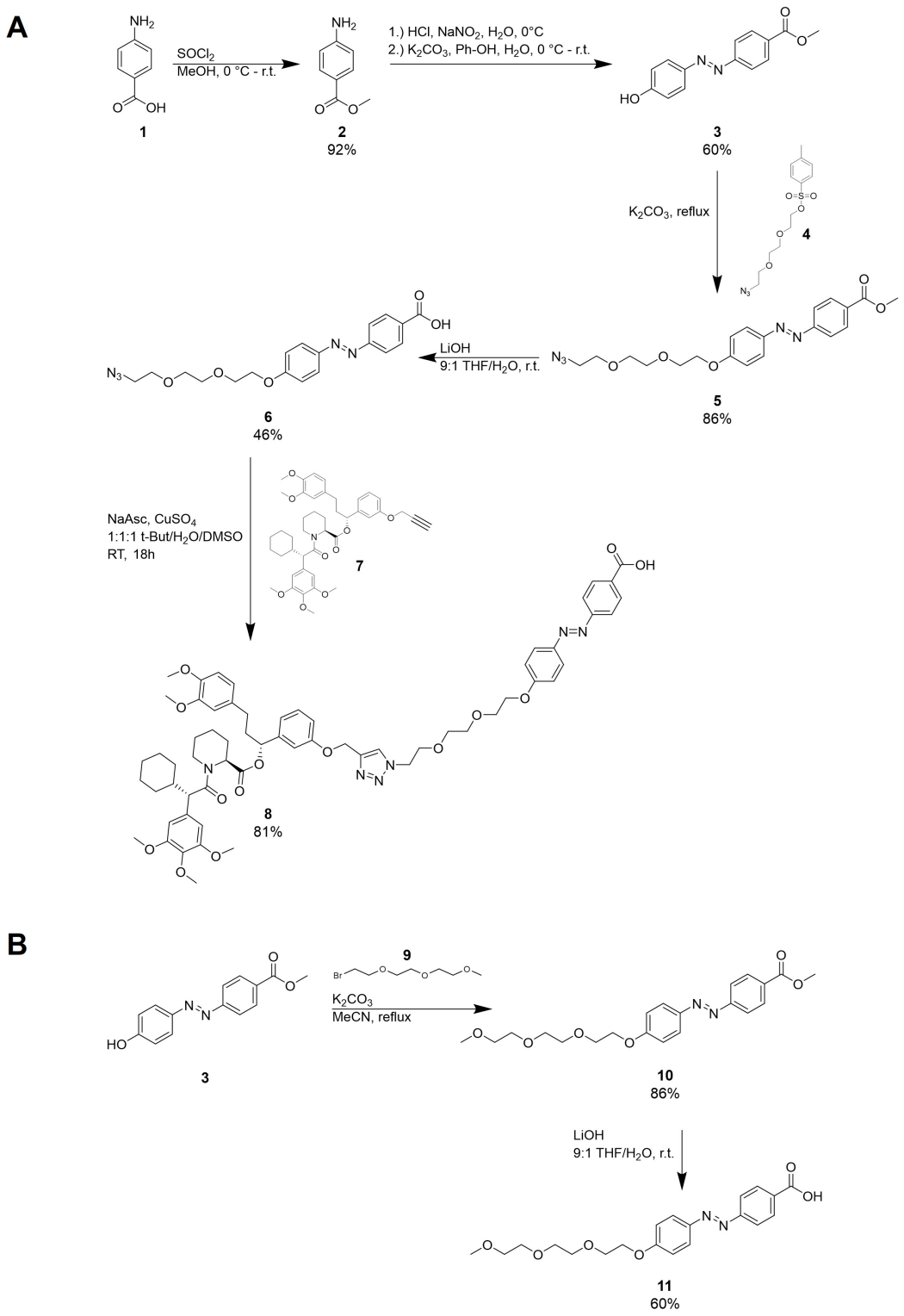
Figure 3.44: Five experimental setups of SAFit1-based affinity chromatography using A: purified full-length FKBP51, B: purified FKBP51^{FK1}, C: HEK293 lysate spiked with purified full-length FKBP51, D: HEK293 lysate spiked with the purified FK1 domain of FKBP51, and E: Unspiked HEK293 lysate as loading samples. The buffers used for each condition are displayed in the lower right corner.

Immobilization of SAFit1 with a cleavable unit on amine-functionalized beads

A possibility to solve the elution issue is the use of a cleavable linker, as described in more detail in the introduction section 1.3.2. In this work a diazobenezne cleavable crosslinker was synthesized in 5 steps (see scheme 3.45A), starting with the protection of the carboxy-group of 4-aminobenzoic acid **1** to methyl 4-aminobenzoate **2** using methanol under acid conditions. In the next step, compound **2** and phenol are used in an azo coupling reaction leading to compound **3**. This reaction occurs in two steps: First, the aromatic amine forms a diazonium salt in the presence of nitrous acid, which then reacts with the phenol to form the azo compound **3**. In order to leave a molecular spacer between the ligand (SAFit1) and the cleavable moiety, a polyethylene glycol (PEG)-based linker was introduced using compound **4** which was kindly provided by Dr. Michael Walz and Dr. Horst W. Schuchmann. This nucleophilic substitution reaction was followed by cleavage of the methyl ester with lithium hydroxide, resulting in compound **6**. Since compound **4** contains an azide group, the SAFit building block **7**(provided by Dr. Tianqi Mao) could be introduced via click reaction, resulting in the formation of the final compound SAFit-CLU (**8**) with an overall yield of approximately 18 % (see scheme 3.45A). For the synthesis of the control, compound **3** was coupled with compound **9** by nucleophilic substitution, followed by ester hydrolysis using LiOH leading to the control compound **11** with an overall yield of 28 % (see scheme 3.45B).

The control with a cleavable linker was first tested in solution for cleavage. For this purpose, sodium dithionite was added to the control compound **11** in FPLC buffer and incubated for 15 min at room temperature. The initial yellow solution lost its color, indicating cleavage of the linker (see figure 3.46A). In addition, the solution of compound **11** before and after the addition of the reducing agent sodium dithionite was measured using an LC-MS system. In both chromatograms (see figure 3.46B), compound **11** was detected by mass spectrometry, but the addition of sodium dithionite resulted in a less intense band (see figure 3.46C). It is noticeable that the chromatogram of the sample without sodium dithionite shows an additional peak around 0.3 min, which can be assigned the same mass as that of the compound **11**. This is most likely the cis isomer of the compound, which, in contrast to the trans isomer, has a different elution time.

Since the in-solution cleavage experiment showed the desired effect, compound **8** (SAFit1-CLU) was immobilized on the amine-functionalized beads under the same procedure as shown in figure 3.43A for SAFit1. As a control, compound **11** was also immobilized on amine-functionalized beads. The modified beads were used next to test whether the first issue of FKBP51 capturing from complex mixtures is buffer dependent and at the same time to investigate the elution efficiency with and without the detergent SDS. In figure 3.47A, pure FKBP51 was tested in the usually used FPLC buffer (20 mM HEPES, pH 8, 20 mM NaCl), showing successful capture of the protein on the SAFit1-CLU Beads when comparing the flow-through samples of the Control Beads with the SAFit1-CLU Beads. However, the Control Beads show non-specific binding of the protein in the Beads1 sample, but this can be washed away. Four elution conditions were also tested. Condition a contains only 25 mM sodium dithionite, while conditions b-d add SDS at concentrations of 0.2 %, 0.5 %, and 1 %, respectively. Only the addition of SDS shows FKBP51 in the elution fractions when analyzed by SDS-PAGE. The most efficient SDS concentration was 1 % as showed in figure 3.47A condition 1d. This suggests that the cleavable linker, when immobilized, is eventually not cleavable. Even though the elution issue has not been solved yet, the issue of capturing could be solved in this experiment as it depends on the buffer. In figure 3.47B, it is evident that when pure FKBP51 is diluted in NETN buffer,



Scheme 3.45: A: synthesis of SAFit1 with a cleavable unit (SAFit1-CLU, **8**). B: synthesis of the control compound (**11**).

which is used as a lysate buffer to prepare HEK293 lysate, the capturing of FKBP51 on the SAFit1-CLU Beads is unsuccessful, as the flow-through samples of the SAFit1-CLU Beads

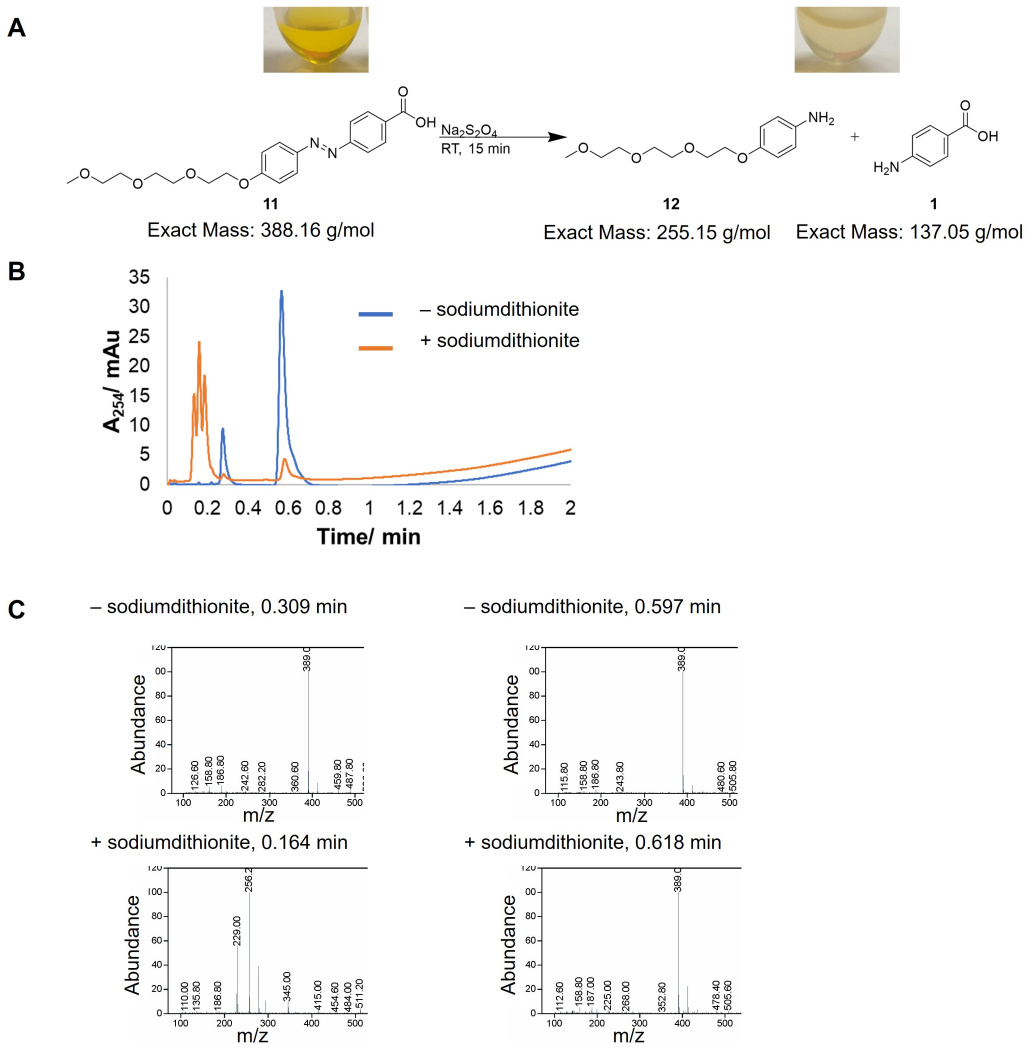


Figure 3.46: A: chemical reaction equation for the cleavage of compoung **11** with sodium dithionite. The originally yellow colored solution turned into a colorless solution after addition of sodium dithionite. B: the samples were analyzed by LC-MS, showing the chromatograms of the sample without sodium dithionite (blue curve) and after cleavage (orange curve). C: respective MS data of the peaks.

and the Control Beads both show a very intense FKBP51 protein band. This issue can be solved by using the commercially available cell lysate buffer M-PER. The NETN buffer contains 100 mM NaCl, 20 mM Tris, pH 8, 0.5 mM EDTA, 0.5 % Nonidet P-40 and freshly added DTT (Dithiothreitol), PMSF (phenylmethylsulfonyl fluoride) and protease inhibitor mix. In contrast, the M-PER buffer has a simple composition consisting of a proprietary detergent in 25 mM bicin buffer with a pH of 7.6. It is likely that the detergent in the NETN buffer Nonidet P-40 is not compatible with ligand-based affinity chromatography experiments, while the M-PER buffer appears to contain a suitable detergent for cell lysis, but still does not interfere with the loading step in affinity chromatography.

After determining that the NETN buffer caused the capture issue, new HEK293 lysate was prepared in M-PER buffer and tested in affinity chromatography experiments with three

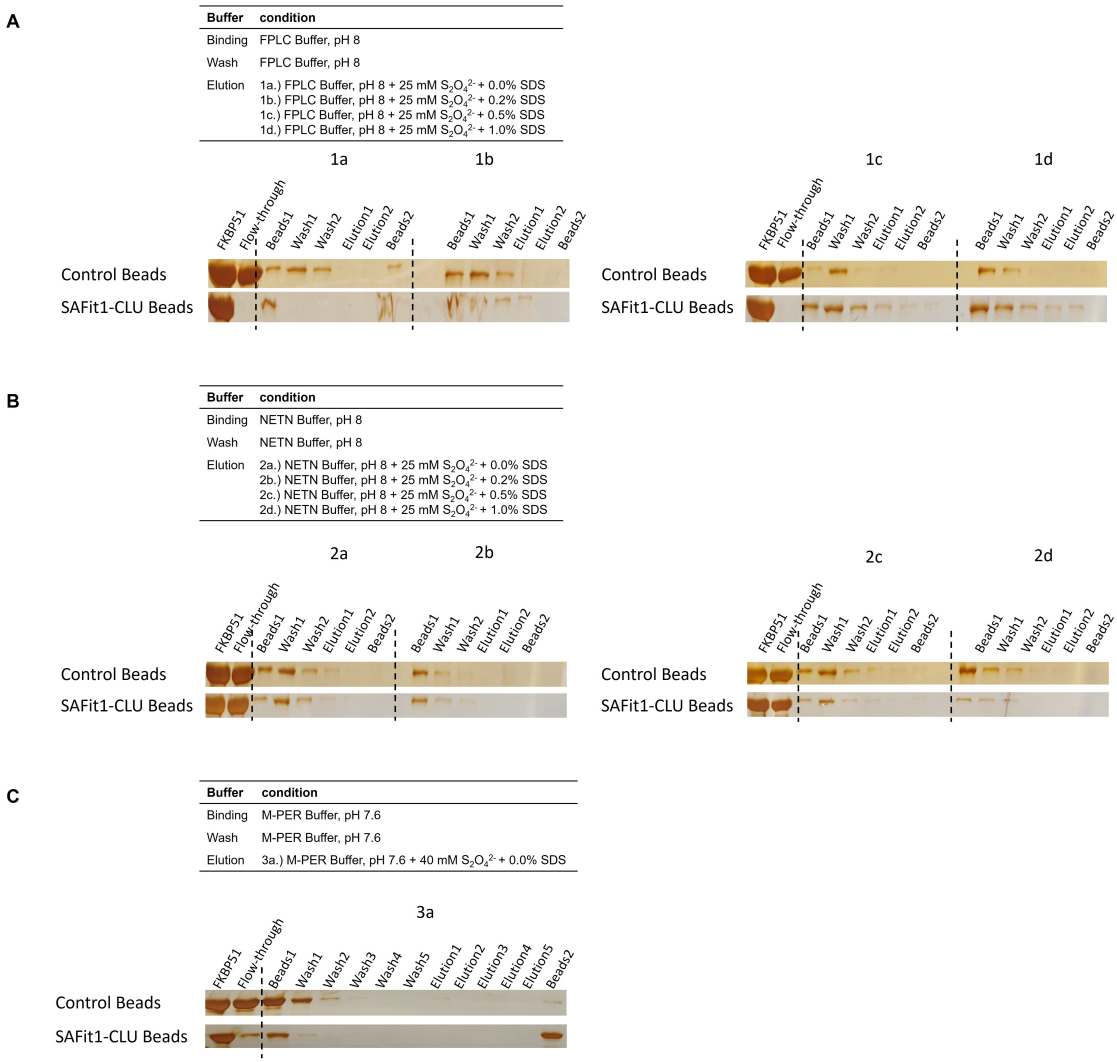


Figure 3.47: SDS-PAGE analysis of Control Beads and SAFit1-CLU Beads in an affinity chromatography experiment with pure FKBP51 (A-C) and 4 different elution conditions (A, B). To test the loading efficiency buffer-dependently, FPLC buffer (A), NETN buffer (B) and commercial M-PER buffer (C) were examined, FPLC buffer (A), NETN buffer (B), and commercially available M-PER buffer (C) were examined. The latter experiment was performed by Simon Reiners during his internship in the Hausch Lab under my supervision.

different elution conditions performed by Simon Reiners. Figure 3.48 shows the SDS-PAGE gels and respective Western Blots for the characterization of the affinity chromatography fractions.The loading of FKBP51 onto SAFit1-CLU Beads was successful, as shown by the comparison of the Flow-through samples between the Control Beads and the SAFit1-CLU Beads, where the FKBP51 signal was less intense in the SAFit1-CLU Beads. However, competitive elution with 5 mM SAFit1 (figure 3.48B) or $FKBP51^{FK1}$ (figure 3.48C) showed no effect, whereas elution with low concentrations of SDS (figure 3.48A) was observed in the Western Blot.

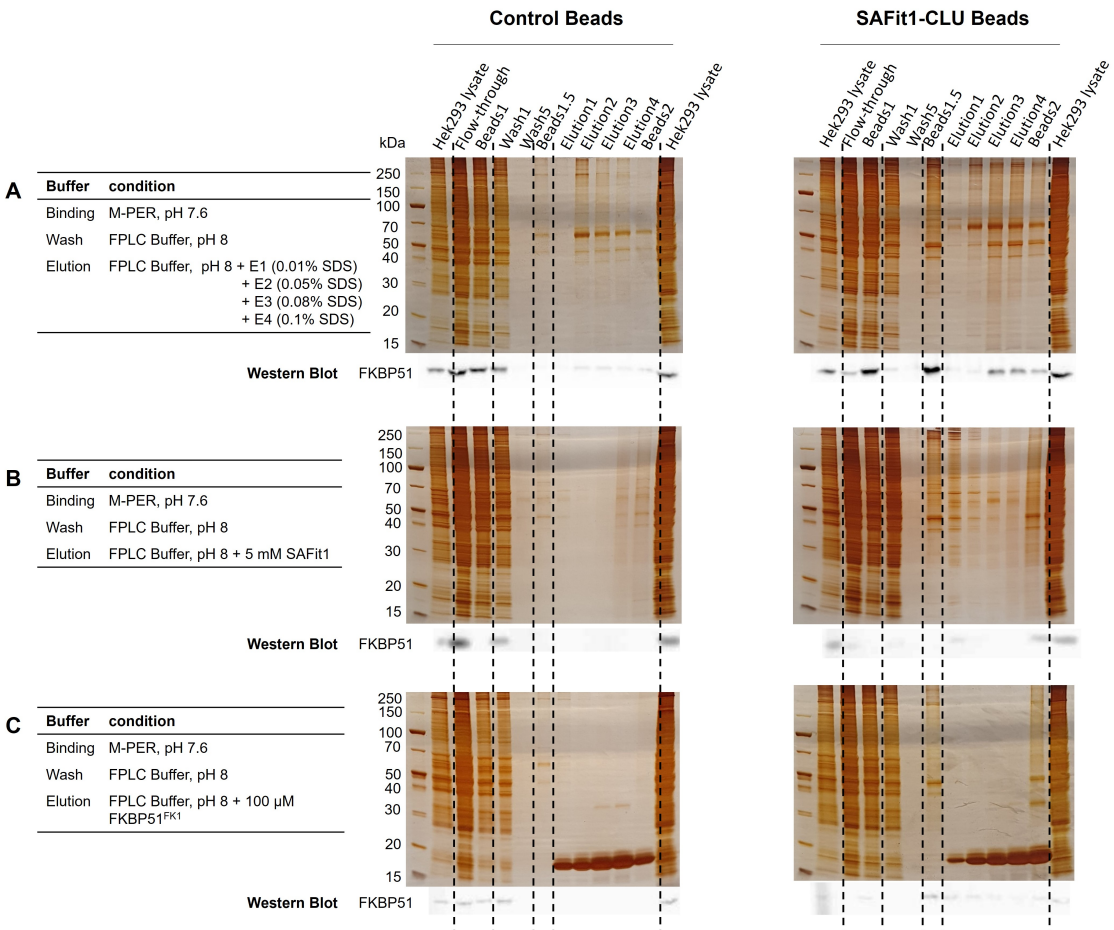


Figure 3.48: SDS-PAGE analysis and Western Blots of affinity chromatography experiments with 3 tested elution conditions: A: SDS, B: 5 mM SAFit1 and C: FKBP51^{FK1}. These experiments were conducted by Simon Reiners.

Immobilization of SAFit1-desthiobiotin on streptavidin-functionalized beads

Since SDS is a known ion suppressor[151,152] and subsequent MS measurements would need the additional step of SDS removal, a new affinity chromatography system avoiding SDS was tested. Additionally, the elution specificity has to be enhanced. Therefore, the SAFit1 analog **7** was clicked to a desthiobiotin analog containing a PEG (polyethylene glycol)-linker (compound **12**), resulting in compound **13** with a yield of 40 %.

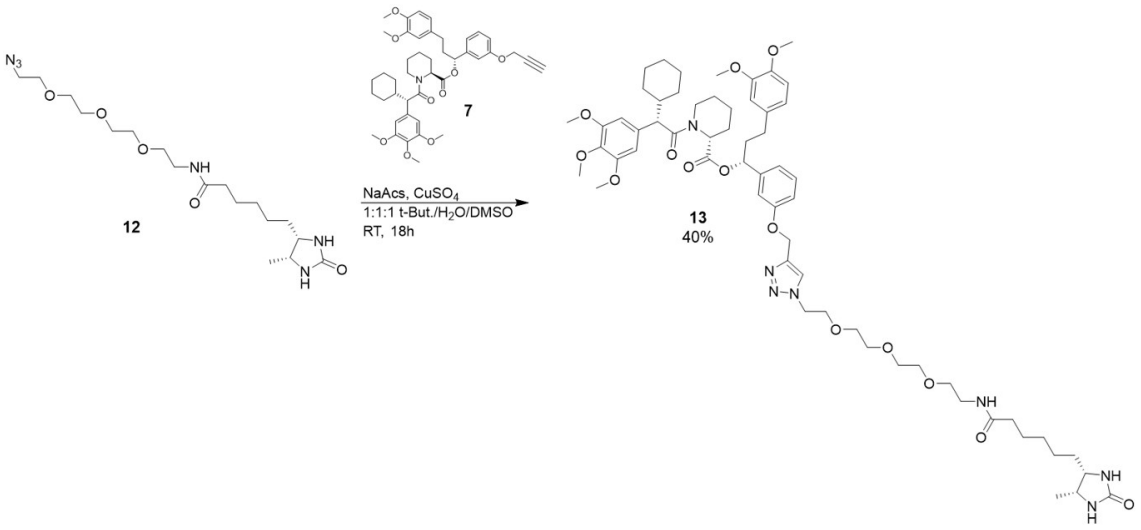


Figure 3.49: Synthesis of compound **13**, a SAFit1 analog coupled to a PEG-linker and a desthiobition moiety.

Desthiobition is a biotin analog that binds streptavidin with a lower binding affinity than biotin. Immobilization of a desthiobitionylated ligand on a streptavidin matrix should allow elution of the ligand with the respective loaded proteins by competitive elution with biotin.

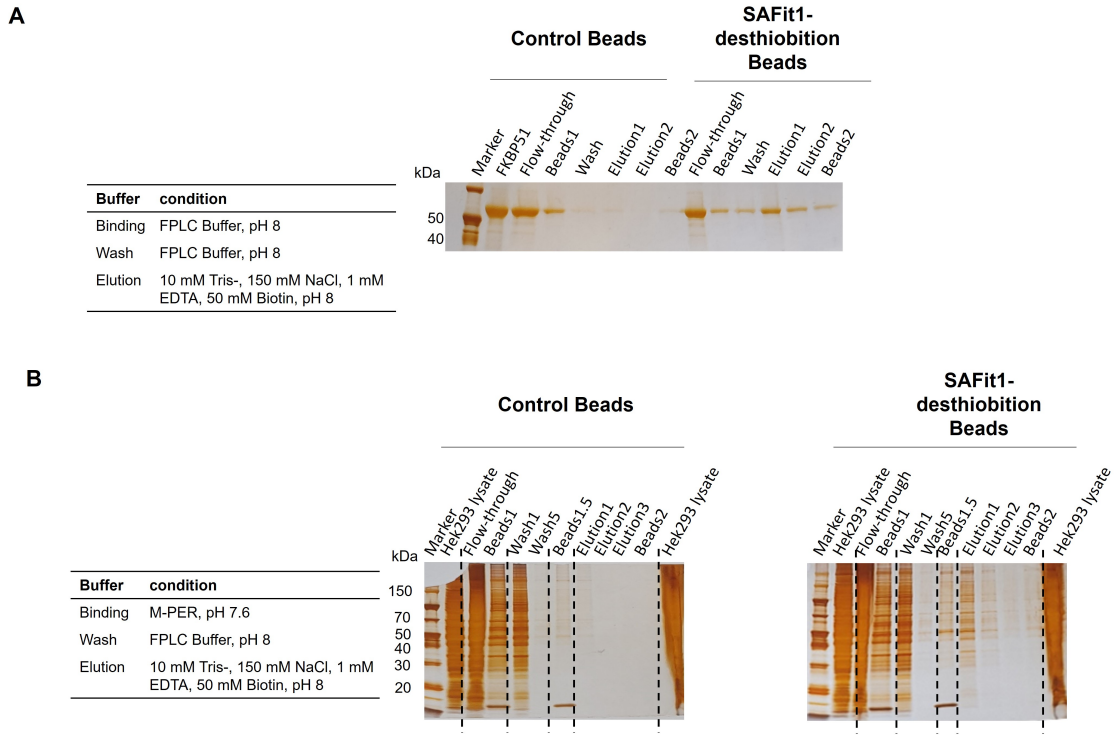


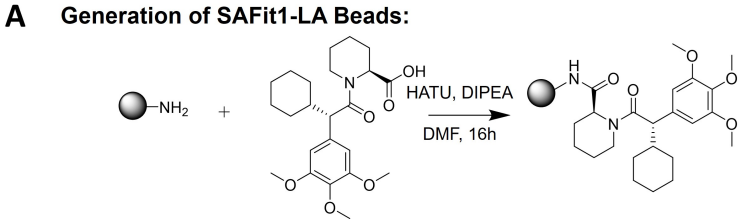
Figure 3.50: SDS-PAGE analysis of samples from the desthiobition-streptavidin affinity chromatography setup with a competitive biotin elution. A: affinity chromatography with pure FKBP51 and B: with HEK293 lysate in M-PER buffer.

Indeed elution with a biotin containing elution buffer seems feasible, comparing the

Control Beads with the SAFit1-desthiobiotin Beads in figure 3.50A for pure FKBP51 protein and 3.50B for HEK293 lysate in M-PER buffer. However, the loading of protein was not complete, since the flow-through sample for the SAFit1-desthiobiotin Beads show a intensive FKBP51 protein band in 3.50A and therefore the conditions in this affinity chromatographic set-up has to be optimized further.

Immobilization of MBa377, a low affinity SAFit1 analoga (SAFit1-LA) on amine-functionalized beads

The same principle as for the desthiobition-streptavidin system with competitive biotin elution, can be applied to SAFit1 by immobilizing a SAFit1 analog with a lower binding affinity to FKBP51 and then eluting with a SAFit1-containing elution buffer.



B

Buffer	condition
Binding	FPLC Buffer, pH 8
Wash	FPLC Buffer, pH 8
Elution	FPLC Buffer, pH 8 + 5 mM SAFit1

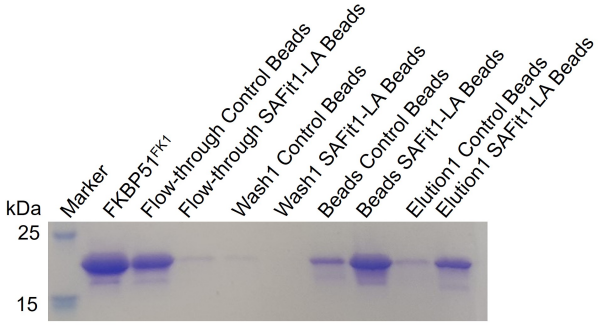


Figure 3.51: A: Immobilization of MBa377 (low binding affinity SAFit1 analog) on amine-functionalized beads by amide bond formation and generation of control beads by acetylation. B: SDS -PAGE analysis of the affinity chromatography samples with pure $FKBP^{FK1}$ and competitive SAFit1 elution.

Despite the lower affinity of the immobilized ligand towards FKBP51, $FKBP51^{FK1}$ could be loaded onto the ligand-modified beads, which is evident when comparing the Flow-through samples of the Control Beads and the SAFit1-LA Beads again. The boiled beads after washing (sample: Beads) further confirm the loading of protein onto the ligand-modified beads, as a significant protein band can be observed on the SDS-PAGE gel for the SAFit1-LA Beads sample, while the corresponding sample for the Control Beads shows

only a low intensity protein band. Moreover, elution with SAFit1 resulted in a distinct protein band in the elution fraction for the SAFit-LA Beads.

3.2.2 Application of ligand-based affinity chromatography

Ligand-based affinity chromatography is a suitable technique for enrichment of the protein of interest from a complex biological mixture or for identification of interaction partners and ligand off-targets. Therefore, the next step was to investigate the ability of the developed SAFit1-based affinity chromatography system to enrich FKBP51 from complex mixtures. The cell lysates tested were HEK293, N2a, Jurkat wild-type, and Jurkat FKBP51 knockout cell lysates. Affinity chromatography samples were analyzed using SDS-PAGE and Western Blotting to detect FKBP51 as well as Hsp90, a known interaction partner of FKBP51. In figure 3.52A the results for HEK293 cell lysate is summarized. Notably, the Beads1 samples (after washing) show an enriched protein band at approximately 50 kDa for the SAFit1 Beads, indicating that it is FKBP51. Western blot confirmed the presence of FKBP51 in the Beads1 samples, especially in contrast to the Control Beads. However, the known interaction partner HSP90 could not be enriched in the HEK293 cell lysate. However, in the N2a cell lysate, the Western Blot for Hsp90 indeed shows a signal in the Beads1 sample, but especially in the Beads2 sample (after washing and elution) (see figure 3.52B). This indicates specific isolation when compared with the corresponding Control Beads samples. However, contrary to expectation, the FKBP51 signal in the Western Blots is weak, which is true for both the Control Beads samples and the SAFit1 Beads. In addition, the SDS-PAGE gels show a high background in the elution sample of the SAFit1 Beads, suggesting that further optimization of the washing conditions is needed. Although the same conditions were used for the experiments with Jurkat lysate, the FKBP51 signals are barely visible in the Western Blots for the wild-type lysate (see figure 3.52C). The first assumption that one of the protein bands in the Beads1 sample of SAFit1 Beads is FKBP51 with a size around 50 kDa is due to the impermanence of the Western Blot. In particular, when comparing the Beads1 or Elution2 samples of the SAFit1 Beads of the Jurkat wild-type lysate with the putative Jurkat FKBP51 knockout lysate. In particular, when comparing the Beads1 or Elution2 samples of the SAFit1 Beads of the Jurkat wildtype lysate with the putative Jurkat FKBP51 knockout lysate, as the protein band pattern in both SDS-PAGE gels are nearly identical.

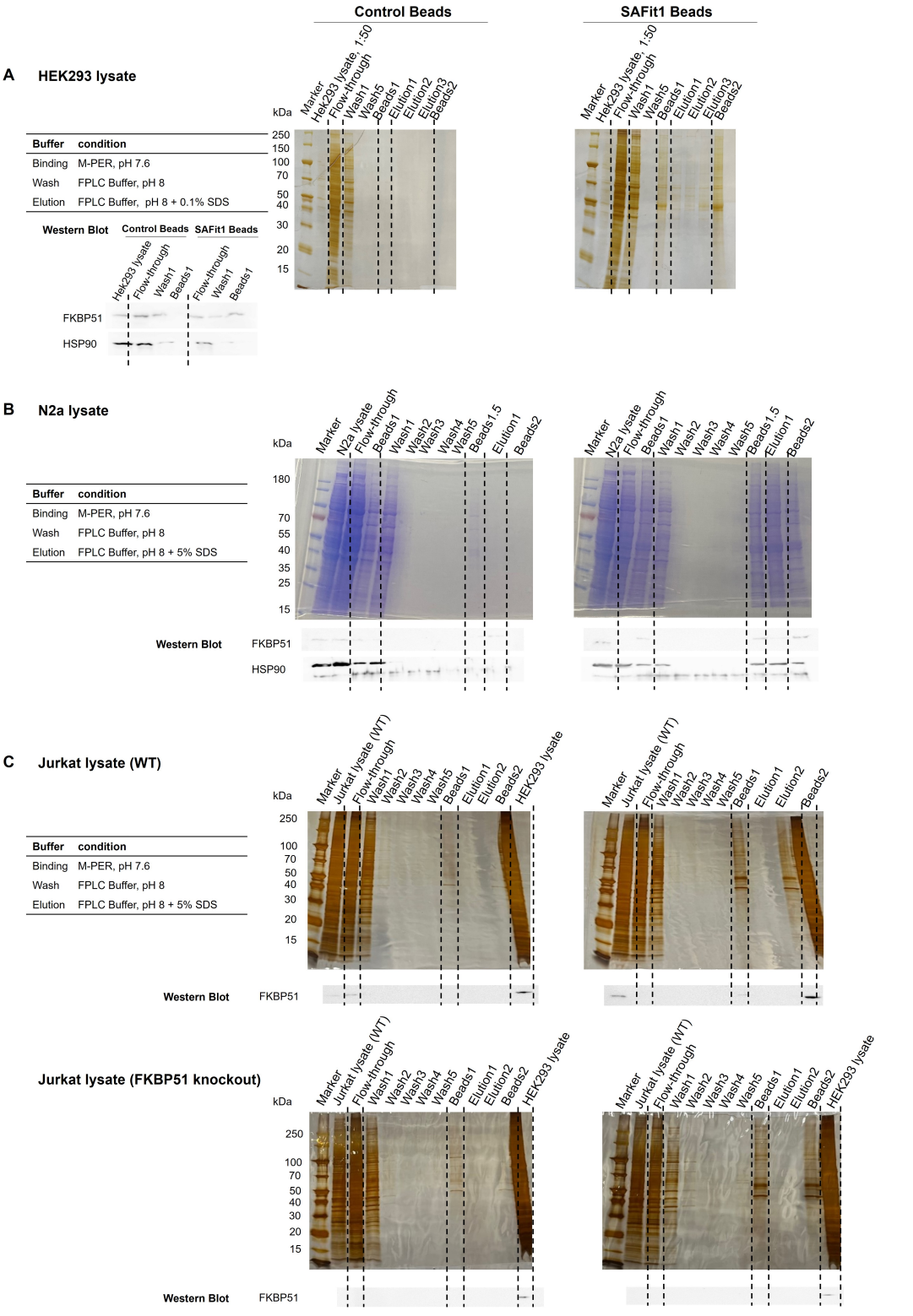
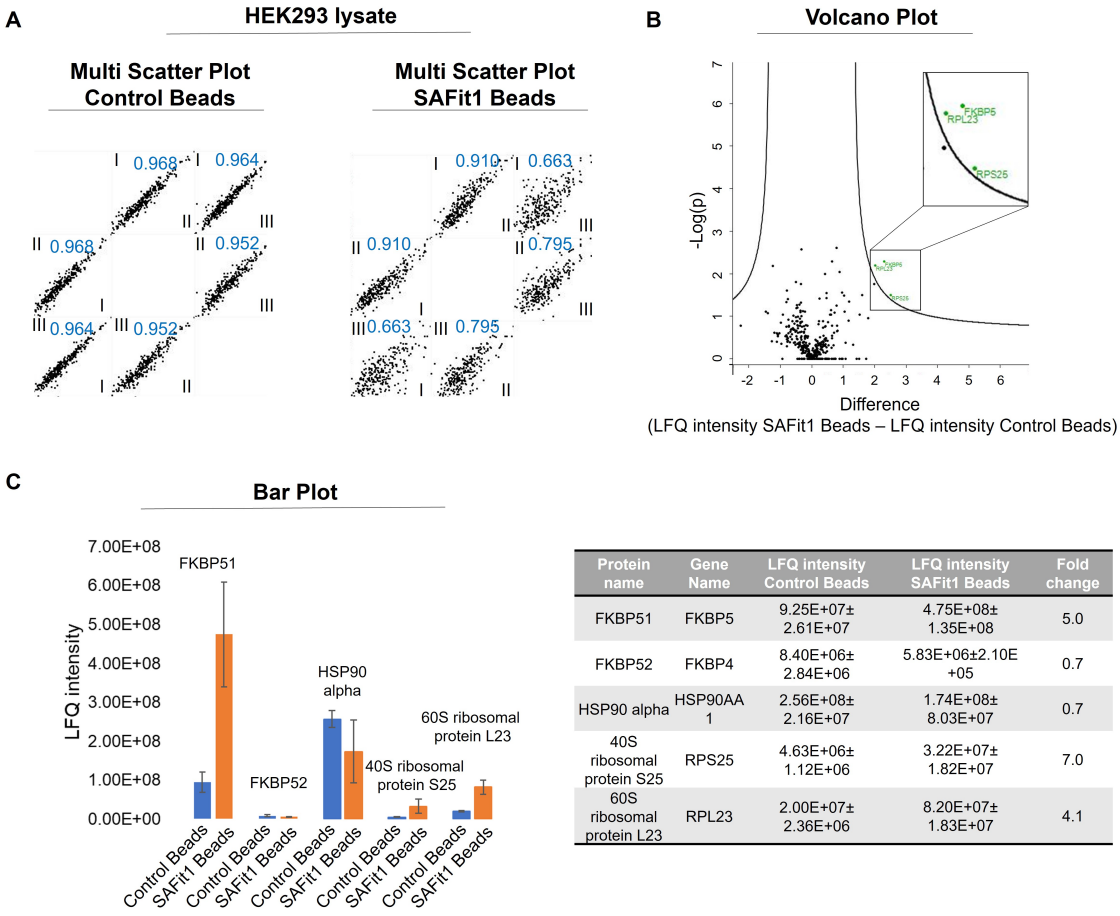


Figure 3.52: SDS-PAGE analysis and Western Blots of SAFit1-based affinity chromatography experiments with HEK293 lysate (A), N2a lysate (B), and Jurkat lysate wildtype and FKBP51 knockout (C).

SAFit1-based affinity chromatography with MS-analysis of the on-bead digestion samples

To obtain an overview of the composition of the samples after affinity chromatography, quantify the proteins of interest, and identify enriched proteins, the samples were subsequently characterized by mass spectrometry. Therefore, after loading the Beads with the complex biological mixture and washing out non-specific interaction partners, a direct on-bead digestion is performed to investigate which proteins interacted with the immobilized ligand compared to the Control Beads. HEK293 lysate was tested first, the results of the on-bead digestion are shown in figure 3.53. Three replicates were generated for each condition (SAFit1 Beads and Control Beads) and numbered with a Roman numerals. The multiscatter plots in figure 3.53A of the corresponding replicates for the Control Bead (left) and the SAFit1 Bead samples (right) show the correlation between the replicates. The Control Bead replicats have a Pearson correlation of up to 0.968, while the SAFit1 Bead replicates have a correlation coefficient of 0.663 between replicates I and III and 0.910 for replicates I and II, indicating a broad variance in replicat II. However, sample reproducibility is sufficient, and the MS data set was further analyzed. To find enriched proteins, a volcano plot was generated as shown in figure 3.53B. The proteins located in the upper right corner represent the enriched proteins, in this case FKBP51, 40S ribosomal protein S25, and 60S ribosomal protein L23. To visualize the enrichment of the proteins, a bar plot was created and shown in figure 3.53C. The LFQ (label-free quantification) intensities were plotted with regard to the sample: Control Beads compared to SAFit1 Beads. Interestingly, FKBP51 shows a higher LFQ intensity in the SAFit1 Beads sample compared to the Control Beads, while FKBP52 is only slightly present under both conditions, as expected, since SAFit1 is a selective ligand that can discriminate between FKBP51 and FKBP52. However, the known interaction partner HSP90 shows no enrichment, suggesting that the experimental setup is not suitable for identification of interaction partners, but rather for the enrichment of FKBP51. Therefore, the two proteins enriched in the SAFit1 Bead samples, 40S ribosomal protein S25 and 60S ribosomal protein L23, may be artifacts. However, further binding validation experiments need to be performed to confidently determine possible interaction partners of FKBP51 or SAFit1 off-targets.

One approach to distinguish between enriched proteins that interact primarily with FKBP51 or directly with the bound ligand SAFit1 is to use wildtype and FKBP51 knockout cell lysates. In this case, Jurkat WT (JHu WT) and Jurkat FKBP51 knockout (JHu 51 ko) cells[153] were available in the Hausch Lab and were analyzed by MS to generate a protein rank plot as shown in figure 3.54. The MS measurements were performed by Dr. Tim Heymann at the Max Planck Institue of Biochemistry. In both cell lines, cytoskeletal proteins such as actin (rank 2) and tubulin (JHu WT rank 6 and JHu 51 ko rank 7), as well as cofilin-1 (JHu WT rank 5 and JHu 51 ko rank 6) and profilin-1 (JHu WT rank 7) are widely distributed. The commonly used housekeeping gene GAPDH is also ranked 3rd in both cell lines. The glucocorticoid receptor, a known interaction partner of FKBP51, is only marginally present in both cell lines (JHu WT rank 2998 and JHu 51 ko rank 3375). FKBP family members FKBP12, FKBP52, and FKBP51 are ranked 95, 178, and 836, respectively, in the JHu WT lysate. In contrast, the ranks for FKBP52 and FKBP12 are approximately the same with ranks 201 and 215 in JHu FKBP51 knockout cells, indicating a lower abundance of FKBP12 compared to JHu wild type. Interestingly, despite the expectation that FKBP51 would not be found in JHu FKBP51 knockout cells, the protein was ranked 2028. Although FKBP51 abundance has decreased compared with the wild-type cell line, it is still detectable with MS. Thus, the initial approach of comparing wild-type cells with

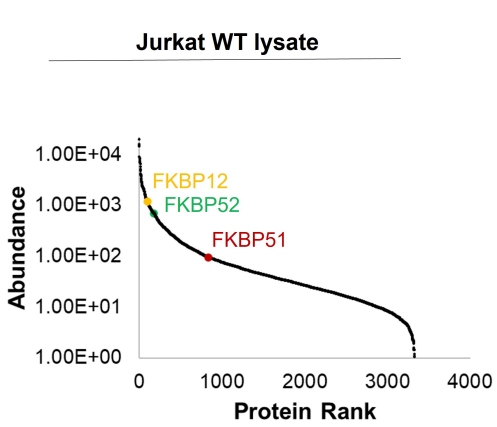


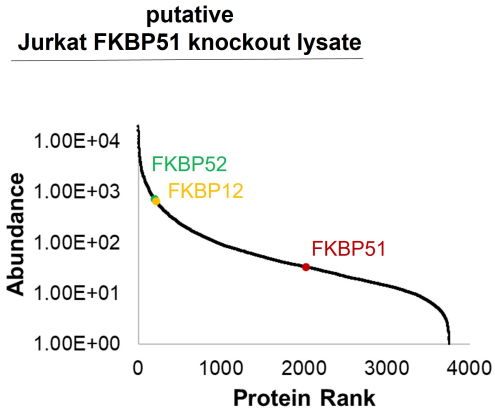
Protein name	Gene Name	LFQ intensity Control Beads	LFQ intensity SAFit1 Beads	Fold change
FKBP51	FKBP5	9.25E+07± 2.61E+07	4.75E+08± 1.35E+08	5.0
FKBP52	FKBP4	8.40E+06± 2.84E+06	5.83E+06±2.10E +05	0.7
HSP90 alpha	HSP90AA 1	2.56E+08± 2.16E+07	1.74E+08± 8.03E+07	0.7
40S ribosomal protein S25	RPS25	4.63E+06± 1.12E+06	3.22E+07± 1.82E+07	7.0
60S ribosomal protein L23	RPL23	2.00E+07± 2.36E+06	8.20E+07± 1.83E+07	4.1

Figure 3.53: SAFit1-based affinity chromatography with HEK293 lysate. After target loading and washing, on-bead digestion was performed and samples were analyzed by MS. Three replicates were generated for each condition (Control Beads and SAFit1 Beads). A: Multiscatter plot of the respective replicates. The Roman numerals indicate the replicate number, while the blue number represents the Pearson correlation. The left side shows the multiscatter plot for the Control Beads and the right side shows the multiscatter plot for the SAFit1 Beads. B: Volcano plot with the enriched proteins highlighted in green. C: Bar plot of LFQ intensities with respect to beads type for FKBP51, FKBP52, Hsp90, 40S ribosomal protein S25 and 60S ribosomal protein L23.

FKBP51 knockout cells to identify possible SAFit1 off-targets is not a reliable option in this case and hence was not performed in this work. Instead, the focus is solely on the enrichment of FKBP51.

Therefore, Jurkat wild-type and Jurkat FKBP51 knockout cell lines were loaded onto both SAFit1-based affinity matrices and fully capped Control Beads. After washing the beads to remove non-specific binders, on-bead digestion with trypsin was performed. Samples were analyzed by MS by Dr. Tim Heymann at the Max Planck Institute of Biochemistry. Figure 3.55A shows the multiscatter plots for all three replicates of each condition. Pearson correlations are all above 0.8, which can be explained by the high biological variance. Additionally, a big factor in regard of sample variance might be an incomplete digestion,





Protein Rank#	Protein name	Gene Name
1	Alpha-enolase	ENO1
2	Actin, cytoplasmic 1	ACTB
3	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
4	Heat shock cognate 71 kDa protein	HSPA8
5	Cofilin-1	CFL1
6	Tubulin beta-4B chain	TUBB
7	Profilin-1	PFN1
8	Elongation factor 1-alpha 1	EEF1A1
9	Heat shock protein HSP 90-beta	HSP90AB1
10	10 kDa heat shock protein, mitochondrial	HSPE1
95	FKBP12	FKBP1A
178	FKBP52	FKBP4
836	**FKBP51**	**FKBP5**
2998	Glucocorticoid receptor	NR3C1

Protein Rank#	Protein name	Gene Name
1	60 kDa heat shock protein, mitochondrial	HSPD1
2	Actin, cytoplasmic 1	ACTB
3	Glyceraldehyde-3-phosphate dehydrogenase	GAPDH
4	Heat shock cognate 71 kDa protein	HSPA8
5	Alpha-enolase	ENO1
6	Cofilin-1	CFL1
7	Tubulin beta-4B chain	TUBB
8	Peptidyl-prolyl cis-trans isomerase A	PPIA
9	MIF4G domain-containing protein	MIF
10	40S ribosomal protein S26	RPS26
11	Heat shock protein HSP 90-beta	HSP90AB1
201	FKBP52	FKBP4
215	FKBP12	FKBP1A
2028	**FKBP51**	**FKBP5**
3375	Glucocorticoid receptor	NR3C1

Figure 3.54: Protein ranking of Jurkat WT (left) and Jurkat FKBP51 knockout total cell lysate (right). The tables summarize the 10 most abundant proteins for both cell lines. The rankings for FKBP12, FKBP52, FKBP51, glucocorticoid receptor, and Hsp90 are also provided. FKBP51 is highlighted in red in both plots.

which can be minimized by optimization.

The Volcano plots for the JHu-51ko lysate show a higher number of significantly enriched proteins compared to the JHu-WT cell line (see figure 3.55B). Interestingly, FKBP12 is significantly enriched in both experiments, which is highlighted in the Volcano plots by a

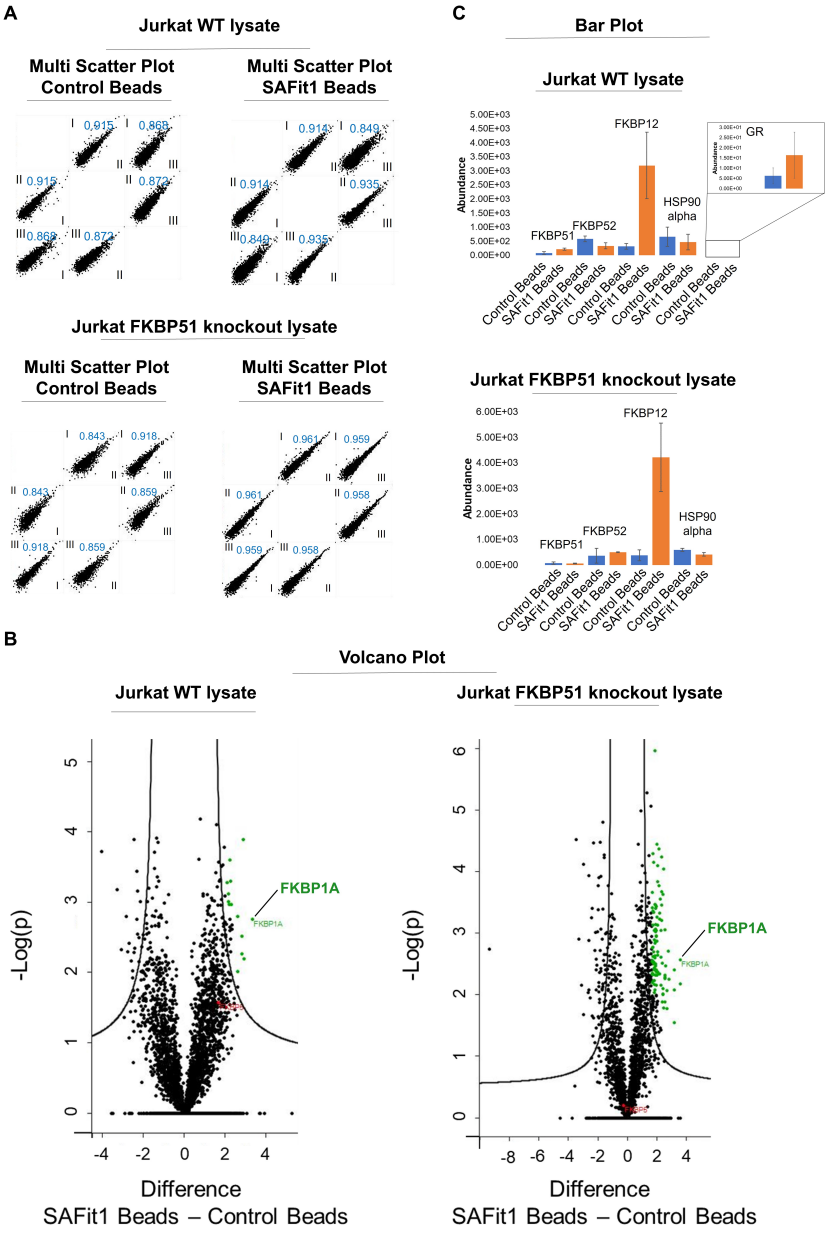


Figure 3.55: SAFit1-based affinity chromatography with JHu WT and JHu FKBP51 knockout lysate. After target loading and washing, on-bead digestion was performed and samples were analyzed by MS. Three replicates were generated for each condition. A: Multiscatter plot of the respective replicates. The Roman numerals indicate the replicate number, while the blue number represents the Pearson correlation. The upper multicatter plots show the results with JHu WT lysate and the multicatter plots below the respective replicates using JHu 51 ko cell lysate. B: Volcano plot with the enriched proteins highlighted in green. C: Bar plot of the abundance of the proteins FKBP51, FKBP52, FKBP12, Hsp90, and the glucocorticoid receptor for both cell line experiments.

green dot and labeled with the gene name FKBP1A. FKBP51, on the other hand, is not

enriched, which is highlighted in red in both Volcano plots. This is further illustrated in the bar plots, where the abundance of the proteins in the respective samples is plotted (see figure 3.55C). The FKBP family members FKBP51, FKBP52 and FKBP12 were examined in more detail, as well as the known FKBP51 interaction partners Hsp90 and the GR. For the JHu-WT cell lysate samples, only a slight increase in FKBP51 abundance is observed in the SAFit1 Beads compared to the Control Beads, while the JHu-51ko samples show low abundance in both beads samples. For FKBP52 and Hsp90, an even slightly higher abundance is observed in the Control Beads than in the SAFit1 Beads, although the relative abundance of both proteins is quite high, with FKBP52 ranking 178th in JHu WT and 201st in JHu 51ko, while Hsp90 is even among the top 11 most abundant proteins in the cell lysate samples. In contrast, GR is only slightly present in cell lysates and therefore does not show a significant increase in SAFit1 Beads. FKBP12, on the other hand, is significantly increased in both cell lines, which may be due to the high abundance of the protein, allowing it to outcompete FKBP51.

SAFit1-LA-based affinity chromatography with MS-analyzed elution samples

In addition, the ligand SAFit-LA with low binding affinity (MBa377) was immobilized and tested in an affinity chromatography experiment using Jurkat wild-type cell lysate and SAFit1 for competitive elution. The elution samples were analyzed by MS by Dr. Tim Heymann. In contrast to the other samples shown in 3.53 and 3.55, the multiscatter plot shows a relatively low correlation of replicates with Pearson correlation coefficients of 0.575 for replicate II and III and 0.579 for replicate I and II when considering the elution samples of SAFit1-LA Beads (see figure 3.56A). None of the other correlation factors are above 0.766, and therefore the standard deviation are found to be rather high in the bar plots in figure 3.56B. FKBP51 appears to be enriched, although the standard deviation is high. However, looking at the protein ranks of the elution samples in comparison, FKBP51 is ranked 701 in the Control Beads elution fraction, while FKBP51 is ranked 61 in the SAFit LA Bead elution fraction as shown in figure 3.56C. FKBP51 can be enriched and eluted in a likely native, intact state, even though the elution protocol still needs to be improved to increase the robustness.

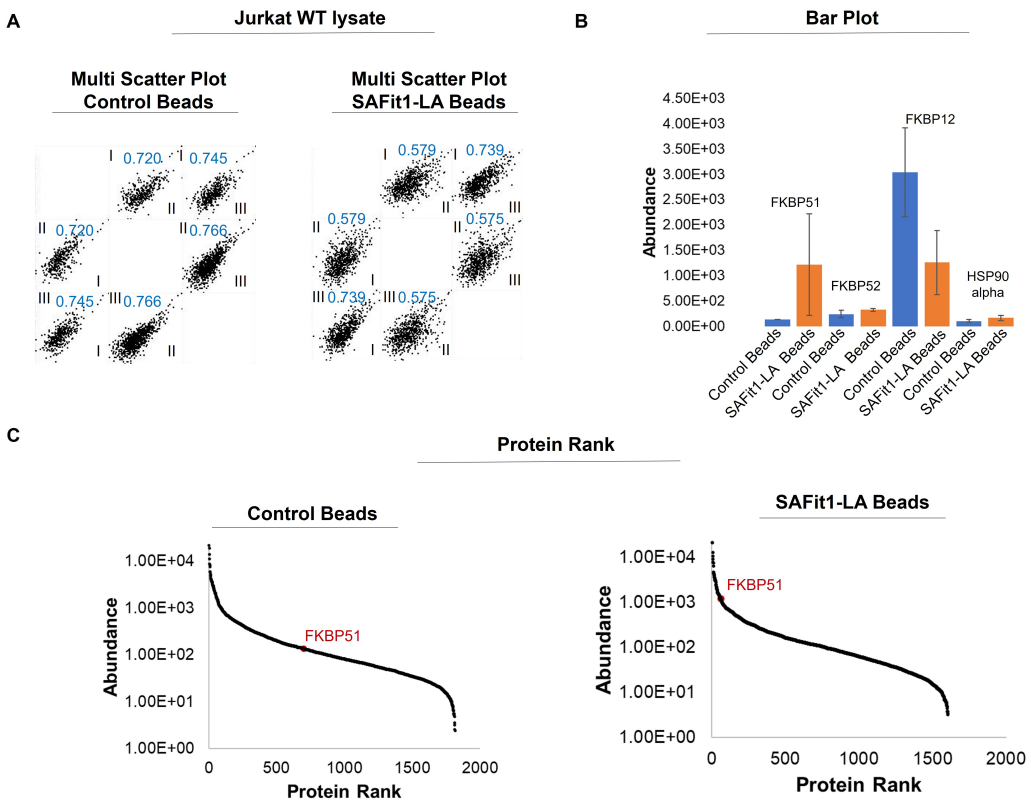


Figure 3.56: SAFit1-LA-based affinity chromatography with JHu WT lysate. After target loading and washing, elution with SAFit1 was performed and the elution samples were analyzed by MS. Three replicates were generated for each condition. A: Multiscatter plot of the respective replicates. The Roman numerals indicate the replicate number, while the blue number represents the Pearson correlation. B: Bar plot of the abundance of the proteins FKBP51, FKBP52, FKBP12 and Hsp90 in the eution samples of SAFit1-LA Beads and Control Beads. C: Protein Rank of elution fraction of Control Beads (left) and SAFit1 Beads (right). FKBP51 is highlighted in red.

4 Conclusion & Outlook

In this work, 14 F67-out stabilized protein variants were generated by 5 different protein engineering approaches.

Table 4.1: Summarized K_d values of a high and low affinity tracer and the stabilized FKBP51^{FK1} variants in comparison to the corresponding control proteins.

Variant	Experiment	$K_{d,\ high\ affinity\ tracer}$	$K_{d,\ low\ affinity\ tracer}$
WT	Rational design	6±0.7	-
		5±0.9	-
	Photocrosslinking	2.8±0.2	-
	Click Chemistry	2.3±0.4	-
	Cysteine-directed	-	4122±209
		-	5181±338
recombinant noCys	Rational&Lactam	-	6741±873
synthetic noCys	Rational&Lactam	-	7443±1850
F67Y	Rational	0.7±0.2	-
	Photocrosslinking	0.3±0.1	-
	Cysteine-directed	-	433±23
F67G	Rational	0.7±0.2	-
F67Y/K58D	Rational	26±5	-
F67Y/K58N	Rational	5±1	-
F67Amb/G64S	Rational	-	285±61
F67Bpa -UV	Photocrosslinking	0.5±0.1	-
F67Bpa +UV	Photocrosslinking	0.2±0.1	-
F67pAzF/G13C	Click-Chemistry	0.2±0.1	-
F67pAzF/G13C + DBCO-MI	Click-Chemistry	0.5±0.1	-
F67E/K58 (i, i+9)	Lactam	-	685±88
F67E/K60 (i, i+7)	Lactam	-	1103±151
F67E/K60Orn (i, i+7)	Lactam	-	172±27
F67E/K60Dab (i, i+7)	Lactam	-	220±12
F67C/K58C, red.	Cysteine-directed	-	3107±439
		-	2299±202
F67C/K58C, ox.	Cysteine-directed	-	890±73
F67C/K58C, DBA CL	Cysteine-directed	-	106±6
F67C/K60C, red.	Cysteine-directed	-	4437±303
		-	1695±202
F67C/K60C, ox.	Cysteine-directed	-	206±14
F67C/K60C, DBA CL	Cysteine-directed	-	350±26

Table 4.1 summarizes the determined K_d for high and low affinity tracers with the

respective stabilized protein variants as well as the control proteins. The first variants developed were tested with a high affinity tracer. However, with increasing binding affinity, the FP assay reaches its limits. For this reason, a low affinity tracer was synthesized to facilitate the detection of improved binding affinity. The highest binding affinity was detected for the DBA crosslinked F67C/K58C variant, but since DBA crosslinking resulted in heterogeneous samples, the winning variant was declared to be F67C/K60C, ox. with a disulfide bridge. This variant proved to be a suitable screening tool in the search for new lead structures for FKBP51. The initial screening with the TRABITA library resulted in 6 positive hits that would have remained undetected without the stabilized FKBP51^{FK1} screening variant. In addition, a second round of screening identified one more fragment hit. This marks the beginning of the search for a new lead structure for FKBP51, as the previously established ligands (iFit series, including SAFit1 and 2) do not have sufficient physicochemical properties and ligand efficiency for small molecule drugs. The chemical structures of the fragment hits are shown in figure 4.1A.

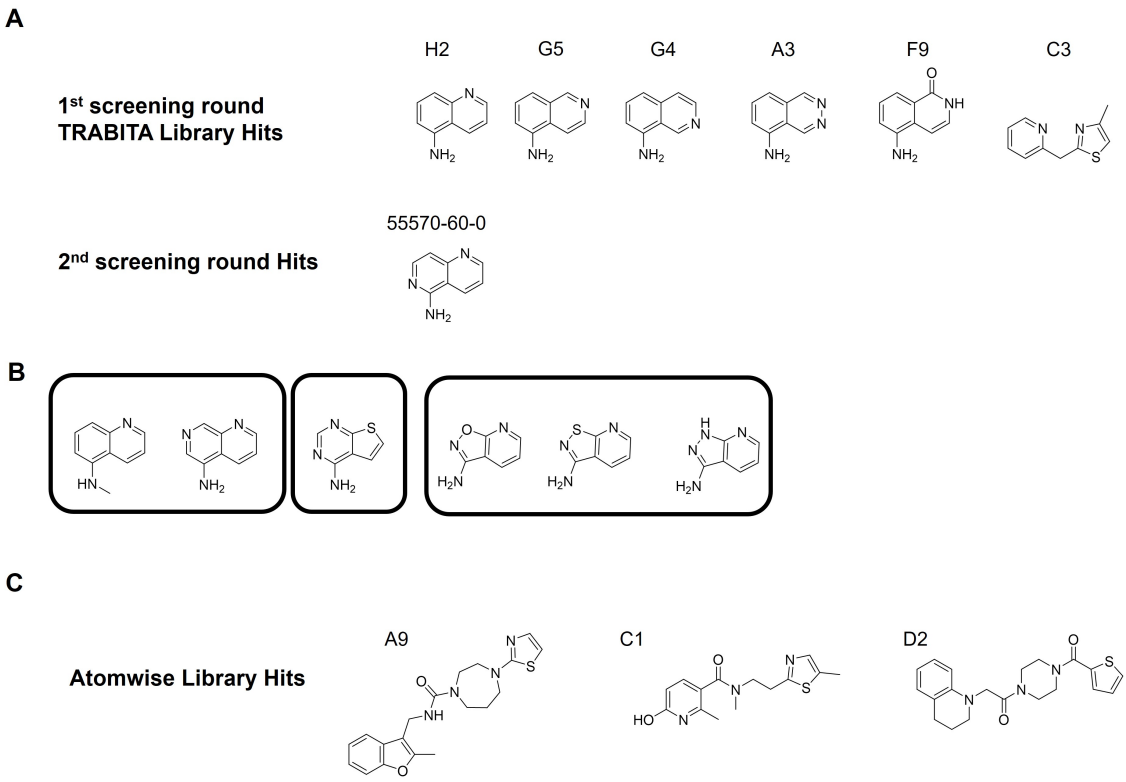


Figure 4.1: A: chemical structures of the yielded hits in two fragment-based screenings with F67C/K60C^{ox}. B: examples of new to test fragments based on the initial hits. C: chemical structures of Atomwise library hits A9, C1 and D2.

The in the screening discovered structures need to be further investigated by testing additional derivatives based on the hits obtained in this work. Possible fragment structures to be tested are summarized in figure 4.1B. Moreover, fragment C3 has a different structure than the other hits from the initial TRABITA library screening, so this structure may be explored in the future as well. In addition, the Atomwise compound library was also tested, resulting in the identification of three hits. The chemical structures of the hits are shown in 4.1C. It is noticeable that the A9 and C1 structures have a thiazole moiety as fragment

C3 from the TRABITA library.

In conclusion, the newly developed screening tool can be used to screen fragment and compound libraries with higher detection sensitivity. Orthogonal screening methods such as crystallographic fragment screening or saturation transfer difference nuclear magnetic resonance spectroscopy (STD-NMR) are conceivable techniques that can be used next to identify a new ligand scaffold for FKBP51.

The second project of this work was to develop a ligand-based affinity chromatography matrix to enrich FKBP51 from complex biological mixtures as well as to identify FKBP51 interaction partners and/or SAFit1 off-targets. Therefore, four ligands were immobilized on polymethacrylic or agarose beads. Two main problems were encountered during the initial experiments to enrich FKBP51 in complex biological mixtures: First, capturing FKBP51 on the ligand-based beads from cell lysates and second, eluting the specifically bound proteins from the beads. Cell lysates in NETN lysate buffer have been shown to contain substances that interfere with the interaction of the immobilized ligand with the protein. In contrast, the commercially available M-PER buffer is a suitable lysate buffer and is compatible with the affinity chromatography system. To solve the elution problem, two options are available. Either harsh elution conditions are applied, such as the use of a strong detergent and/or heat, or a SAFit1 analog with low affinity is immobilized so that elution can be performed with SAFit1. The initial affinity chromatography experiments succeeded in enriching FKBP51 from HEK293 lysate, but the background is still quite high. Examination of Jurkat wild-type (JHu WT) and Jurkat FKBP51 knockout cell lysates (JHu 51ko) revealed that FKBP51 was relatively low populated in the JHu WT cell line, and although no FKBP51 signal was expected in the FKBP51 knockout cell line, the protein could still be detected by MS. Thus, the first steps for a FKBP51 ligand-based affinity chromatography system are set, but still the protocol needs to be further optimized in terms of washing conditions and reproducibility.

5 Experimental section

5.1 Site directed mutagenesis

To introduce mutation in the protein sequence, the SPRINP-PCR (Single-Primer Reactions IN Parallel polymerase chain reaction) apoproach is used. The method is based on performing two spatially separated PCR reactions, one with the forward primer and the other with the reverse primer. The obtained PCR products were combined and parental plasmid strands were digested before transformation in *E.coli*. The protocol was adjusted based on literature.[154]

5.1.1 SPRINP-PCR

The used primer for introducing the respective mutations are summarized in the following table:

Mutation(s)	Primer
F67G	fw AGCAACGGCAAAAAA GGC GATAGCAGCCATGAT rv ATCATGGCTGCTATC GCC TTTTTTGCCGTTGCT
F67Y/K58D	fw GTGTATGTGCACTAT GAT GGCAAACTGAGC rv GCTCAGTTTGCC ATC ATAGTGCACATACAC
F67Y/K58N	fw GTGTATGTGCACTAT AAT GGCAAACTGAGC rv GCTCAGTTTGCC AAT ATAGTGCACATACAC
F67TAG	fw CTGAGCAACGGCAAAAAA TAG GATAGCAGCCATGATCGT rv ACGATCATGGCTGCTATC CTA TTTTTTGCCGTTGCTCAG
F67TAG/G13C	fw CTGAGCAACGGCAAAAAA TAG GATAGCAGCCATGATCGT rv ACGATCATGGCTGCTATC CTA TTTTTTGCCGTTGCTCAG
F67C/K58C	fw GTGTATGTGCACTAT TGC GGCAAACTGAGCAACGGCAAAAAATGC rv GCATTTTTTGCCGTTGCTCAGTTTGCC GCA ATAGTGCACATACAC
F67C/K60C	fw GTGTATGTGCACTATAAAGGC TGC CTGAGCAACGGCAAAAAATGC rv GCATTTTTTGCCGTTGCTCAG GCA GCCTTTATAGTGCACATACAC
F67C/K65C	fw GGCAAACTGAGCAACGGC TGC AAATGCGATAGCAGCCATGAT rv ATCATGGCTGCTATCGCATTT GCA GCCGTTGCTCAGTTTGCC
F67C/S69C	fw GAGCAACGGCAAAAAATGCGAT TGC AGCCATGATCGTAATGAAC rv GTTCATTACGATCATGGCT GCA ATCGCATTTTTTGCCGTTGCTC
F67C/G13C	fw CACAGAGAACAGATTGGTGGT TGC GCCCCTGCAACCGTTACC rv GGTAACGGTTGCAGGGGC GCA ACCACCAATCTGTTCTCTGTG

Additionally the template plasmids are listed in the following, with the template plasmids having an internal Hausch Lab number starting with HP (Hausch Plasmid). HP794, HP795 and HP1202 have a pET-14b-vector (selection marker: Ampicillin (100 µg/mL)) coding for FKBP51$^{FK1,\ 16\text{-}140}$, whereas HP683 and HP743 are based on a pET-30b-vector (selection

marker: Kanamycin (50 µg/mL)) coding for $FKBP51^{FK1,\ 1\text{-}140}$. All plasmids abolished the wildtype cysteines (C103A/C107I). HP1202 additionally has the F67C mutation and codes for the $FKBP51^{FK1,\ 16\text{-}140}$ noCys variant.

Mutation(s)	Template
F67G	HP794
F67Y/K58D	HP795
F67Y/K58N	HP683
F67TAG	HP683
F67TAG/G13C	HP743
F67C/K58C	HP1202
F67C/K60C	HP1202
F67C/K65C	HP794
F67C/S69C	HP1202
F67C/G13C	HP 1202

For SPRINP-PCR, first, the PCR reaction mixtures were prepared. Therefore, 10 µL Q5 reaction buffer (B9028A, NEB), 10 µL GC Enhancer buffer (B9027S, NEB), 1 µL deoxynucleotide solution mix (N0447S, NEB, 200 µM), 1 µL template (80n g), 24 µL sterile MQ-water, 0.5 µL Q5-DNA high fidelity polymerase (M0491S, NEB, 2000 U/mL) and 2.5 µL forward or reverse primer were mixed together. The PCR reaction was conducted with an initial denaturation step at 95 °C for 5min, afterwards the PCR cycle of denaturation (95 °C, 30 s), annealing (60 °C, 1 min) and extension (72 °C, 1 min) was repeated 30 times. 40 µL of each reaction mixtures (forward primer and reverse primer PCR products) were combined and treated under the following temperature programm:

T/ °C	t/ min
95	5
90	1
80	1
70	0.5
60	0.5
50	0.5
40	0.5
37	0.5

0.5 µL DpnI (R0176S, NEB, 20 000 U/mL) were added for digestion of the methylated, parental strands. The mixture was incubated at 37 °C for 5 h before transformation in *E.coli*.

5.1.2 Transformation

10 ng DNA (prepared as described in section 5.1.3) were mixed with 70 µL competent DH5α cells and incubated for 30 min on ice. The heat shock was performed afterwards by incubating the cells at 42 °C for 45 s. 250 µL SOC medium (0.5 % yeat extract, 2 % tryptone, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$ and 20 mM Glucose) was added and the cells were incubated for 1-2 h at 37,°C. 100 µL of the cell suspension was plated on agar plates with the respective antibiotic. The plate was incubated at 37 °C overnight.

5.1.3 Plasmid Preparation

A colony from the agar plate was picked and inoculated in 5-6 mL LB medium containing the respective antibiotic and incubated overnight at 37 °C. The cells were pelleted and the FastGene Plasmid Mini Kit from Nippogenetics was used for plasmid preparations as described by the manufacturer. The plasmid concentration was determined photometrically and the sequencing was conducted by the Microsynth AG. The plasmid with the correct sequence was secured by preapring the respective glycerol stock. Therefore, the plasmid was once again transformed into competent *E.coli* cells (DH5α or BL21DE3) and 750 µL of the bacterial culture was mixed with 750 µL sterile glycerol and stored at -80 °C.

5.2 Protein expression

For protein expression in *E.coli*, the plasmid coding for the protein of interest has to be transformed into an appropriate bacterial strain, which were chemically competent.

5.2.1 Generation of chemically competent cells

Inoculate 5 mL of LB medium with the respective bacterial strain and incubate at 37 °C overnight. This preculture was used to innoculate 250 mL of LB medium and incubate at 37 °C until an OD_{600} of 0.5-0.6 was reached. Afterwards the bacterial cells were centrifuged at 4 °C and 2000 g for 15 min. The cell pellet was carefully resuspended in 50 mL TBF1 buffer (100 mM $CaCl_2$ in MQ water) and centrifuged again at 4 °C and 2000 g for 15 min. The pellet was now resuspended in 15 mL TBF2 buffer (100 mM $CaCl_2$, 15 % glycerin in MQ buffer). The suspension was aliquoted and directly frozen in liquid nitrogen. The aliquots were stored at -80 °C. The cells may now be used for transformation. The protocol for the transformation follows as explained in section 5.1.2.

5.2.2 Subsequential transformation for amber codon suppression

In case of amber codon suppression, two plasmids have to be transformed into a competent bacterial strain. Therefore, the plasmid encoding the tRNA/aminoacyl-tRNA synthetase pair (selection marker: Chloramphenicol (25 µg/mL)) was first transformed into a chemically competent BL21(DE3) strain (commercially purchased by ThermoFisher, EC0114) according to protocol in section 5.1.2. Subsequently the transformed bacterial cells were made chemically competent again, by following the protocol in section 5.2.1. Now, the second plasmid, coding for the protein of interest with an respective TAG mutation was transformed into the bacterial cells, again following the protocol in section 5.1.2.

5.2.3 Protein expression

After the transformation of the plasmid(s), a preculture was prepared by adding the transformed cells to a flask containing 50-200 mL of LB media and the respective antibiotic(s) (Ampicilin: 100 µg/mL, Kanamycin: 50 µg/mL and Chloramphenicol: 25 µg/mL). The preculture was incubated overnight at 37°C and with 150 rpm, before inoculating 1-4 L of LB media to an OD_{600} of 0.1 with the preculture. As soon as a OD_{600} of approximately 0.5 was reached, the culture was induced by the addition of 1 mM isopropyl 1-thio-D-galactopyranoside (IPTG, Roth, 206-703-0) and incubated overnight at 25 °C and with

150 rpm. However, in case of amber suppression, 0.2 % L-arabinose (L-Ara, Roth, 226-214-6) and 1 mM of the respective unnatural amino acid (para-azidophenylalanine (abcr) or para-benzoylphenylalanine (Sigma)) was added to the bacterial growth medium. The culture was incubated for 5 h at 37 °C. After protein expression, the cells were harvested by centrifugation (10 000 g, 10 min, 4 °C) and stored at -20 °C until lysis.

5.2.4 Lysis of bacterial cells

The cells were resuspended in 15 mL lysis buffer (20 mM HEPES pH=8, 300 mM NaCl) and 2 mM ethylenediaminetetraacetic acid (EDTA, Roth, 205-358-3), 1 mM phenylmethylsulfonyl fluoride (PMSF, Roth, 206-350-2), and 1 mg/mL lysozyme (Roth, 235-747-3) were added. The mixture was incubated for 2 h at 4 °C on a rolling device. Afterwards 0.1 mg/mL DNaseI (NEB, M0303S) and 100 mM $MgSO_4$ (Roth, 231-298-2) were added, and the tube was shaken at 4 °C on a rolling device for another hour. Afterwards, the samples were sonicated 3x 3 min. Cellular debris was removed by centrifugation (13 000 g, 15 min, 4 °C).

5.2.5 Purification of proteins

IMAC & desalting

The supernatant from the last step of lysis was used in metal ion affinity chromatography (IMAC) to obtain the pure, His-tagged protein of interest (Equilibration Buffer: 20 mM HEPES pH 8, 500 mM NaCl and if necessary 1 mM TCEP; Washing Buffer: 20 mM HEPES pH 8, 500 mM NaCl, 10 mM imidazole and if necessary 1 mM TCEP; Elution Buffer: 20 mM HEPES pH 8, 500 mM NaCl, 300 mM imidazole and if necessary 1 mM TCEP). To remove the imidazole (Roth, 206-019-2), the fractions with the highest protein concentration (measured photometrically at A_{280}) were combined and desalted using the HiPrep 26/10 Desalting column on an Äkta Pure Protein Purification System (Cytiva) at a flow rate of 5 mL/min with the Equilibration Buffer (20 mM HEPES pH 8, 500 mM NaCl). For small volumes, the desalting was done by utilizing ready-to-use PD-10 columns (Cytiva 17-0851-01).

Removal of the protein tag

The His- and/or Sumo-Tags were removed by the addition of Sumo protease (1:100 (m/m)) or TEV protease (1:20 (m/m)) and incubation overnight at 4°C on a rolling device. To remove the cleaved His- and Sumo-Tag from the sample, a reverse IMAC was performed, collecting the supernatant.

SEC

Additionally, the protein was further purified by SEC (size-exclusion chromatography) on a Äkta Pure Protein Purification System (Cytiva) using the HiLoad 16/600 Superdex 75 pg column with a flow rate of 0.5 mL/min (running buffer: 20 mM HEPES pH 8, 20 mM NaCl). The fractions containing the pure protein (confirmed by SDS-PAGE) were concentrated using Amicon Ultra 2 mL Centrifugal Filter (Sigma, UFC201024). The proteins were aliquoted and stored at -80°C.

WCX

The weak cation exchange chromatography was performed on a 1260 Infinity II Bio-Inert LC System fitted with a Agilent Bio MAb, NPS, 4.6x250 mm, PK column. Four solvents were freshly prepared: A= MQ-water, B= NaCl (3350 mM), C= MES (28.5 mM) and D= MES - Na (3.55 mM). The start condition was 34.8 % A, 3.0 % B, 31.2 % C and 31.0 %D and within 20 min the end condition of 10.6 % A, 29.9 % B, 31.6 % C and 27.9 % D was reached. The proteins to be purified were diluted in the starting condition to a 1 mg/mL solution and per run 20-100 µL were injected. The fractions containing the desired protein (confirmed by SDS-PAGE and/or MS) were concentrated using Amicon Ultra 2 mL Centrifugal Filter (Sigma, UFC201024). The protein was aliquoted and stored at -80°C.

5.3 Biophysical methods

5.3.1 Fluorescence polarization assay

To determine the binding affinity of a tracer and the respective protein variants, fluorescence polarization binding assays were performed.[148]

Active site titration

In a 384-well assay plate (black, flat bottom, low binding) a serial dilution of the investigational protein variants were prepared in FP-Assay buffer (20 mM HEPES pH 8, 150 mM NaCl, 0.015 % Triton X-100). Therefore, 60 µL of the starting protein solution was pipetted in well A. The rows B-P were filled with 30 µL FP-Assay buffer and a 1:2 serial dilution was carried out from row A to O with a carryover of 30 µL. Then, 10 µL tracer solution was added in each well. In case of an active site titration the tracer was added at an end concentration of 30 nM. All proteins were tested in a technical triplicate. The plates were incubated for 30 min at room temperature and the fluorescence polarization was measured on a plate reader (Gain 75, excitation wavelength 485 nm and emission wavelength 535 nm). The fluorescence polarization values were normalized with respect to the highest signal. The 4-Parameter fitting equation was used to determine the EC_{50} from the active site titration, defined here as the total concentration at which free and bound tracer were present in equal quantities:

$$y = min + \frac{(max - min)}{1 + (\frac{x}{EC_{50}})^{\text{-Hillslope}}} \tag{5.1}$$

Min represents the lower plateau, max the upper plateau and the EC_{50} is the half maximal effective concentration.

Binding curves

For the binding curves, the protein concentration was adjusted according to the active site titration. The protocol was carried out such as for an active site titration, but with a tracer end concentration of 1 nM. The binding affinity was determined according to the following equation with $c_{AST,protein}$ being the concentration of the protein determined by

active site titration, $c_{UV,protein}$ is the concentration of the protein determined by UV and c_{Tracer} the concentration of tracer. EC_{50} is the half maximal effective concentration.

$$c_{\text{AST, protein}} = \frac{c_{\text{Tracer}} \cdot c_{\text{UV, protein}}}{2 \cdot EC_{50}} \tag{5.2}$$

To determine the binding affinity of the tracer with the respective protein variants, the recorded binding curves were fitted by a model, which was described by Wang et. al in 1992.[155]

$$Y = \frac{A}{E} \cdot 0.5 \cdot \left(X + E + \frac{1}{K} - \sqrt{\left(\frac{X + E + 1}{K} \right)^2 - 4EX} \right) + B \tag{5.3}$$

Y is the polarization, X the concentration of the protein, E the concentration of the ligand, K_d the dissociation constant, A the amplitude and B the bottom plateau.

Competitive FP-Assay

For the competitive FP-Assay a serial dilution of the ligand or fragment in the 384 plate was prepared. Therefore, 30 µL of the initial ligand solution was added in well A and 30 µL of FP-Assay buffer were filled from B-P. The serial 1:2 dilution was carried out in the plate with a carryover of 30 µL. Subsequently 10 µL of master mix was added consisting of a protein and tracer mixture in FP-Assay buffer. The master mix for the FKN22 assay contained a tracer end concentration of 1 nM and a protein end concentration of 5 nM. To determine the dissociation constant of the examined ligand the binding curves were fitted with the following equation which was described 1995 by WANG ET AL. for competitive binding of two different ligands to the protein.[156]

$$Y = A \cdot \frac{2 \cdot \sqrt{a^2 - 3b} \cdot cos\frac{\theta}{3} - a}{3 \cdot K1 + (2 \cdot \sqrt{a^2 - 3b} \cdot cos\frac{\theta}{3} - a)} + B \tag{5.4}$$

$$\begin{aligned}
&with \\
&\theta = arccos\frac{-2a^3 + 9ab - 27c}{2 \cdot \sqrt{(a^2 - 3b)^3}} \\
&a = K1 + K2 + L0 + X - P0 \\
&b = K2 \cdot (L0 - P0) + K1 \cdot (X - P0) + K1 \cdot K2 \\
&c = -K1 \cdot K2 \cdot P0
\end{aligned}$$

Y is the polarization, X the concentration of the ligand, K1 the dissociation constant of the tracer, whereas K2 is the dissociation constant of the ligand. A is the amplitude, B the bottom plateau, L0 the tracer concentration and P0 the protein concentration.

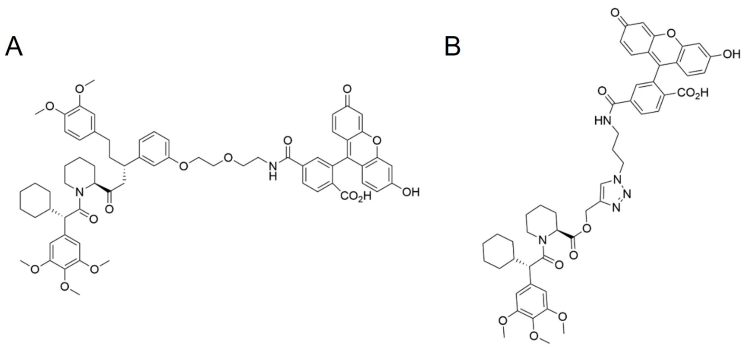


Figure 5.1: A: SAFit1 based tracer coupled to fluorescein. B: Low affinity tracer SAFit1-LA coupled to fluorescein.

The SAFit1-based fluorescein tracer (see figure 5.1) was kindly provided by Dr. Andreas Voll. The low affinity tracer (see figure 5.1) was synthesized within this work as described in the following:

4-((3-(4-((((S)-1-((S)-2-Cyclohexyl-2-(3,4,5-trimethoxyphenyl)acetyl)piperidine-2-carbonyl)oxy)methyl)-1H-1,2,3-triazol-1-yl)propyl)carbamoyl)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic acid (figure 5.1B)

4-((3-Azidopropyl)carbamoyl)-2-(6-hydroxy-3-oxo-3H-xanthen-9-yl)benzoic acid (6-Carboxy fluorescein azide, Roth) (10.01 mg, 22 µmol, 1.00 eq.), **7** (9.99 mg, 22 µmol, 1 eq.), sodium ascorbat (4.33 mg, 22 µmol, 1 eq.) and cooper sulfate (5.45 mg, 22 µmol, 1 eq.) were dissolved in 1:1:1 tert-butanol:water:DMSO and stirred for 16 h at room temperature under argon atmosphere. The reaction solution was afterwards extracted with ethyl acetate (3 x 10 mL) and washed once with brine (50 mL) before combining and drying the organic layers with $MgSO_4$. After solvent removal under reduced pressure, the crude product was purified by RP-HPLC on an Interchim puriFlash 5.250 system fitted with a preperative LC column (Luna-5-C182-250X2, C18). The run was performed with a flow rate of 16 mL/min and a 40-70 %B within 6 CV (column volume) gradient, where solvent A was 0.1 % TFA in water and solvent B was 0.1 % TFA in acetonitrile. The fractions with the pure product were combined and the solvent was removed under reduced pressure and lyophilized. A yellow solid was obtained.
Yield: 14.8 mg (16.17 µmol, 74 %)
LC-MS (5 – 100 % solvent B, 2 min) R_t = 2.089 min, purity (220 nm): 99 %, purity (254 nm): 87 % m/z: calculated = 916.38 [M+H]+, found = 916.40 [M+H]+.
HRMS (ESI): calculated = 916.37635[M+H]+, found = 916.37690 [M+H]+, error [ppm] = 0.6.

5.3.2 Thermal Shift Assay

1 µL of the fragment or ligand (10 mM stock in DMSO) was mixed with 20 µL protein to an protein end concentration of 60 µM. The samples were filled in a capillary (Prometheus Standard Capillaries (400 Count), Nanotemper, SKU: PR-C002) and the melting curves were measured on a Prometheus NT.48 device.

5.4 Protein characterization

5.4.1 SDS-PAGE

For the SDS-PAGE (Sodium dodecylsulfate polyacrylamide gel electrophoresis) gels of the affinity chromatogaphy experiments, commercially available gradient gels were used (Invitrogen™ Novex™ WedgeWell™, 4-20 %). For the other gels 14 % gels were preapred in empty gel cassettes, mini, 1.0 mm from Thermo Fisher. The running gel for 2 gels was prepared by mixing 6.65 mL acrylamide/bisacrylamide 30 (Roth, A124.2), 5.34 mL 1 M Tris pH=8.8, 2 mL MQ-water, 140 µL SDS 10 %, 84 µL APS 10 % (ammonium peroxide sulfate, Roth, 231-786-5) and 17 µL TEMED (tetramethylethylenediamine, Roth, 203-744-6). The stacking gel for 3 gels consists of 5 mL stacking gel buffer (44 mL acrylamide/bisacrylamide 30 (Roth, A124.2), 178.5 mL MQ-water, 25 mL 0.6 M Tris pH 6.8, 2.5 mL SDS 10 %), 25 µL APS and 10 µL TEMED. The sample preparation was conducted by mixing 15 µL of protein solution (concentration range: 20-100 µM) with 5 µL 4x Lämmli buffer (400 mM Tris pH 6.8, 4 % SDS, 50 % glycerol, 0.01 % bromphenol and 5 % β-mercaptoethanol) and denaturing the samples at 95 °C for 10 min. Subsequently the samples were loaded on the gel and a voltage of 200 V was applied. After the run, the gels were stained either by Coomassie stain or by silver staining as described in the next sections. As protein ladder the PageRuler^{TM} Prestained Protein Ladder, 10-180 kDa (Thermofisher, 26616) or the PageRule^{TM} unstained Protein Ladder, Broad Range 5-250 kDa (Thermofisher, 26630) was used.

Coomassie staining

The gel was mixed with Coomassie stain (40 % (v/v) ethanol, 10 % (v/v) acetic acid in water with 1 g/L Coomassie Brilliant Blue R250) for at least 30 min. Afterwards the gels were destained with destaining solution (40 % (v/v) ethanol, 10 % (v/v) acetic acid in water) for 30 min. If necessary the destaining step was repeated with water or destaining solution. The gels can be stored in MQ-water for several days or dried.

Silver staining

After the SDS-PAGE run the gels were incubated in fixation solution (30 % (v/v) ethoanol, 10 % (v/v) acetic acid) over night (16-18 h). The gels were washed 2x with 20 % (v/v) ethanol for 10 min and afterwards 2x with MQ-water for 10 min. Subsequently, the gels were soaked for 1 min in 0.8 mM $Na_2S_2O_3$ (sodium thiosulfate, Sigma, 217263) solution before rinsing them 2x in MQ-water for 1 min. For impregnation the gels were incubated in 12 mM $AgNO_3$ (Roth, 7908.1) solution for at least 20 min and up to 2 h. At this step the gel might turn yellow. To visualize the protein bands the gels were next rinsed in water and put in development solution (3 % (w/v) K_2CO_3 (Alfa Aesar, 11467967), 250 µL formalin/L (37 % ROTIPURAN, Roth, 4979.1)and 125 µL 10 % (w/v) $Na_2S_2O_3$) until the desired staining degree was reached. The gels were then transfered to the stop solution (4 % (v/v) Tris, 2 % (v/v) acetic acid) for at least 30 min. The gels can be stored in MQ-water for several days or dried.

5.4.2 Western Blotting

Western blotting requires the transfer of proteins from SDS gels to a nitrocellulose membrane. Therefore, 2 blots per chamber were prepared as follows: extra thick filter paper (Thermo, 88620), transfer membrane (Amersham™ Protran™ 0.45 µm, GE Healthcare Life Sciences), SDS-PAGE gel, extra thick filter paper (Thermo, 88620). The blots were soaked in Western Blot buffer (Tris, SDS, Glycin, 20 % (v/v) ethanol). The transfer took place for approximately 10 min (1x blot U=30 V, I= max. or 2x blots U=60 V, I= max.) in the Turbo Blotter system from Biorad. The unspecific binding sites on the membrane were blocked by incubating in 5 % (w/v) milk powder solution (Roth, T145.3) in TBS buffer for at least 30 min. The membrane was afterwards incubated in the 1st antibody solution (for FKBP51: rabbit anti-FKBP5/FKBP51 Antibody (polyclonal, whole IgG; Bethyl), used: 1:1000 in 5 % (w/v) milk powder in TBS buffer and for Hsp90: Hsp90alpha/beta rabbit mAB, Bethyl sc-7947, used: 1:1000 in 5 % (w/v) BSA in TBS buffer) at 4 °C overnight on a rolling device. The membrane was then washed in TBS buffer, in TBS-T(0.1 % Tween) buffer, and again in TBS buffer for 5 min each. The membranes were transfered into the 2nd antibody solution (goat anti-rabbit IgG-heavy and light chain antibody HRP Conjugated (Bethyl), used: 1:1000 in 5 % (w/v) milk powder in TBS) and were incubated at room temperature for 1-2 h on a rolling device. Finally, the membrane was washed again in TBS, TBST-T(0.1 % Tween) and TBS buffer for 5 min each. Proteins were detected by chemiluminescence with 1:1 mixture of solution A and B (Immobilon™ Western Chemiluminescent HRP Substrate; Millipore) using a LAS-3000 detection system.

5.4.3 LC-MS

For synthesized proteins

For the characterization of the synthesized proteins, a 0.1 mg/mL solution was prepared in 50 % HPLC acetonitrile in water with 0.1 % TFA. The solutions were measured on an Agilent 6545 Accurate-Mass Q-TOF LCMS system after filtering of the solutions. Solvent A was 0.1 % formic acid in MQ water and solvent B was 0.1 % formic acid in HPLC grade acetonitrile . Method used: C3-1-95 in 15 min using the Agilent Zorbax 300SB-C3 column (2.1 mm × 150 mm, 5-µm particle size).

For recombinantly proteins

The recombinantly proteins measured by Thomas Nehls, on the other hand, were diluted in 0.2 % formic acid in 95:5 water:acetonitrile (v/v) to a protein concentration of 100 nM. The measurements were performed on a Synapt XS from Waters with a Acquity UPLC Protein BEH C4 column (1.7 µm particle size, 300 Å pore size, 50 x 2.1,mm, Waters) at 25 °C and an injection volume of 50 µL. Solvent A was 0.2 % formic acid in water (LC-MS grade) and solvent B was 0.2 % formic acid in acetonitrile (LC-MS grade). The gradient started at 95 % A and 5 % B and was maintained for 1 minute. Between 1 and 5 minutes, the gradient was increased to 50 % B. Over 5 to 10 min, 100 % B was reached and held for 4 min. Within 6 s, the gradient was again adjusted to 5 % B for 6 min to re-equilibrate the column. Subsequently, the sample was ionized with an ESI source in positive mode. The source voltage was 3 kV, with a Cone Voltage of 50 V, a Source Offset of 20 V, at a Source Temperature of 70 °C and Desolvation Temperature of 200 °C.

5.4.4 Analytical HPLC

The analytical HPLC measurements were performed on an Agilent Technologies 1200 Series system. The column used was Kinetex 2.6 µm (C18, 100 Å, 100 x 2.1 mm) with solvent A being 0.1 % TFA in MQ water and solvent B being 0.1 % TFA in HPLC grade acetonitrile. A linear gradient of 5-65 % solvent B in 60 min was used with a flow rate of 0.375 mL/min. The samples were prepared by dissolving the anaylte in 50 % solvent B in solvent A to a concentration of 1 mg/mL. The samples were filtered and 5 µL of each sample were injected.

5.4.5 ETD measurements

For Top-Down electron transfer dissociation (ETD) measurements, the analyte was dissolved in 200 mM ammonium acetate (pH 7, LC-MS grade) at a final concentration of 30 µM resulting in 50 µL protein solution which was mixed with 22 µL acetonitrile, 0.3 % formic acid (solution ratio 70:30 H_2O:ACN with approximately 0.1 % formic acid). The samples were measured by direct infusion in a nanoESI at a Synapt XS from Waters. Therefore, 10 µL protein solution was loaded in a nanoESI capillary, which was self-pulled using the Micropipette Puller Model P-97 from Sutter Instrument. The Capillary Voltage was set between 1.0 and 1.5 kV, the Sampling Cone at 50 V, the Source Offset at 20 V, the Source Temperature at 100 °C,the Cone Gas Flow at 20 L/h, and the Purge Gas Flow at 300 mL/h for the nanoESI-Source. 1,4-dicyanobenzene was used as ETD reagent with a Make-Up gas flow of 50 mL/min, flowing through the Glow Discharge at a current of 55 µA and a voltage of 0.9 kV. The trap cell was filled with anion radicals at an interval of one second for a duration of 0.1 seconds, whereby the pressure of the trap cell was $5.5 \cdot 10^{-2}$ mbar with a trap RF of 450 kHz. The transfer cell was used with Argon at a pressure of approximately $4.0 \cdot 10^{-2}$ mbar and a Collision Energy of 5 V. The TOF (time of flight) was set in resolution mode with a mass range of 50 to 2000 m/z, and measurements were performed for at least 20 min. The instrument was calibrated with sodium iodide. The evaluation was done with the software MASH Native.[157]

5.4.6 SEC-MS – native & denatured

For size exclusion chromatography - mass spectrometry measurements, the protein solutions were first diluted in 200 mM ammonium acetate buffer to a concentration of 10 µM. For the denatured measurements the protein solutions were diluted to a final protein concentration of 1 µM in 0.2 % formic acid in 30:70 Acetonitrile:Water (LC-MS grade). The SEC-MS measurements were performed on the Synapt XS from Waters, consisting of a PAL RTC Autosampler from Leap, a UHPLC with µBinary Pump and Auxiliary Pump, the HDX Manager and a bioZen SEC-2 column with SiO_2 particles (1.8 µm particle size, 150 Å pore size, 150 x 2.1 mm, Phenomenex). The oven temperature was set on 20 °C. 2 µL of sample was injected and the run was performed with an isocratic gradient of 35 µL/min with 50 mM ammonium acetate in water (LC-MS grade) for the native measurement and 0.2 % formic acid in water (LC-MS grade) for the denatured measurements. After chromatography the samples were ionized with an ESI source (positive mode, source voltage 3 kV, Cone Voltage 50 V, Source Offset 20 V, Source Temperature 70 °C and Desolvation Temperature of 200 °C).

5.4.7 IM-MS and CCS

For IM-MS (ion mobility mass spectrometry) measurements, protein analytes were first diluted in 200 mM ammonium acetate buffer to a concentration of 10 µM. The samples were measured with direct-infusion nanoESI at a Synapt G2-S from Waters with 10 µL of sample (settings: Capillary Voltage was between 1.0 and 2.0 kV, Sampling Cone 20 V, Source Offset 50 V, Source Temperature 30 °C, Cone Gas Flow 200 L/h, and Purge Gas Flow 300 mL/h). For the IM-MS measurements, helium at 180 mL/h was used as the cooling gas and nitrogen at 90 mL/h as the drift gas for the IMS cell. Three measurements with different wave velocities of 800, 1000 and 1150 m/s were conducted. The mobilograms were extracted with CIUSuite 2[158] For CCS calibration, bovine ubiquitin, equine horse cytochrome C, and equine holomyoglobin were measured. A logarithmic regression model was chosen for calibration.[136]

5.4.8 CIU

The protein samples were diluted in 200 mM ammonium acetate buffer to an initial protein concentration of 10 µM. The samples were measured with direct-infusion nanoESI. 10 µL protein solution was loaded in a self-pulled nanoESI capillary. A Micropipette Puller Model P-97 from Sutter Instrument was used to generate the capillaries. In the CIU experiment, helium at 180 mL/h was used as the cooling gas and N_2 at 90 mL/h as the drift gas for the IMS cell. The TOF was in resolution mode. The mass range was from 400 to 5000 m/z. The MS instrument was mass calibrated with sodium iodide to 5000 m/z. In this measurement, a charge state was isolated with the quadrupole and the arrival time distribution was measured. The initial collision energy of the trap was 4 V. Gradually, the trap CE was increased to 10 V, 15 V, 20 V and further until the signal of the protein was below the LOD. After that, the measurement was stopped. CIUSuite 2 was used for the evaluation.[158]

5.4.9 Protein Crystallization

Complexes were prepared by refolding F67E/K58 (i, i+9) and F67E/K60Orn (i, i+7) according to the protocol in section 5.6.7 . The obtained proteins in FPLC buffer (20 mM HEPES pH 8, 20 mM NaCl) were concentrated using Amicon Ultra 2 mL Centrifugal Filters and the samples were mixed with a slight excess of SAFit1 previously dissolved in 20 mM DMSO. The complexes were afterwards purified by size exclusion chromatography in FPLC buffer (20 mM HEPES pH 8, 20 mM NaCl) and the fractions at an elution time of approx. 13.5 min were concentrated to a protein concentration of 3 mg/mL for F67E/K60Orn (i, i+7) and 1 mg/mL for F67E/K58 (i, i+9). In case of F67C/K60C^{ox} the protein was crystallized in its apo form at a concentration of 40 mg/mL.

Crystallization was conducted according to the hanging drop vapour-diffusion method at room temperature, equilibrating mixtures of 1 µl protein (complex) and 1 µl reservoir against 500 µl reservoir solution. Crystals were obtained from reservoir solutions containing 32 % PEG-3350 for F67E/K58 (i, i+9) and 30 % PEG-3350 for F67E/K60Orn (i, i+7), 0.2 M NH4-acetate, 0.1 M HEPES-NaOH pH 7.5. For F67C/K60C^{ox} crystals were obtained from reservoir solution containing 30 % PEG1000, 0.1 M SPG pH8. Crystals were fished, cryoprotected with 30 % PEG-3350, 0.2 M NH4-acetate, 0.1 M HEPES-NaOH pH 7.5 and 10 % ethylene glycol and flash frozen in liquid nitrogen.

The crystallographic experiments were performed on the BL14.1 beamline at the Helmholtz-Zentrum BESSY II synchrotron, Berlin, Germany.[159] Diffraction data we-

re integrated with XDS and further processed with the implemented programs of the CCP4i and CCP4i2 interface.[160–163] With Aimless[160,164,165] the data reduction was conducted and crystal structures were solved by molecular replacement using Phaser.[166] Coot and Refmac5 were used for iterative model improvement and refinement.[167–172] The dictionary for SAFit1 was generated with PRODRG implemented in CCP4i.[173] The lactam bridge was generated with AceDRG.[174] Residues facing solvent channels without detectable side chain density were truncated.

Refinement table

PDB entry	**8PJ8 FKBP51FK1 F67E/K60Orn (i, i+7) SAFit1**	**8PJA FKBP51FK1 F67E/K58 (i, i+9) SAFit1**
Data collection		
Beamline	BESSY II (BL14.1)	BESSY II (BL14.1)
Wavelength	λ = 0.9184 Å	λ = 0.9184 Å
Space group	P 2_1 2_1 2_1	P 2_1 2_1 2_1
Cell dimensions		
a, b, c (Å)	45.22, 48.37, 57.20	42.06, 48.54, 56.03
α, β, γ (°)	90, 90, 90	90, 90, 90
Resolution (Å)	36.96-1.50 (1.53-1.50)	36.72-1.60 (1.63-1.60)
R_{merge}	0.058 (1.156)	0.083 (1.536)
R_{pim}	0.030 (0.621)	0.032 (0.801)
$I/\sigma(I)$	18.7 (1.9)	14.9 (1.4)
CC1/2	0.999 (0.729)	0.999 (0.527)
Completeness (%)	97.6 (97.1)	99.7 (99.0)
Redundancy	8.5 (8.2)	8.5 (8.7)
Refinement		
Resolution (Å)	36.96-1.50	36.72-1.60
No. of reflections	20089	15615
R_{work}/R_{free} (%)	19.3/21.6	19.6/21.9
Total number of atoms	2148	2068
Average B, all atoms (Å^2)	25.0	25.0
R.m.s. deviations		
Bond lengths (Å)	0.0122	0.0107
Bond angles (°)	1.794	1.751
Ramachandran plot		
Favoured	121 (98%)	119 (96%)
Allowed	2 (2%)	5 (4%)
Outlier	0	0

5.5 Protein crosslinking

5.5.1 Photocrosslinking

200 µL of $FKBP51^{FK1}$ F67Bpa (0.5 mg/mL) was transferred in a 96-well plate (Greiner, PS, clear, non-binding). The plate was incubated on ice under UV irradiation at 365 nm (2x15W, VL-215.L, 365 nm) for 30 min.

5.5.2 CuAAC

CuAAC (copper-catalyzed azide-alkyne cycloaddition) conditions were screened and analyzed by SDS-PAGE (see figure 3.13). Therefore, the protein $FKBP51^{FK1}$ E140C coupled to propargylmaleimide was used as a model protein. 10 µL protein (20 µM) solution was mixed with 10 µL buffer (native: 20 mM HEPES pH 7, 20 mM NaCl or denatured: 6 M urea in 20 mM HEPES pH 7m 20 mM NaCl). Stock solutions of copper sulfate (2 mM) and sodium ascorbate (5 mM) were freshly prepared in MQ-water. Additionally, a 10 mM stock solution of THPTA (tris-hydroxypropyltriazolylmethylamine) and a 1 mM stock solution of Azide-PEG (10 kDa) was prepared in MQ-water. A mixture of THPTA (4 µL) and $CuSO_4$ (1 µL) was added to 10 µL of the protein solution before Azide-PEG (1 µL) and finally sodium ascorbate (1 µL) were added. Moreover, 3 µL t-butanol were added to the respective samples. In total 60 conditions were tested with a reaction volume of 20 µL and an incubation time of 16 h at 25 °C. The respective end concentrations of the reagents for each condition were shown in figure 3.13. The stock solutions were diluted accordingly.

5.5.3 PEGylation

In this work, PEGylation was used as an indirect detection to identify free thiol groups (Maleimide-PEG, 5 or 10 kDa), azide moieties (DBCO-PEG, 10 kDa) or DBCO groups (Azide-PEG, 10 kDa). The PEGylation leads to an increase in protein mass and therefore a shift in SDS-PAGE analysis. Therefore, 5 mM stocks of MI-PEG, Azide-PEG and DBCO-PEG were prepared. The investigational protein samples were mixed with a 5x excess of the respective PEG reagent. The samples were then incubated at 25 °C for 2 h or for 16 h at 4 °C before loading on an SDS gel.

5.5.4 Crosslinking with DBCO-MI

For the DBCO-MI crosslinking, the protein $FKBP51^{FK1}$ F67pAzF/G13C or as a control $FKBP51^{FK1}$ G13C was diluted in FPLC buffer (20 mM HEPES pH 8, 20 mM NaCl) to a protein concentration of 20 µM. Afterwards, DBCO-MI (stock in DMSO: 250 mM) was added in a 5 fold excess. The mixture was incubated for 16 h-64 h at 4 °C. To remove excess of small molecules and reduce the volume, the samples were concentrated in Amicon Ultra 2 mL Centrifugal Filters, MWCO 10 kDa.

5.5.5 Cysteine directed crosslinking with BMB

A 20 mM stock of the crosslinker BMB (1,4-bismaleimidobutane) was prepared in DMSO. The proteins to be crosslinked, in this work $FKBP51^{FK1}$ F67C/G13C and as controls noCys, E140C and F67C were diluted to a protein concentration of either 20 µM or 100 µM in phosphate buffer pH 6.8. BMB was added in an equimolar amount. The reaction mixture

was incubated for 16 h at 25 °C. To remove excess of small molecules and reduce the volume, the samples were concentrated in Amicon Ultra 2 mL Centrifugal Filters, MWCO 10 kDa. Subsequently, maleimide PEGylation (10 kDa) was performed according to section 5.5.3 either for indirect detection of free thiol groups, in which case only a small amount of the protein sample was treated with MI-PEG, or the entire sample was MI-PEGylated to increase the protein mass of unreacted proteins for easier separation. In the latter case, an SEC run was performed as described in section 5.2.5. Samples were analyzed by SDS-PAGE (section 5.4.1) and/or by measuring intact mass (section 5.4.3).

5.5.6 Cysteine directed crosslinking with DBA

A 200 mM stock of the crosslinker DBA (dibromoacetone) was prepared in DMSO. The proteins to be crosslinked, in this work FKBP51^{FK1} F67C/K60C, F67C/K65C and F67C/S69C were diluted to a protein concentration of 20 µM in phosphate buffer pH 6.8. DBA was added in a 25x excess, resulting in an final DBA concentration of 500 µM. The reaction mixture was incubated for 64 h at 4 °C. To remove excess of small molecules and reduce the volume, the samples were concentrated in Amicon Ultra 2 mL Centrifugal Filters, MWCO 10 kDa. Subsequently, maleimide PEGylation (10 kDa) was performed according to section 5.5.3 either for indirect detection of free thiol groups, in which case only a small amount of the protein sample was treated with MI-PEG, or the entire sample was MI-PEGylated to increase the protein mass of unreacted proteins for easier separation. In the latter case, an SEC run was performed as described in section 5.2.5. Samples were analyzed by SDS-PAGE (section 5.4.1) and/or by measuring intact mass (section 5.4.3). Since MS measurements showed that the proteins with the double cysteines form disulfide bridges, leading to the cysteine side chains to be unavailable for crosslinking with DBA, new conditions were screened, as shown in table 3.3. Condition 18 showed the most satisfactory result. The proteins F67C/K58C and F67C/K60C were diluted in coupling buffer (50 mM Tris pH 7.2, 150 mM NaBr) and 500 µM TCEP (10 mM stock in MQ-water) and 2000 µM DBA (80 mM stock in DMSO) were added. The reaction mixtures were incubated for 16 h at 25 °C. To remove excess of small molecules and reduce the volume, the samples were concentrated in Amicon Ultra 2 mL Centrifugal Filters, MWCO 10 kDa. Subsequently, the solution was incubated with 100 µL maleimide-resin (PureCube, Maleimide Activated Agarose, 51103, CubeBiotech) at 25 °C. The flow through was collected and was next incubated with 250 mg of activated Thiol–Sepharose (Sigma, 4BT8512). The resin was activated by swelling 250 mg of the resin in 1 mL MQ water for 15 min at room temperature and then incubating the swollen resin with 1 mM TCEP solution. Incubation of the activated resin with the sample was 16 h at room temperature before characterizing the flow through by FP assay as described in section 5.3.1. Throughout the experiment, small amounts of sample were taken for MS analysis (section 5.4.3).

5.5.7 Disulfid bridge formation

After protein production and purification of the doubleCys variants FKBP51^{FK1} F67C/K58C and F67C/K60C according to section 5.2, the fractions after SEC containing the pure protein of interest were incubated at room temperature overnight (approximately 16 h). The formation of a disulfide bond was checked by mass spectrometry (see section 5.4.3). When the disulfid bridging was successfull, the protein solutions were concentrated in Amicon Ultra 2 mL Centrifugal Filters, MWCO 10 kDa.

5.6 Protein Synthesis

All reagents were used without further purification. The Fmoc protected amino acids were purchased from CreoSalus, Novabiochem, ChemImpex and ChemPrep. As solid support H-Rink amide ChemMatrix LL was used (PCAS BioMatrix, Loading: and bead size: 100-200 mesh dry). The activators HATU (O-(7-azabenzotriazol-1-yl)-N,N,N′,N′-tetramethyluronium hexafluorophosphate) and PyAOP (7-azabenzotriazol-1-yloxy) tripyrrolidinophosphonium hexafluorophosphate) were purchased from ChemImpex. DCM (dichlormethane) and DMF (N,N-diisopropylethylamine) as solvents were purchased from EMD Millipore and each DMF bottle contained an AldraAmine trapping packet from Sigma-Aldrich. DIPEA (Diisopropylethylamine), piperidine, TFA (trifluoroacetic acid), FA (formic acid), DMSO (dimethyl sulfoxide), Pd(PPh3)4 (Tetrakis(triphenylphosphine)palladium(0)), PhSiH3 (phenyl silane), EDT (1,2- ethanedithiol), thioanisole, phenol, acetonitrile (HPLC grade), were purchased from Sigma-Aldrich. Water was deionized by a Milli-Q Reference water purification system from Millipore.

5.6.1 Automated peptide flow synthesis (AFPS)

For protein synthesis the self-build automated flow peptide synthesizer (Peptidator) in the Pentelute Lab was used.[143] Approximately 130 mg of H-Rink amide ChemMatrix resin (loading 0.18 mmol/g) were swollen in amine free DMF for 15 min and washed alternating with DCM and DMF before the beads were loaded into the reactor, a fritted syringe (6 mL), in the Peptidator. The general flow rate was 40 mL/min, the temperature in the loop equals 90 °C and 85-90 °C in the reactor. The first step was a 60 s wash step at elevated temperatures with 40 mL of amine free DMF. The coupling-deprotection cycle was repeated for all additional monomers used in the sequences. The sequence starts at position A16, the first three amino acids GAP were residues of the tag of the recombinant protein. In addition, the mutation A19T was introduced, as this mutation has been shown to be important for crystallization.[45] The wildtype cysteines were abolished (C103A/C107I) and the methionines were replaced by norleucin (M48NLeu, M97NLeu), to protect the protein from oxidation. For the lactam bridged variants F67 was repalaced by Fmoc-Glu(OAll)-OH. For F67E/K58 (i, i+9) in position 58 and for F67E/K60 (i, i+7) in position 60 Fmoc-Lys(Alloc)-OH was incorporated. For variants F67E/K60Orn (i, i+7) and F67E/K60Dab (i, i+7) in position 60 Fmoc-Orn(Alloc)-OH and Fmoc-Dab(Alloc)-OH, respectively, were incorporated. The N-terminus was protected by using Boc-Gly-OH as the last amino acid. For the F67Amb/G64S variant the unnatural amino acid building block 4-aminomethylbenzoic acid was introduced in position 67. Here, the N-terminus was protected by biotinylation with Biotin-PEG_4-propionic acid.

5.6.2 Biotinylation

Variant F67Amb/G64S was biotinylated with Biotin-PEG_4-propionic acid after the synthesis. Therefore, the beads (0.022 mmol, 1 eq.) were mixed with Biotin-PEG_4-propionic acid (0.117 mmol, 5 eq.), PyAOP (0.117 mmol, 5 eq.) and DIPEA (0.235 mmol, 10 eq.) in 1 mL amine free DMF. The incubation was performed over night for approximately 16 h at room temperature on a rolling device. Afterwards, the beads were washed thoroughly with amine free DMF and DCM, before drying in an vacuum chamber.

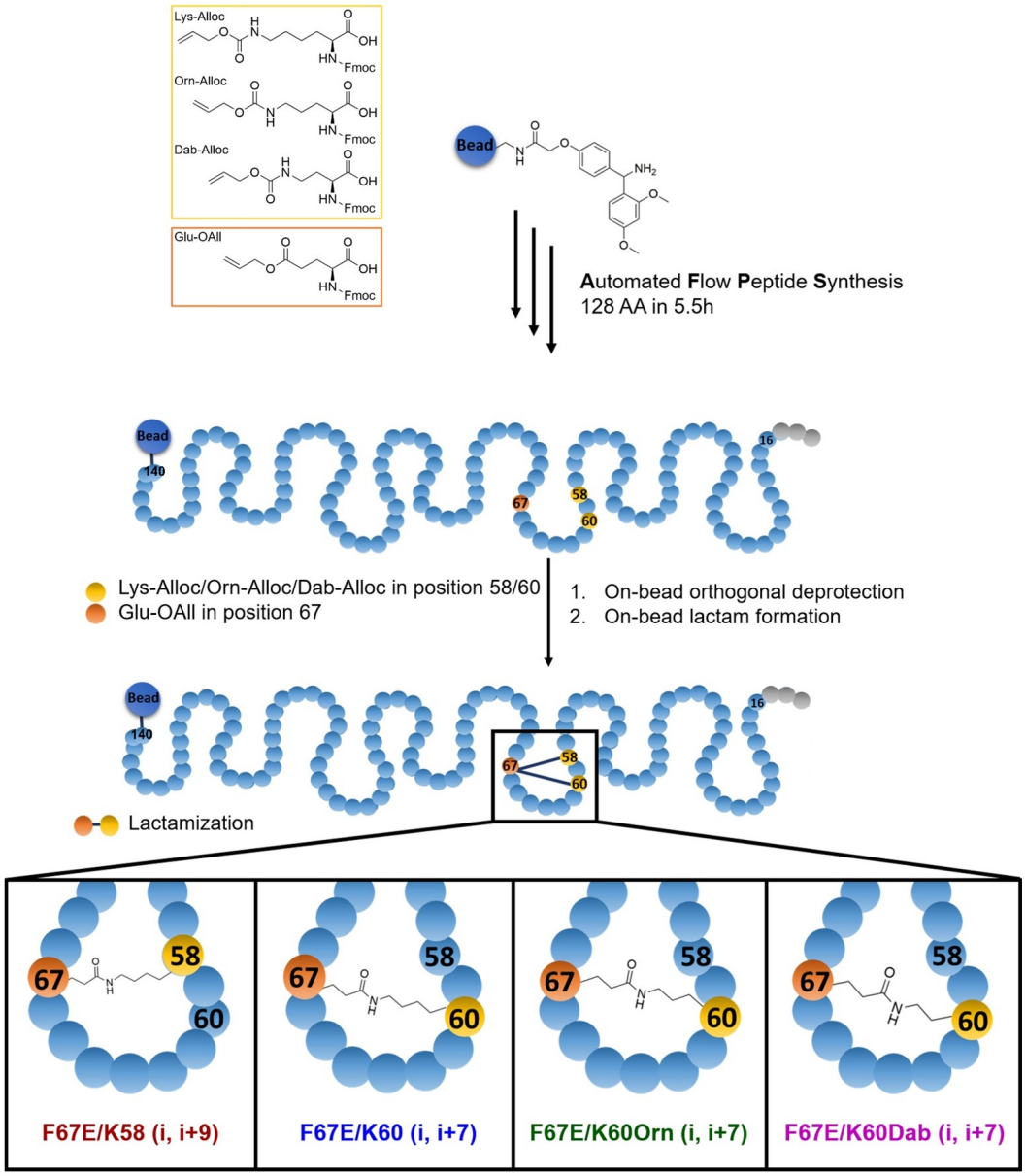


Figure 5.2: Total synthesis of lactam-bridged FKBP51^{FK1} variants by automated flow peptide synthesis (AFPS). AFPS allows the site-specific incorporation of orthogonal protected amino acids, for subsequent on-bead orthogonal deprotection and lactamization. This way four variants were synthesized with different ring sizes.

5.6.3 Orthogonal on-bead deprotection of Alloc and OAll groups

After synthesis, the beads (loading: 0.0221 mmol, 1 eq.) were washed thoroughly with amine free DMF (3 x 5 mL) and DCM (3 x 5 mL) before drying in a vacuum chamber. To the beads was added a mixture of 5.1 mg $Pd(PPh_3)_4$ (0.00442 mmol, 0.2 eq.) in 4 mL DCM and 54 µL $PhSiH_3$ (0.442 mmol, 20 eq.). The suspension was incubated for 30 min at room temperature. This step was repeated. The beads were then washed with DCM (3

x 5 mL) and subsequently DMF (3 x 5 mL).

5.6.4 Lactamization

The beads (loading: 0.0221 mmol, 1 eq.) were mixed with a solution of 57.6 mg PyAOP (0.1105 mmol, 5 eq.) and 80 µL DIPEA (0.442 mmol, 20 eq.) in 2 mL amine free DMF. The reaction mixture was incubated for 4 h at room temperature on a rolling device. The beads were washed 3x for 5 min each with DMF and DCM before they were then dried in a vacuum chamber.

5.6.5 Cleavage

To remove the remaining protecting groups and release the proteins from the solid support, the dried beads (200 mg) were mixed with 5 mL cleavage solution (82. % TFA, 5 % water, 5 % phenol, 5 % thioanisole, 2.5 % EDT). The mixture was incubated on a rolling device for 3 h at room temperature. Precipitation of the proteins was achieved by the addition of cleavage solution in 40-45 mL ice cold diethyl ether. The precipitate was collected by centrifugation and traces of diethyl ether and TFA were removed as much as possible by evaporation with a nitrogen gas flow. Finally, the proteins were dissolved in 50 % acetonitrile with 0.1 % TFA in water and dried by lyophilization.

5.6.6 Preparative mass directed RP-HPLC

The lyophilized crude proteins (30-40 mg) were dissolved in 5 mL loading buffer (6 M guanidine hydrochloride, 100 mM Tris at pH 8). Before loading on the column, the samples were filtrated with a Nylon 0.22 µm syringe filter. A gradient of 28 % to 48 % solvent B in 100 min was used. Solvent A was MQ water with 0.1 % TFA and Solvent B was HPLC grade acetonitrile with 0.1 % TFA. The semipreparative Agilent Zorbax 300SB-C3 column (9.4 mm × 250 mm, 5-µm particle size) was used at a temperature of 60 °C and with a flow rate of 4 mL/min. The pure fractions were combined, lyophilized, weighted and used for further characterization.

5.6.7 Refolding of the synthesized FKBP51 variants

For the refolding, 1 mg of the synthetic protein was dissolved in 250 µL denaturing buffer (6 M guanidine HCl, 100 mM Tris pH 8) and incubated at room temperature for 15 min on a rolling device. In a 96-well plate 190 µL of refolding buffer (50 mM Tris-Cl pH 8.5, 9.6 mM NaCl, 0.4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.5 M arginine, 0.4 M sucrose, 0.75 M guanidine HCl) was filled in each well (approximately in 25 wells). 10 µL of the protein solution was added slowly in each well under stirring before the plate was incubated for 1 h at 4°C. Afterwards, the fractions were combined and the buffer was exchanged to 20 mM HEPES pH 8, 20 mM NaCl and concentrated using Amicon® Ultra 2 mL Centrifugal Filters, MWCO 10 kDa.

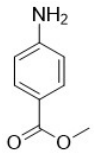
5.7 Ligand-based affinity chromatography

5.7.1 Synthesis

The characterization of the synthesized compounds was conducted on the Agilent 1260 Infinity II System with an EC-C18 column (Poroshell 120 3 mm × 150 mm, 2.7 µm) and a diode array detector connected to a 6125B MSD single quadrupole detector. Solvent A was 0.1 % formic acid in water and solvent B 0.1 % formic acid in HPLC grade acetonitrile. Furthermore, the compounds were analyzed by nuclear magnetic resonance spectroscopy (NMR) at the NMR facility at TUDa (Technische Universität Darmstadt). The instrument used was a Bruker AC300, AR300 or DRX500. TLC (thin layer chromatography) was conducted on aluminium plates coated with silica 60 F_{254nm} from Merck. Visualization of the compound spots was achieved by UV light and/or by staining the TLC plate with Hanessian or ninhydrin stain solution.

The chemical structures are summarized in figure 7.1 in the appendix.

Methyl 4-aminobenzoate (2)



4-aminobenzoic acid (2 g, 15 mmol, 1.00 eq.) was dissolved in 60 mL methanol and cooled to 4 °C. Then thionyl chloride (3.2 mL, 44 mmol, 3.00 eq.) was added dropwise. The resulting solution was allowed to warm to room temperature and was stirred for 16 h. After complete conversion of the starting material, the reaction mixture was first neutralized by the addition of saturated $NaHCO_3$ solution and extracted three times with DCM. The combined organic layers, were washed with NaOH (60 mL, 3 M) and dried over $MgSO_4$. Finally, the solvent was removed under reduced pressure and compound **2** was obtained as a white solid.

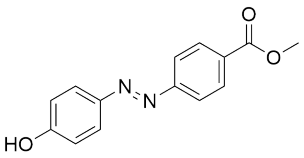
Yield: 2 g (13.24 mmol, 91 %)

TLC (CH/EA, 2:1 + 3 % Trimethylamine): R_f = 0.23

^{1}H NMR (500 MHz, DMSO-d6) δ 7.68 – 7.65 (m, 2H), 6.60 – 6.58 (m, 2H), 5.96 (s, 2H), 3.75 (s, 3H).

^{13}C NMR (126 MHz, DMSO) δ 166.35, 153.45, 131.05, 115.78, 112.65, 51.08.

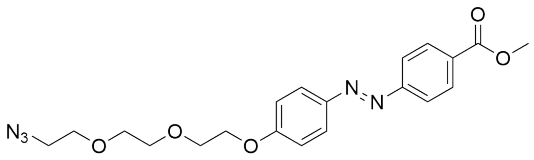
Methyl (E)-4-((4-hydroxyphenyl)diazenyl)benzoate (3)



2 (1.5 g, 10 mmol, 1.00 eq.) was dissolved in 12.5 mL water and 1.25 mL conc. HCl. The resulting solution was cooled to 4 °C before a solution of sodium nitrite (0.69 g, 10 mmol, 1 eq.) in 19 mL water was added. Then, a pre-cooled solution of phenol (1.03 g, 11 mmol,

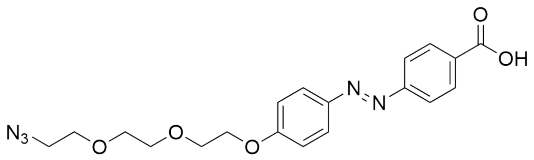
1.1 eq.) in 15 mL water was added and the reaction mixture was stirred for 30 min at 4 °C. Finally, K_2CO_3 (4 g, 29 mmol, 3 eq.) was added and the resulting reaction mixture was stirred for 18 h at room temperature. After the solvent was removed under reduced pressure, the crude product was purified by automated flash chromatography (cyclohexane/ethyl acetate 5:1). The obtained product was further purified by recrystallizing from a 1:1 water:ethanol solution, yielding an orange solid.
Yield: 1.5 g (5.85 mmol, 60 %)
TLC (CH/EA, 5:1): R_f = 0.20
^{1}H NMR (500 MHz, DMSO-d6) δ 10.48 (s, 1H), 8.15 – 8.11 (m, 2H), 7.93 – 7.89 (m, 2H), 7.88 – 7.84 (m, 2H), 7.00 – 6.96 (m, 2H), 3.90 (s, 3H).
^{13}C NMR (126 MHz, DMSO) δ 165.70, 161.75, 154.77, 145.33, 130.57, 130.43, 125.36, 122.20, 116.08, 52.30.
LC-MS (50 – 100 % solvent B, 2 min) R_t = 0.886 min, purity (220 nm): 89 %, m/z: calculated = 257.08 [M+H]+, found = 257.10 [M+H]+.

Methyl (E)-4-((4-(2-(2-(2-azidoethoxy)ethoxy)ethoxy)phenyl)diazenyl)benzoate (5)



3 (60 mg, 234 µmol, 1.00 eq.), **4** (77.12 mg, 234 µmol, 1 eq.) and K_2CO_3 (129.44 mg, 937 µmol, 4 eq.) were dissolved in 15 mL acetonitrile and the resulting reaction mixture was heated to reflux overnight. The reaction mixture was allowed to cool to room temperature and was diluted with DCM. The organic layer was then washed with water (2 × 30 mL) followed by brine (2 × 30 mL) and the combined organic layers were dried over $MgSO_4$. Then the solvent was removed under reduced pressure and the crude product was purified by column chromatography (cyclohexane/ethyl acetate 4:1). An orange solid was obtained.
Yield: 83.6 mg (202.34 µmol, 86 %)
TLC (CH/EA, 4:1): R_f = 0.33
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.21 – 8.16 (m, 2H), 7.98 – 7.94 (m, 2H), 7.94 – 7.91 (m, 2H), 7.10 – 7.03 (m, 2H), 4.26 – 4.23 (m, 2H), 3.97 (s, 3H), 3.95 – 3.92 (m, 2H), 3.79 – 3.76 (m, 2H), 3.73 – 3.70 (m, 4H), 3.41 (t, J = 6.9, 3.3 Hz, 2H).
^{13}C NMR (126 MHz, CDCl3) δ 166.62, 161.91, 155.37, 147.12, 131.22, 130.58, 125.15, 122.37, 114.96, 70.96, 70.78, 70.13, 69.71, 67.81, 52.26, 50.71.

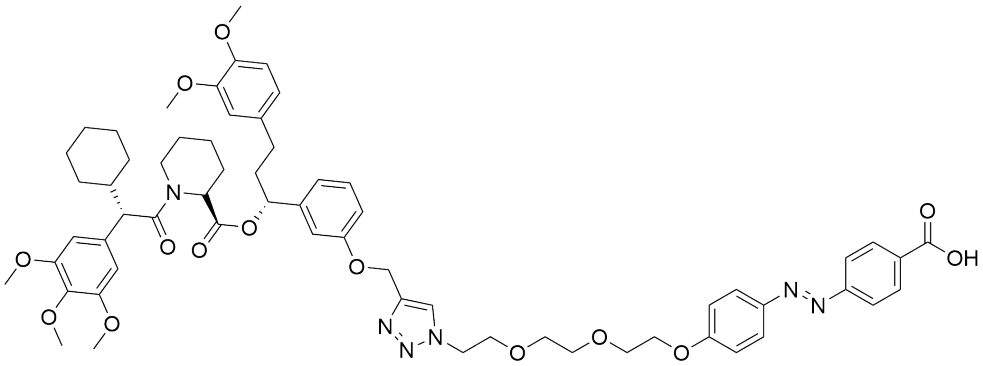
(E)-4-((4-(2-(2-(2-Azidoethoxy)ethoxy)ethoxy)phenyl)diazenyl)benzoic acid (6)



5 (83.6 mg, 202 µmol, 1.00 eq.) was dissolved in 1:1 THF:water before LiOH (239.5 mg, 10 mmol, 50 eq.) was added. The reaction mixture was stirred at room temperature for approximately 16 h. The crude product was obtained after extraction with ethyl acetate

(3 x 10 mL) which was subsequently purified by RP-HPLC on an Interchim puriFlash 5.250 system fitted with a preperative LC column (US5C18HQ-250/212, 5 µm, C18, 250 x 21.2 mm). The run was performed with a flow rate of 24 mL/min and a 50-100 %B within 6 CV (column volume) gradient, where solvent A was 0.1 % TFA in water and solvent B was 0.1 % TFA in acetonitrile. The fractions with the pure product were combined and the solvent was removed under reduced pressure. 37.7 mg of an orange solid was obtained.
Yield: 37.7 mg (93.94 µmol, 46 %)
TLC (CH/EA, 1:1 + 1 % formic acid): R_f = 0.3
^{1}H NMR (500 MHz, Chloroform-d) δ 8.31 – 8.17 (m, 2H), 8.00 – 7.90 (m, 4H), 7.10 – 6.98 (m, 2H), 4.29 – 4.20 (m, 2H), 3.99 – 3.87 (m, 2H), 3.81 – 3.73 (m, 2H), 3.71 – 3.64 (m, 4H), 3.41 – 3.35 (m, 2H).
^{13}C NMR (126 MHz, CDCl3) δ 162.17, 131.41, 125.40, 122.62, 115.14, 71.13, 70.95, 70.30, 69.88, 67.98, 50.87.
LC-MS (50 – 100 % solvent B, 2 min) R_t = 0.662 min & 1.649 min (cis-trans isomers), purity (220 nm): 95 %, m/z: calculated = 400.15 [M+H]+,found = 400.20 [M+H]+.

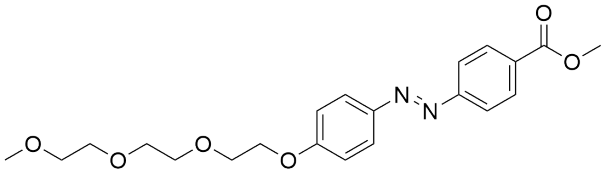
4-((E)-(4-(2-(2-(2-(4-((3-((R)-1-(((R)-1-((S)-2-Cyclohexyl-2-(3,4,5-trimethoxyphenyl)acetyl)piperidine-2-carbonyl)oxy)-3-(3,4-dimethoxyphenyl)propyl)phenoxy)methyl)-1H-1,2,3-triazol-1-yl)ethoxy)ethoxy)ethoxy)phenyl)diazenyl)benzoic acid (8)



6 (30.0 mg, 75.11 µmol, 1.00 eq.), **7** (54.67 mg, 75 µmol, 1 eq.), sodium ascorbat (5.95 mg, 30 µmol, 0.4 eq.) and copper sulfate (3 mg, 12 µmol, 0.4 eq.) were dissolved in 1:1:1 tert-butanol:water:DMSO and stirred for 16 h at room temperature under argon atmosphere. The reaction mixture was then diluted in water and acidified using 1 M HCl. The resulting mixture was afterwards extracted with ethyl acetate (3 x 10 mL) and washed once with brine (50 mL) before combining and drying the organic layers with $MgSO_4$. After solvent removal under reduced pressure, the crude product was purified by RP-HPLC on an Interchim puriFlash 5.250 system fitted with a preperative LC column (US5C18HQ-250/212, 5 µm, C18, 250 x 21.2 mm). The run was performed with a flow rate of 16 mL/min and a 30-100 %B within 6 CV (column volume) gradient, where solvent A was 0.1 % TFA in water and solvent B was 0.1 % TFA in acetonitrile. The fraction with the pure product were combined and the solvent was removed under reduced pressure and lyophilized. 69 mg of an orange solid was obtained.
Yield: 69.0 mg (61.25 µmol, 81 %)
TLC (CH/EA, 1:1 + 1 % formic acid): R_f = 0.1
^{1}H NMR (500 MHz, $CDCl_3$) δ 8.26 – 8.17 (m, 2H), 7.92 – 7.91 (m, 2H), 7.91 – 7.90 (m,

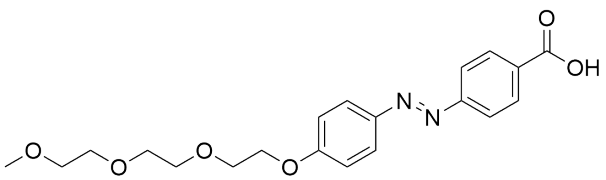
3H), 7.11 (t, J = 7.9 Hz, 1H), 7.04 – 6.98 (m, 3H), 6.82 (dd, J = 8.2, 2.5 Hz, 1H), 6.76 – 6.73 (m, 1H), 6.72 (t, J = 2.0 Hz, 1H), 6.61 (d, J = 6.9 Hz, 2H), 6.50 (s, 2H), 6.42 (s, 1H), 5.56 (dd, J = 8.0, 5.8 Hz, 1H), 5.48 (d, J = 5.5 Hz, 1H), 5.18 (s, 2H), 4.59 – 4.55 (m, 2H), 4.22 – 4.17 (m, 2H), 3.99 (d, J = 14.0 Hz, 1H), 3.92 (t, J = 5.0 Hz, 2H), 3.85 (s, 3H), 3.83 (s, 6H), 3.82 (s, 2H), 3.77 (s, 2H), 3.71 (s, 6H), 3.66 (dd, J = 6.2, 2.9 Hz, 2H), 3.41 (d, J = 10.0 Hz, 1H), 2.64 – 2.51 (m, 1H), 2.52 – 2.39 (m, 1H), 2.39 – 2.23 (m, 2H), 2.16 – 2.01 (m, 2H), 1.88 – 1.76 (m, 2H), 1.72 – 1.58 (m, 5H), 1.58 – 1.50 (m, 1H), 1.48 – 1.36 (m, 1H), 1.34 – 1.22 (m, 2H), 1.20 – 1.07 (m, 2H), 1.03 – 0.85 (m, 1H), 0.81 – 0.68 (m, 1H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 173.14, 171.32, 170.28, 170.11, 161.85, 157.93, 155.87, 153.13, 148.83, 147.29, 147.14, 143.44, 141.78, 136.89, 133.36, 133.22, 131.28, 129.74, 125.27, 124.59, 122.48, 120.28, 119.20, 114.91, 113.90, 113.46, 111.83, 111.37, 105.76, 105.26, 75.83, 70.73, 70.58, 69.64, 69.14, 67.73, 61.19, 60.78, 56.30, 56.05, 55.92, 55.81, 55.04, 52.42, 50.91, 43.99, 41.14, 37.76, 32.70, 30.99, 30.62, 26.73, 26.68, 26.49, 26.35, 26.12, 26.00, 25.45, 20.82.
LC-MS (70 – 100 % solvent B, 2 min) R_t = 0.828 min & 1.525 min (cis-trans isomers), purity (220 nm): 99 %, m/z: calculated = 1127.53 [M+H]+, found = 1127.50 [M+H]+.

Methyl (E)-4-((4-(2-(2-(2-methoxyethoxy)ethoxy)ethoxy)phenyl)diazenyl)benzoate (10)



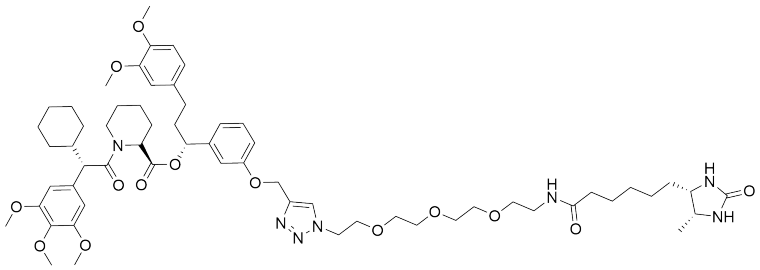
3 (63.67 mg, 249 µmol, 1.00 eq.), **9** (62 µL, 249 µmol, 1 eq.) and K_2CO_3 (137.37 mg, 994 µmol, 4 eq.) were dissolved in acetonitril and stirred for 16 h at reflux. The reaction solution was afterwards extracted with DCM (3 x 30 mL) and washed once with water and brine (each 50 mL) before combining and drying the organic layers with $MgSO_4$. The crude product was purified by manual column chromatography with 2:1 cyclohexane: ethyl acetate as eluent and 15 g silica (SiO_2, Macherey−Nagel, particle size 0.04−0.063 mm). An orange solid was obtained.
Yield: 60.0 mg (149.19 µmol, 86 %)
TLC (CH/EA, 2:1): R_f = 0.2
LC-MS (50 – 100 % solvent B, 2 min) R_t = 0.361 min & 0.768 min (cis-trans isomers), purity (220 nm): 99 %, m/z: calculated = 403.18 [M+H]+, found = 403.20 [M+H]+.

(E)-4-((4-(2-(2-(2-Methoxyethoxy)ethoxy)ethoxy)phenyl)diazenyl)benzoic acid (11)



10 (60.0 mg, 149 µmol, 1.00 eq.) was dissolved in 1:1 THF:water before LiOH (239.5 mg, 10 mmol, 67 eq.) was added. The reaction mixture was stirred at room temperature for approximately 16 h. The reaction mixture was then diluted in water and acidified using 1 M HCl. The crude product was obtained after extraction with ethyl acetate (3 x 10 mL) and washing with HCl and brine (1 x 50 mL). Subsequently, the crude product was purified by RP-HPLC on an Interchim puriFlash 5.250 system fitted with a preperative LC column (US5C18HQ-250/212, 5 µm, C18, 250 x 21.2 mm). The run was performed with a flow rate of 24 mL/min and a 50-100 %B within 6 CV (column volume) gradient, where solvent A was 0.1 % TFA in water and solvent B was 0.1 % TFA in acetonitrile. The fractions with the pure product were combined and the solvent was removed under reduced pressure. 35.0 mg of an orange solid was obtained.
Yield: 35.0 mg (90.17 µmol, 60 %)
TLC (CH/EA, 1:1 + 1 % formic acid): R_f = 0.28
^{1}H NMR (500 MHz, DMSO-d6) δ 8.15 – 8.09 (m, 2H), 7.95 – 7.86 (m, 4H), 7.19 – 7.14 (m, 2H), 4.24 – 4.21 (m, 2H), 3.80 – 3.76 (m, 2H), 3.61 – 3.58 (m, 2H), 3.55 – 3.50 (m, 4H), 3.44 – 3.40 (m, 2H), 3.24 – 3.22 (m, 3H).
^{13}C NMR (126 MHz, DMSO) δ 166.77, 161.87, 154.47, 146.25, 132.21, 130.58, 124.99, 122.23, 115.23, 71.27, 69.97, 69.80, 69.61, 68.79, 67.75, 58.04.
LC-MS (50 – 100 % solvent B, 2 min) R_t = 0.282 min & 0.575 min (cis-trans isomers), purity (220 nm): 99 %, m/z: calculated = 389.16 [M+H]+, found = 389.00 [M+H]+.

(R)-3-(3,4-Dimethoxyphenyl)-1-(3-((1-(18-((4S,5R)-5-methyl-2-oxoimidazolidin-4-yl)-13-oxo-3,6,9-trioxa-12-azaoctadecyl)-1H-1,2,3-triazol-4-yl)methoxy)phenyl)propyl (R)-1-((S)-2-cyclohexyl-2-(3,4,5-trimethoxyphenyl)acetyl)piperidine-2-carboxylate (13)



12 (18.14 mg, 44 µmol, 1.00 eq.), **7** (31.86 mg, 44 µmol, 1 eq.), sodium ascorbat (3.47 mg, 17.5 µmol, 0.4 eq.) and cooper sulfate (4.37 mg, 17.5 µmol, 0.4 eq.) were dissolved in 1:1:1 tert-butanol:water:DMSO and stirred for 16 h at room temperature under argon atmosphere. The reaction solution was afterwards extracted with ethyl acetate (3 x 10 mL) and washed once with brine (50 mL) before combining and drying the organic layers with $MgSO_4$. After solvent removal under reduced pressure, the crude product was purified by RP-HPLC on an Interchim puriFlash 5.250 system fitted with a preperative LC column (US5C18HQ-250/212, 5 µm, C18, 250 x 21.2 mm). The run was performed with a flow rate of 16 mL/min and a 50-100 %B within 6 CV (column volume) gradient, where solvent A was 0.1 % TFA in water and solvent B was 0.1 % TFA in acetonitrile. The fractions with the pure product were combined and the solvent was removed under reduced pressure and lyophilized. 20 mg of a white solid was obtained.
Yield: 20 mg (1.57 µmol, 40 %)
^{1}H NMR (500 MHz, $CDCl_3$) δ 7.87 – 7.80 (m, 1H), 7.13 (t, J = 7.9 Hz, 1H), 7.01 – 6.94

(m, 1H), 6.90 – 6.84 (m, 1H), 6.81 – 6.73 (m, 2H), 6.68 – 6.61 (m, 2H), 6.51 – 6.41 (m, 4H), 5.56 (dd, J = 8.1, 5.6 Hz, 1H), 5.46 (d, J = 5.4 Hz, 1H), 5.16 (s, 3H), 4.57 (t, J = 5.1 Hz, 4H), 3.90 (t, J = 5.1 Hz, 3H), 3.86 – 3.80 (m, 9H), 3.77 (s, 3H), 3.70 (s, 5H), 3.59 – 3.53 (m, 9H), 3.45 – 3.37 (m, 2H), 2.85 – 2.74 (m, 1H), 2.64 – 2.51 (m, 1H), 2.49 – 2.42 (m, 1H), 2.40 – 2.32 (m, 1H), 2.26 – 2.18 (m, 2H), 2.13 – 2.02 (m, 2H), 2.00 – 1.91 (m, 1H), 1.89 – 1.78 (m, 2H), 1.72 – 1.57 (m, 9H), 1.38 – 1.26 (m, 9H), 1.18 – 1.04 (m, 6H).
^{13}C NMR (126 MHz, $CDCl_3$) δ 174.90, 173.08, 170.81, 158.45, 153.72, 153.45, 149.26, 147.72, 142.33, 137.23, 133.89, 133.80, 130.12, 124.69, 119.43, 114.22, 113.77, 112.25, 111.76, 106.18, 105.69, 76.10, 70.86, 70.75, 70.70, 70.43, 70.08, 69.69, 62.02, 61.15, 56.89, 56.70, 56.43, 56.32, 56.24, 55.37, 52.59, 52.39, 50.91, 44.17, 41.70, 39.84, 38.36, 36.15, 33.13, 31.42, 31.04, 29.46, 28.95, 27.15, 26.92, 26.57, 26.50, 26.12, 25.89, 25.59, 21.33, 15.75.
LC-MS (50 – 100 % solvent B, 2 min) R_t = 0.994 min, purity (220 nm): 99 %, m/z: calculated = 1142.63 [M+H]+, found = 1142.60 [M+H]+.

5.7.2 Generation of cell lysates

Adherent cell lines such as HEK293 and N2a cells were cultured in TC-dish 150, Standard (Sarstedt, 83.3903) or TC-dish 100, Standard (Sarstedt, 83.3902). The used media was DMEM, Dulbecco´s Modified Eagle Medium ([+] 4.5 g/L D-Glucose, L-Glutamin, [+] Pyruvat, Gibco, Life Technologies, 41966-052, 500 mL) where 10 % (v/v) FBS (Fetal Bovine Serum, Gibco, Life Technologies, 10500-064) and 1 % (v/v) Pen/Strep (Pencillin/Streptomycin, Gibco, Life Technologies, 5140-122) were added. The suspension cell lines Jurkat WT and Jurkat FKBP51 knockout were cultured in CELLSTAR Filter Cap Suspension Culture Flasks (690195, 50 mL), Tissue Culture Flask 150 from TRP (90151) or Flask T75 from Eppendorf (0030711122). The media used was RPMI Medium 1640 ((1X), [+] 4.5 g/L D-Glucose, [+] L-Glutamin, [+] 1.5g/L Sodium Bicarbonate, [+] 110mg/L Sodium Pyruvat, Gibco, Life Technologies, A10491-01) where 10 % (v/v) FBS, 1 % (v/v) Pen/Strep and 1X GlutaMAY (100X, Gibco, 35050-038) were added. For detachment of cells from the surface, trypsin (0.25 % Trypsin-EDTA (1x), Gibco, Life Technologies, 25200-056) was used, and cells were washed with DPBS (1x), Dulbecco‘s Phosphate Buffered Saline ([-] $CaCl_2$, [-] $MgCl_2$). Trypan blue stain was used from Gibco, Life Technologies (15250-061). Cells were incubated at 37°C & 5 % CO_2. All media and supplements were prewarmed in a water bath at 37 °C.

Thawing of cell lines

An aliquot of the desired frozen cell line was thawed for 1 min at 37 °C in the water bath, before the transfer to 10 mL media in a 15 mL Falcon tube. The Falcon tube was inverted 2-3x and incubated for 5 min. Afterwards the suspension was centrifuged for 2 min at 1000 rpm at room temperature. The supernatant was discarded and 15 mL fresh culture growth media was added. Adherent cell lines were seeded on TC-dishes and suspension cell lines in culture flasks.

Seeding of cells

For N2a and HEK293: The cells were harvested by first removing the medium and washing the cells on the TC-dish 100 with DPBS. Then, 1 mL of Trypsin-EDTA solution was

distributed on each TC-dish. The plates were incubated for 5 min at 37 °C and 5 % CO_2. The Trypsin reaction for detachment was stopped by the addition of 7 mL culture growth medium in each TC-dish. The cells were resuspended and transfered to a 50 mL tube. The suspension was centrifuged at 1000 rpm for 3 min at room temperature. The supernatant was discarded and 30 mL of fresh cell growth medium was added. The cell pellet was carefully resuspended and seeded in a new TC-dish. For Jurkat cells the same protocol was conducted, but without the detachment step, since these cells were suspension cells. Also, the cells were cultured in culturing flasks.

Counting cells

To count the cells, 15 µL of cell suspension and 15 µL of trypan blue were mixed and transferred into a Neubauer counting chamber. Under the microscope, the cells were counted and the following equation was used to determine the cell number:

$$Cell\,number\,per\,mL = \frac{counted\,cell\,number \cdot 2(= Dilution\,with\,trypanblau)}{number\,of\,counted\,largesquares \cdot 10000(= chamber\,factor)}$$

Harvesting and lysis of the cells

Suspension cell lines were harvested by centrifuging the cell suspension at 120 g for 7 min at room temperature. The adherent cell lines were harvested by Trypsin-EDTA solution as described in the subsection Seeding of cells. The cells were washed with pre-cooled DPBS 4-5 times. Afterwards, the lysis of the cell pellets was achieved by the addition 2-5 mL of lysis buffer (NETN(100 mM NaCl, 20 mM Tris, pH 8, 0.5 mM EDTA, 0.5 % Nonidet P-40) or commercially available M-PER (ThermoFisher, 1 tablet protease inhibitor cocktail/10 mL (cOmplete Tablets, Mini, EasyPack, Roche, 04 693 124 001), 1 mM PMSF and 1 mM DTT were freshly added). The mixture was incubated for 1 h at 4 °C on a rolling device. To remove the precipitate, the mixture was centrifuged at 13 000 g and 4 °C for 30 min. The supernatant was aliquoted in pre-cooled protein low binding tubes (Sarstedt) and were frozen in liquid nitrogen. The cell lysate samples were stored at -80 °C.

Total protein determination by BCA assay

The total protein concentration of the cell lysates was determined by BCA assay (Bicinchonin acid). The BSA standard pre-dilution from ThermoFisher (Pierce, Bovine Serum Albumin Standard Pre-Diluted Set, 23208) was used to generate an accurate standard curve for protein quantification. Cell lysates were diluted in 1:20 and 1:40 ratios. 25 µL of each sample was transferred into a 96-well plate (microplate, 96 well, PS, F-bottom, clear) as duplicates and 200 µL BCA reagent from the Pierce Rapid Gold BCA Protein Assay Kit (ThermoFisher, A53227) were added. The plate was incubated at room temperature for 5 min before the absorption values at 480 nm were obtained with the TECAN plate reader.

5.7.3 Immobilization of a ligand on a solid support

The amine functionalized beads (TOYOPEARL AF-Amino-650M, Tosoh Bioscience, 100 µmol/mL±30 ligand density) were equilibrated in DMF. Next, HATU (500 mM) and ligand (20 mM) stocks in DMF were prepared. The ligand to be immobilized (0.01 eq. for 1 %

loading) was then mixed with HATU (1.1 eq.) and DIPEA (5 eq.) in DMF. The resulting mixture was added to the equilibrated beads (1:1 (w/v)). The suspension was incubated for 16 h at 50 °C. The beads were washed three times with DMF and the capping solution (30 eq. Ac_2O and 30 eq. DIPEA in DMF) was added next. The suspension was incubated for 1 h at 50 °C. To detect free amine groups, a Kaiser test was performed. Therefore, three solutions were prepared: A= 1 g ninhydrin in 20 mL EtOH, B= 5 g phenol in 2.5 mL EtOH and C= 16.5 mg KCN in pyridine. A small amount of beads (10 µL) to be tested were transferred to a new tube and additionally untreated beads were tested as negative control and water was tested as positive control. To the samples, 10 µL of each solution (A, B & C) was added and the resulting mixtures were incubated for 5 min at 100 °C. A blue color indicates free amine groups, while a yellow-reddish color means that no free amine groups were present in the sample. After a successful Kaiser test, the beads were thoroughly washed with 10 CV DMF and 20 CV buffer (usually: 20 mM HEPES pH 8, 20 mM NaCl).

5.7.4 Protocol for affinity chromatography

The beads immobilized with the ligand of interest were incubated with cell lysate (at least 100 ng total protein concentration in lysates) or pure protein (equimolar to ligand) for 1 h at 4 °C on a rolling device. Next, the beads were washed with 10 CV washing buffer as indicated for each experiment, followed by either an elution step or direct on-bead digestion. For the elution, the beads were mixed with the respective elution buffer and were incubated at room temperature for 15-30 min. This step was repeated 3-5 times. The samples were then analyzed by SDS-PAGE (see section 5.4.1) and/or Western Blotting (see section 5.4.2). For the on-bead digestion on the other hand, lysis buffer was prepared (6 M guanidine hydrochloride, 10 mM TCEP, 40 mM chloracetamide in 100 mM Tris pH 8.5 buffer) and mixed with the beads (1:1 (w/v)). The suspensions were incubated at 95 °C for 10 min followed by a 1:10 dilution step with 100 mM Tris pH 8.5 buffer. To each sample 1 µL trypsin (1 mg/mL) was added and incubated at 37 °C for at least 12 h and maximal 24 h. To each sample 1 Vol.% TFA was added and the peptides were purified with SDB-RPS Stage tips (Empore products, ThermoFisher) as described by the manufacturer. For the in-house measurements, the obtained peptides were dissolved in 10 µL solvent A (water with 2 % acetonitrile and 0.1 % formic acid) and measured on an Orbitrap XL connected to an easy nLC1200 system fitted with a C18 column (ReproSil-pur C18 1.9 µm silica column). A 90 min gradient was used, starting at 0 % solvent B (80 % acetonitrile and 20 % water with 0.1 % formic acid) and increasing to 40 % B within 50 min and then to 60 % within the next 20 min. In the following 10 min 100 % B was reached and held for the last 10 min of the gradient. MaxQuant version 2.0.1 was used to analyze the raw data with the Swiss-Prot data base with the default settings. Data processing and visualization was performed with Perseus 2.0.10.0. For the measurements performed by Dr. Tim Heymann at the Max-Planck Institute for Biochemistry, the peptides were measured on a Sciex 7600 ZenoTOF coupled to an Evosep One LC from Evosep Biosystems and Zeno SWATH method with 60 variable windows in the mass range 400-900 m/z with a TOF MSMS accumulation time of 13 ms. All samples were analyzed with the 60 samples per day method (21 min gradient) with 0.1 % formic acid in water as solvent A and 0.1 % formic acid in ACN as solvent B. The sample was loaded onto C18 Evotip pure according to the manufacturers instructions before LC/MS analysis. Raw files were analyzed using DIA-NN version 1.8.1 and searched against the reviewed human proteome (UniProt, November 2021, 20,360 entries without isoforms) with trypsin as digestion enzymes and up to 1

missed cleavage with a predicted library using the smart profiling option for precursors with a charge state of 1-4 from 300-1800 and fragments from 200-1800 m/z. Cysteine carbamidomethylation was set as fixed modification. The MS and MS1 accuracy was set to 20 ppm, match between runs and heuristic protein inference were enabled and the quantification strategy was set to robust LC (high precision). All data analysis was performed with the pg.matrix.tsv file that contained proteins with <1 % FDR.

6 Abbreviations

Abbreviation	Explanation
1,3-DBA	1,3-Dibromoacetone
2-PDS	2,2'-Bispyridyl disulfide
Ac_2O	Acetic anhydride
ACTH	Adrenocorticotropic hormone
ADME	Administration, distribution, metabolism, and excretion
AFPS	Automated Peptide Flow Synthesis
APDM	4,4'-Azobenzene dimaleimide
AST	Active site titration
BMB	1,4-Bismaleimidobutane
CCS	Collision cross-section
CIU	Collision induced unfolding
CL	Crosslink(er)
CNS	Central nervous system
CRH	Corticotropin-releasing hormone
CSDS	Chronic social defeat stress
CuAAC	Copper-catalyzed azide-alkyne cycloaddition
Cyp40	Cyclophilin40
Dab	Diaminobutyric acid
DBCO-MI	Dibenzocyclooctinmaleimide)
DIPEA	N,N-Diisopropylethylamine
DMF	Dimethylformamide
DTNB	5,5'-Dithiobis-(2-nitrobenzoic acid) (Ellman's reagent)
DTT	Dithiothreitol
EC_{50}	Half maximal effective concentration
EDTA	Ethylenediaminetetraacetic acid
ETD	Electron transfer dissociation
FBDD	Fragment-based drug discovery
FDA	U.S. Food and Drug Administration
FKBP51	FK506 binding protein 51
$FKBP51^{FK1}$	FK1 domain (1-140aa) of FKBP51
FSM	Fluorescein-5-maleimide
FP-assay	Fluorescence polarization assay
GRE	Glucocorticoid response elements
HATU	Hexafluorophosphate azabenzotriazole tetramethyl uronium
HEK293	Human Embryonic Kidney 293-cells
HEPES	4-(2-Hydroxyethyl)-1-piperazineethanesulfonic acid
HOP	Hsp70-Hsp90 organizing protein
HP(number)	Hausch Lab plasmid (number)
HPA-axis	Hypothalamic–pituitary–adrenal axis

Hsp90	Heat shock protein 90
HTS	Highthroughput screening
IMAC	Immobilized Metal Ion Affinity Chromatography
IM-MS	Ion mobility mass spectrometry
IPTG	Isopropyl β-D-1-thiogalactopyranoside
JHu WT	Jurkat wildtype cells
JHu 51 ko	Jurkat FKBP51 knockout cells
K_d	Dissociation constant
L-Ara	L-arabinose
LC-MS	Liquid chromatography tandem mass spectrometry
MES	2-(N-morpholino)ethanesulfonic acid
MI-PEG	Methoxypolyethylene glycol maleimide
MI-resin	Maleimide functionalized resin
M_w	Molecular weight
MWCO	Molecular weight cut-off
N2a cells	Mouse neuroblastoma cells
NCL	Native chemical ligation
NLe	Norleucine
Orn	Ornithine
ox.	oxidized
pAzF	para-Azidophenylalanine
pBpa	para-Benzoylphenylalanine
PDI	Protein disulfide isomerase
PEG	Polyethylene glycol
PK profile	Pharmacokinetic profile
PP5	Serine/threonine protein phosphates 5
PPIase	Peptidyl-prolyl cis-trans isomerase
PTM	Post-translational modification
red.	reduced
RP-HPLC	Reversed Phase-High Performance Liquid Chromatography
SAFit1	Selective antagonist for FKBP51 by induced fit 1
SAFit1-CLU	SAFit1 with cleavable unit
SAFit1-LA	SAFit1 analoga with low binding affinity
SDS-PAGE	Sodium dodecyl sulfate polyacrylamide gel electrophoresis
SEC	Size exclusion chromatography
SH-PEG	O-[2-(3-Mercaptopropionylamino)ethyl]-O'-methylpolyethylene glycol
SH-resin	Sulfhydryl funtionalized resin
SPAAC	Strained promoted azide-alkyne cycloaddition
SPPS	Solid-phase peptide synthesis
SPRINP-PCR	Single primer in parallel – polymerase chain reaction
SrtA	Sortase A
TCEP	Tris(2-carboxyethyl)phosphine
THPTA	Tris(3- hydroxypropyltriazolylmethyl)amine
T_m	Melting temperature
TNB	5-Thio-2-nitrobenzoic acid
TPR	Tetratricopeptiderepeat
TRABITA	Transient binding pockets
TRAPP webserver	Transient Pockets in Proteins web server

Tris	Tris[hydroxymethyl]aminomethane
UV radiation	Ultraviolet radiation
WCX	Weak cation exchange chromatography

7 Appendix

7.1 Summary of chemical structures 1-13

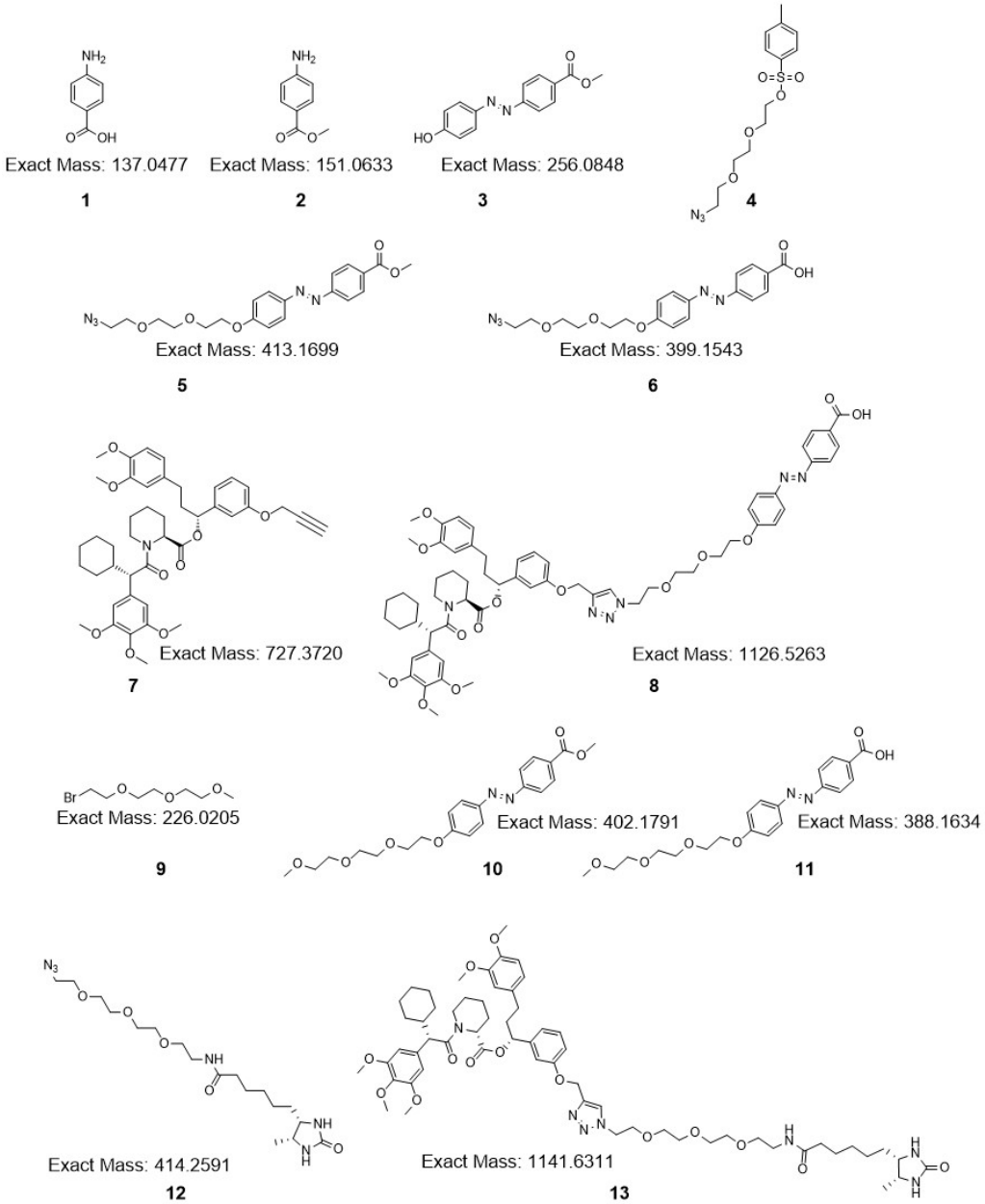


Figure 7.1: Chemical structures **1**-**13**.

7.2 Thermal shift assay data

Table 7.1: Determined melting temperatures (T_m) of FKBP51^{FK1} F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM) from the TRABITA library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	48.1	-	50.4	-
SAFit1	64.3	16.2	64.3	13.9
A2	49.7	1.6	52.1	1.7
B2	49.6	1.5	53.2	2.8
C2	48.0	-0. 1	50.0	-0.4
D2	-	-	42.0	-8.4
E2	48.8	0.7	50.9	0.5
F2	51.9	3.8	54.7	4.3
G2	48.1	0.0	52.4	2.0
H2	-	-	52.6	2.2
A3	47.1	-1.0	53.6	3.2
B3	47.8	-0.3	52.6	2.2
C3	50.3	2.2	53.0	2.6
D3	47.7	-0.4	50.2	-0.2
E3	47.8	-0.3	49.8	-0.6
F3	48.3	0.2	50.0	-0.4
G3	47.7	-0.4	51.0	0.6
H3	48.4	0.3	52.5	2.1
A4	48.9	0.8	50.4	0.0
B4	47.7	-0.4	49.5	-0.9
C4	47.7	-0.4	49.9	-0.5
D4	47.9	-0.2	47.6	-2.8
E4	47.8	-0.3	50.0	-0.4
F4	47.7	-0.4	49.9	-0.5
G4	47.6	-0.5	53.3	2.9
H4	48.0	-0.1	54.2	3.8
A5	45.3	-2.8	49.8	-0.6
B5	47.8	-0.3	-	-
C5	47.8	-0.3	55.1	4.7
D5	46.8	-1.3	49.7	-0.7
E5	47.7	-0.4	50.5	0.1
F5	48.2	0.1	49.9	-0.5
G5	48.7	0.6	53.3	2.9
H5	47.8	-0.3	49.9	-0.5
A6	47.6	-0.5	49.5	-0.9
B6	-	-	-	-
C6	48.4	0.3	51.2	0.8

Table 7.1: Determined melting temperatures (T_m) of FKBP51^{FK1} F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM) from the TRABITA library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	48.1	-	50.4	-
SAFit1	64.3	16.2	64.3	13.9
D6	47.8	-0.3	50.1	-0.3
E6	48.0	-0.1	52.6	2.2
F6	47.8	-0.3	49.6	-0.8
G6	47.0	-1.1	54.5	4.1
H6	46.4	-1.7	39.8	-10.6
A7	48.0	-0.1	51.5	1.1
B7	47.8	-0.3	50.2	-0.2
C7	-	-	-	-
D7	47.8	-0.3	50.1	-0.3
E7	48.1	0.0	49.7	-0.7
F7	47.7	-0.4	53.0	2.6
G7	48.2	0.1	51.3	0.9
H7	47.6	-0.5	50.0	-0.4
A8	49.9	1.8	52.8	2.4
B8	47.7	-0.4	50.0	-0.4
C8	47.5	-0.6	52.8	2.4
D8	47.4	-0.7	49.4	-1.0
E8	47.8	-0.3	50.0	-0.4
F8	-		-	-
G8	47.8	-0.3	50.1	-0.3
H8	49.5	1.4	52.8	2.4
A9	47.5	-0.6	53.5	3.1
B9	47.7	-0.4	50.3	-0.1
C9	47.7	-0.4	50.2	-0.2
D9	47.3	-0.8	53.4	3.0
E9	47.8	-0.3	50.1	-0.3
F9	47.7	-0.4	53.2	2.8
G9	47.9	-0.2	49.7	-0.7
H9	47.5	-0.6	47.5	-2.9
A10	47.7	-0.4	49.7	-0.7
B10	47.6	-0.5	49.8	-0.6
C10	47.6	-0.5	49.8	-0.6
D10	47.4	-0.7	50.2	-0.2
E10	47.6	-0.5	50.0	-0.4
F10	47.6	-0.5	49.9	-0.5
G10	47.7	-0.4	50.6	0.2

Table 7.2: Determined melting temperatures (T_m) of FKBP51^{FK1} F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM) from the Atomwise library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	47.6	-	50.1	-
SAFit1	64.0	16.4	62.8	12.7
A1	47.7	0.1	50.4	0.3
B1	47.5	-0.1	50.5	0.4
C1	47.6	0.0	53.2	3.1
D1	47.5	-0.1	50.3	0.2
E1	47.6	0.0	50.4	0.3
F1	47.3	-0.3	50.7	0.6
G1	47.4	-0.2	50.2	0.1
H1	48.4	0.8	56.9	6.8
A2	47.5	-0.1	50.6	0.5
B2	47.7	0.1	50.1	0.0
C2	47.5	-0.1	50.0	-0.1
D2	47.2	-0.4	51.2	1.1
E2	45.6	-2.0	49.5	-0.6
F2	46.9	-0.7	50.1	0.0
G2	47.5	-0.1	50.1	0.0
H2	47.6	0.0	50.5	0.4
A3	47.8	0.2	51.0	0.9
B3	48.1	0.5	50.1	0.0
C3	47.6	0.0	50.2	0.1
D3	47.7	0.1	50.4	0.3
E3	47.5	-0.1	50.2	0.1
F3	48.1	0.5	51.8	1.7
G3	48.8	1.2	52.5	2.4
H3	47.2	-0.4	53.1	3.0
A4	48.6	1.0	50.3	0.2
B4	50.7	3.1	53.7	3.6
C4	47.3	-0.3	53.8	3.7
D4	47.8	0.2	50.2	0.1
E4	48.4	0.8	-	-
F4	48.2	0.6	50.2	0.1
G4	48.6	1.0	50.2	0.1
H4	48.0	0.4	50.3	0.2
A5	46.9	-0.7	50.5	0.4
B5	47.5	-0.1	49.6	-0.5
C5	47.5	-0.1	50.4	0.3
D5	47.5	-0.1	50.1	0.0

Table 7.2: Determined melting temperatures (T_m) of FKBP51^{FK1} F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM) from the Atomwise library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	47.6	-	50.1	-
SAFit1	64.0	16.4	62.8	12.7
E5	48.0	0.4	49.7	-0.4
F5	47.8	0.2	50.0	-0.1
G5	47.3	-0.3	51.6	1.5
H5	47.1	-0.5	50.2	0.1
A6	47.2	-0.4	50.0	-0.1
B6	47.3	-0.3	50.0	-0.1
C6	47.5	-0.1	52.9	2.8
D6	48.2	0.6	50.9	0.8
E6	47.3	-0.3	50.2	0.1
F6	47.5	-0.1	50.6	0.5
G6	47.2	-0.4	48.8	-1.3
H6	-	-	-	-
A7	47.6	0.0	50.7	0.6
B7	47.6	0.0	52.4	2.3
C7	49.2	1.6	50.7	0.6
D7	49.0	1.4	51.0	0.9
E7	47.0	-0.6	50.3	0.2
F7	47.3	-0.3	49.9	-0.2
G7	47.4	-0.2	49.8	-0.3
H7	47.7	0.1	51.3	1.2
A8	48.8	1.2	51.0	0.9
B8	52.8	5.2	39.5	-10.6
C8	47.1	-0.5	50.1	0.0
D8	48.0	0.4	50.2	0.1
E8	47.6	0.0	50.8	0.7
F8	48.9	1.3	49.5	-0.6
G8	47.8	0.2	50.1	0.0
H8	47.7	0.1	50.1	0.0
A9	48.7	1.1	51.6	1.5
B9	47.4	-0.2	50.5	0.4
C9	47.5	-0.1	50.5	0.4
D9	47.6	0.0	50.2	0.1
E9	-	-	32.5	-17.6
F9	48.1	0.5	51.9	1.8
G9	47.0	-0.6	50.2	0.1
H9	47.4	-0.2	50.5	0.4

Table 7.2: Determined melting temperatures (T_m) of FKBP51^{FK1} F67Y and F67C/K60C^{ox} (each 60 µM) with ligand/fragments (500 µM) from the Atomwise library. ΔT_m shows the difference in T_m of the DMSO control compared to the corresponding ligand or fragment.

Fragment/Compound	F67Y		F67C/K60C^{ox}	
	T_m/ °C	ΔT_m/ °C	T_m/ °C	ΔT_m/ °C
DMSO	47.6	-	50.1	-
SAFit1	64.0	16.4	62.8	12.7
A10	47.7	0.1	50.8	0.7
B10	48.9	1.3	49.7	-0.4
C10	47.5	-0.1	49.9	-0.2
D10	47.5	-0.1	50.0	-0.1
E10	52.9	5.3	54.6	4.5
F10	49.0	1.4	52.2	2.1
G10	46.9	-0.7	51.8	1.7
H10	46.6	-1.0	50.1	0.0
A11	47.6	0.0	50.1	0.0
B11	52.4	4.8	53.7	3.6
C11	47.3	-0.3	50.0	-0.1
D11	47.8	0.2	51.2	1.1
E11	47.6	0.0	49.8	-0.3
F11	47.1	-0.5	49.8	-0.3
G11	49.4	1.8	53.6	3.5
H11	46.8	-0.8	50.3	0.2
A12	48.0	0.4	48.7	-1.4
B12	47.8	0.2	50.1	0.0
C12	47.4	-0.2	49.4	-0.7
D12	47.4	-0.2	51.3	1.2
E12	47.9	0.3	50.1	0.0
F12	47.4	-0.2	49.7	-0.4
G12	48.0	0.4	50.2	0.1
H12	47.8	0.2	51.2	1.1

A fragment was considered a positive hit if the difference in melting temperature compared to the DMSO control sample was greater than 1.5 °C and the curve displayed the characteristic sigmoidal shape.

8 References

[1] I. H. G. S. Consortium, *nature* **2001**, *409*, 860–921.

[2] L. M. Smith, N. L. Kelleher, *Nature methods* **2013**, *10*, 186–187.

[3] L. M. Smith, J. N. Agar, J. Chamot-Rooke, P. O. Danis, Y. Ge, J. A. Loo, L. Paša-Tolić, Y. O. Tsybin, N. L. Kelleher, Consortium for Top-Down Proteomics, *Science Advances* **2021**, *7*, eabk0734.

[4] J. Loscalzo, *The FASEB Journal* **2023**, *37*, e22660.

[5] J. Drews, *science* **2000**, *287*, 1960–1964.

[6] A. L. Hopkins, C. R. Groom, *Nature reviews Drug discovery* **2002**, *1*, 727–730.

[7] A. P. Russ, S. Lampel, *Drug discovery today* **2005**, *10*, 1607–1610.

[8] C. J. Radoux, F. Vianello, J. McGreig, N. Desai, A. Bradley, *Frontiers in Bioinformatics* **2022**, 84.

[9] R. Fischer, *nature* **1894**, *50*, 272–280.

[10] Koshland Jr de, *Proceedings of the National Academy of Sciences* **1958**, *44*, 98–104.

[11] D. Plewczyński, K. Ginalski, *Cellular & molecular biology letters* **2009**, *14*, 1–22.

[12] T. M. Przytycka, M. Singh, D. K. Slonim, *Briefings in bioinformatics* **2010**, *11*, 15–29.

[13] X. Du, Y. Li, Y.-L. Xia, S.-M. Ai, J. Liang, P. Sang, X.-L. Ji, S.-Q. Liu, *International journal of molecular sciences* **2016**, *17*, 144.

[14] G. Schreiber, A. E. Keating, *Current opinion in structural biology* **2011**, *21*, 50–61.

[15] D. M. Jameson, J. C. Croney, *Combinatorial chemistry & high throughput screening* **2003**, *6*, 167–176.

[16] A. M. Rossi, C. W. Taylor, *Nature protocols* **2011**, *6*, 365–387.

[17] W. Becker, K. C. Bhattiprolu, N. Gubensäk, K. Zangger, *ChemPhysChem* **2018**, *19*, 895–906.

[18] E. Freire, O. L. Mayorga, M. Straume, *Analytical chemistry* **1990**, *62*, 950A–959A.

[19] S. G. Patching, *Biochimica et Biophysica Acta (BBA)-Biomembranes* **2014**, *1838*, 43–55.

[20] A. D. Vogt, E. Di Cera, *Biochemistry* **2012**, *51*, 5894–5902.

[21] Koshland Jr de, *Proceedings of the National Academy of Sciences* **1958**, *44*, 98–104.

[22] D. E. Koshland Jr, *Angewandte Chemie International Edition in English* **1995**, *33*, 2375–2378.

[23] J.-P. Changeux, S. Edelstein, *F1000 biology reports* **2011**, *3*.

[24] P. Csermely, R. Palotai, R. Nussinov, *Nature Precedings* **2010**, 1.

[25] A. D. Vogt, E. Di Cera, *Biochemistry* **2013**, *52*, 5723–5729.

[26] A. D. Vogt, N. Pozzi, Z. Chen, E. Di Cera, *Biophysical chemistry* **2014**, *186*, 13–21.

[27] P. J. Tummino, R. A. Copeland, *Biochemistry* **2008**, *47*, 5481–5492.

[28] A. Stank, D. B. Kokh, J. C. Fuller, R. C. Wade, *Accounts of Chemical Research* **2016**, *49*, 809–815.

[29] M. R. Arkin, M. Randal, W. L. DeLano, J. Hyde, T. N. Luong, J. D. Oslob, D. R. Raphael, L. Taylor, J. Wang, R. S. McDowell, *Proceedings of the National Academy of Sciences* **2003**, *100*, 1603–1608.

[30] J. Hyde, A. C. Braisted, M. Randal, M. R. Arkin, *Biochemistry* **2003**, *42*, 6475–6483.

[31] P. Priya, A. Maity, S. Ghosh Dastidar, *Proteins: Structure Function and Bioinformatics* **2017**, *85*, 1567–1579.

[32] B. E. Sleebs, P. E. Czabotar, W. J. Fairbrother, W. D. Fairlie, J. A. Flygare, D. C. S. Huang, W. J. A. Kersten, M. F. T. Koehler, G. Lessene, K. Lowes, *Journal of medicinal chemistry* **2011**, *54*, 1914–1926.

[33] G. Lessene, P. E. Czabotar, B. E. Sleebs, K. Zobel, K. N. Lowes, J. M. Adams, J. B. Baell, P. M. Colman, K. Deshayes, W. J. Fairbrother, *Nature chemical biology* **2013**, *9*, 390–397.

[34] S. Tassin–Moindrot, A. Caille, J.-.-P. Douliez, D. Marion, F. Vovelle, *European Journal of Biochemistry* **2000**, *267*, 1117–1124.

[35] C. Pargellis, L. Tong, L. Churchill, P. F. Cirillo, T. Gilmore, A. G. Graham, P. M. Grob, E. R. Hickey, N. Moss, S. Pav, *Nature structural biology* **2002**, *9*, 268–272.

[36] C. B. Kang, Y. Hong, S. Dhe-Paganon, H. S. Yoon, *Neurosignals* **2008**, *16*, 318–325.

[37] B. M. Dunyak, J. E. Gestwicki, *Journal of medicinal chemistry* **2016**, *59*, 9622–9644.

[38] X. Feng, S. Pomplun, F. Hausch, *Current Molecular Pharmacology* **2016**, *9*, 27–36.

[39] K. A. Ellsworth, L. Wang, **2014**.

[40] A. Hähle, S. Merz, C. Meyners, F. Hausch, *Biomolecules* **2019**, *9*, 35.

[41] J. A. Somarelli, S. Y. Lee, J. Skolnick, R. J. Herrera, *Proteins: Structure Function and Bioinformatics* **2008**, *72*, 197–208.

[42] L. A. Stechschulte, E. R. Sanchez, *Current opinion in pharmacology* **2011**, *11*, 332–337.

[43] M. V. Schmidt, M. Paez–Pereda, F. Holsboer, F. Hausch, *ChemMedChem* **2012**, *7*, 1351–1359.

[44] C. R. Sinars, J. Cheung-Flynn, R. A. Rimerman, J. G. Scammell, D. F. Smith, J. Clardy, *Proceedings of the National Academy of Sciences* **2003**, *100*, 868–873.

[45] A. Bracher, C. Kozany, A.-K. Thost, F. Hausch, *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 549–559.

[46] R. Kumar, M. Moche, B. Winblad, P. F. Pavlov, *Scientific reports* **2017**, *7*, 14288.

[47] J. Hartmann, K. V. Wagner, C. Liebl, S. H. Scharf, X.-D. Wang, M. Wolf, F. Hausch, T. Rein, U. Schmidt, C. Touma, *Neuropharmacology* **2012**, *62*, 332–339.

[48] G. M. Wochnik, J. Ruegg, G. A. Abel, U. Schmidt, F. Holsboer, T. Rein, *Journal of Biological Chemistry* **2005**, *280*, 4609–4616.

[49] E. B. Binder, D. Salyakina, P. Lichtner, G. M. Wochnik, M. Ising, B. Pütz, S. Papiol, S. Seaman, S. Lucae, M. A. Kohli, *Nature genetics* **2004**, *36*, 1319–1325.

[50] J. G. Scammell, W. B. Denny, D. L. Valentine, D. F. Smith, *General and comparative endocrinology* **2001**, *124*, 152–165.

[51] E. B. Binder, *Psychoneuroendocrinology* **2009**, *34*, S186–S195.

[52] A. S. Zannas, E. B. Binder, *Genes Brain and Behavior* **2014**, *13*, 25–37.

[53] M. Maiarù, K. K. Tochiki, M. B. Cox, L. V. Annan, C. G. Bell, X. Feng, F. Hausch, S. M. Géranton, *Science translational medicine* **2016**, *8*, 325ra19–325ra19.

[54] N. Matosin, T. Halldorsdottir, E. B. Binder, *Biological psychiatry* **2018**, *83*, 821–830.

[55] Y. Hernandez-Diaz, T. B. González-Castro, C. A. Tovilla-Zárate, I. E. Juárez-Rojop, M. L. López-Narváez, N. Perez-Hernandez, J. M. Rodríguez-Pérez, A. D. Genis-Mendoza, *Psychiatry research* **2019**, *271*, 658–668.

[56] L. A. Stechschulte, B. Qiu, M. Warrier, T. D. Hinds Jr, M. Zhang, H. Gu, Y. Xu, S. S. Khuder, L. Russo, S. M. Najjar u. a., *Endocrinology* **2016**, *157*, 3888–3900.

[57] G. Balsevich, A. S. Häusl, C. W. Meyer, S. Karamihalev, X. Feng, M. L. Pöhlmann, C. Dournes, A. Uribe-Marino, S. Santarelli, C. Labermaier, *Nature Communications* **2017**, *8*, 1725.

[58] M. Maiarù, O. B. Morgan, T. Mao, M. Breitsamer, H. Bamber, M. Pöhlmann, M. V. Schmidt, G. Winter, F. Hausch, S. M. Géranton, *Pain* **2018**, *159*, 1224.

[59] S. Gaali, A. Kirschner, S. Cuboni, J. Hartmann, C. Kozany, G. Balsevich, C. Namendorf, P. Fernandez-Vizarra, C. Sippel, A. S. Zannas, *Nature chemical biology* **2015**, *11*, 33–37.

[60] X. Feng, C. Sippel, A. Bracher, F. Hausch, *Journal of medicinal chemistry* **2015**, *58*, 7796–7806.

[61] X. Barril, *Wiley Interdisciplinary Reviews: Computational Molecular Science* **2013**, *3*, 327–338.

[62] F. D. Makurvet, *Medicine in Drug Discovery* **2021**, *9*, 100075.

[63] P. J. Declerck, *GaBI J* **2012**, *1*, 13–16.

[64] S. Vajda, D. Beglov, A. E. Wakefield, M. Egbert, A. Whitty, *Current opinion in chemical biology* **2018**, *44*, 1–8.

[65] D. B. Kokh, P. Czodrowski, F. Rippmann, R. C. Wade, *Journal of Chemical Theory and Computation* **2016**, *12*, 4100–4113.

[66] G. Rodgers, C. Austin, J. Anderson, A. Pawlyk, C. Colvis, R. Margolis, J. Baker, *Nature reviews Drug discovery* **2018**, *17*, 301–302.

[67] Y. Yuan, J. Pei, L. Lai, *Current pharmaceutical design* **2013**, *19*, 2326–2333.

[68] A. Volkamer, D. Kuhn, T. Grombacher, F. Rippmann, M. Rarey, *Journal of chemical information and modeling* **2012**, *52*, 360–372.

[69] X. Zheng, L. Gan, E. Wang, J. Wang, *The AAPS journal* **2013**, *15*, 228–241.

[70] A. Stank, D. B. Kokh, M. Horn, E. Sizikova, R. Neil, J. Panecka, S. Richter, R. C. Wade, *Nucleic Acids Research* **2017**, *45*, W325–W330.

[71] C. W. Murray, D. C. Rees, *Nature Chemistry* **2009**, *1*, 187–192.

[72] T. H. Keller, A. Pichota, Z. Yin, *Current opinion in chemical biology* **2006**, *10*, 357–361.

[73] H. Kubinyi, *Die Pharmazie* **1995**, *50*, 647–662.

[74] P. Kirsch, A. M. Hartman, A. K. H. Hirsch, M. Empting, *Molecules* **2019**, *24*, 4309.

[75] P. J. Hajduk, J. Greer, *Nature reviews Drug discovery* **2007**, *6*, 211–219.

[76] M. Bartolowits, V. J. Davisson, Considerations of Protein Subpockets in Fragment–Based Drug Design, **2016**.

[77] T. Wang, W. H. Bisson, P. Mäser, L. Scapozza, D. Picard, *Journal of medicinal chemistry* **2014**, *57*, 2524–2535.

[78] G. A. Petit, B. Mohanty, R. M. McMahon, S. Nebl, D. H. Hilko, K. L. Wilde, M. J. Scanlon, J. L. Martin, M. A. Halili, *Acta Crystallographica Section D: Structural Biology* **2022**, *78*, 75–90.

[79] D. Joseph-McCarthy, A. J. Campbell, G. Kern, D. Moustakas, *Journal of chemical information and modeling* **2014**, *54*, 693–704.

[80] D. Nguyen, X. Xie, S. Jakobi, F. Terwesten, A. Metz, T. X. P. Nguyen, V. A. Palchykov, A. Heine, K. Reuter, G. Klebe, *ACS Chemical Biology* **2021**, *16*, 1090–1098.

[81] E. A. Hoyt, P. M. Cal, B. L. Oliveira, G. J. L. Bernardes, *Nature Reviews Chemistry* **2019**, *3*, 147–171.

[82] S. S. Wong, D. M. Jameson, *Chemistry of protein and nucleic acid cross-linking and conjugation*, CRC Press, **2011**.

[83] K. Maruyama, H. Nagasawa, A. Suzuki, *Peptides* **1999**, *20*, 881–884.

[84] A. Habeeb in *Methods in enzymology*, *Bd. 25*, Elsevier, **1972**, S. 457–464.

[85] M. B. Hansen, F. Hubálek, T. Skrydstrup, T. Hoeg–Jensen, *Chemistry–A European Journal* **2016**, *22*, 1572–1576.

[86] J. Ohata, M. K. Miller, C. M. Mountain, F. Vohidov, Z. T. Ball, *Angewandte Chemie International Edition* **2018**, *57*, 2827–2830.

[87] D. Hung-ChieháChou, *Chemical Science* **2017**, *8*, 2717–2722.

[88] L. Coleman, J. Bork, H. Dunn, *The Journal of Organic Chemistry* **1959**, *24*, 135–136.

[89] S. S. Husain, G. Lowe, *Biochemical Journal* **1968**, *108*, 855–859.

[90] H. Ozawa, *The Journal of Biochemistry* **1967**, *62*, 531–536.

[91] D. M. Segal, E. Hurwitz, *Biochemistry* **1976**, *15*, 5253–5258.

[92] R. F. Ludueña, M. C. Roach, P. P. Trcka, S. Weintraub, *Analytical Biochemistry* **1981**, *117*, 76–80.

[93] S. B. Gunnoo, A. Madder, *ChemBioChem* **2016**, *17*, 529–553.

[94] R. Thomas, *Chemical Communications* **2011**, *47*, 12068–12070.

[95] I. Annis, B. Hargittai, G. Barany, *Methods in enzymology* **1997**, *289*, 198–221.

[96] G. Bulaj, *Biotechnology advances* **2005**, *23*, 87–92.

[97] V. R. Pattabiraman, J. W. Bode, *nature* **2011**, *480*, 471–479.

[98] W. Wang, S. C. Khojasteh, D. Su, *Molecules* **2021**, *26*, 3338.

[99] G. K. T. Nguyen, A. Kam, S. Loo, A. E. Jansson, L. X. Pan, J. P. Tam, *Journal of the American Chemical Society* **2015**, *137*, 15398–15401.

[100] A. Russo, C. Aiello, P. Grieco, D. Marasco, *Current Medicinal Chemistry* **2016**, *23*, 748–762.

[101] K. Tarasava, E. Freisinger, *Protein Engineering Design & Selection* **2014**, *27*, 481–488.

[102] K. Ha, J.-C. M. Monbaliu, B. C. Williams, G. G. Pillai, C. E. Ocampo, M. Zeller, C. V. Stevens, A. R. Katritzky, *Organic & Biomolecular Chemistry* **2012**, *10*, 8055–8058.

[103] Z. Wu, X. Guo, Z. Guo, *Chemical Communications* **2011**, *47*, 9218–9220.

[104] S. H. Joo, *Biomolecules & therapeutics* **2012**, *20*, 19.

[105] M. Zheng, L. Zheng, P. Zhang, J. Li, Y. Zhang, *Molecules* **2015**.

[106] L. Wang, A. Brock, B. Herberich, P. G. Schultz, *Science* **2001**, *292*, 498–500.

[107] M. A. Shandell, Z. Tan, V. W. Cornish, *Biochemistry* **2021**, *60*, 3455–3469.

[108] P. K. Mishra, C.-M. Yoo, E. Hong, H. W. Rhee, *ChemBioChem* **2020**, *21*, 924–932.

[109] I. Coin, *Current opinion in chemical biology* **2018**, *46*, 156–163.

[110] E. Smith, I. Collins, *Future medicinal chemistry* **2015**, *7*, 159–183.

[111] J. W. Chin, S. W. Santoro, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Journal of the American Chemical Society* **2002**, *124*, 9026–9027.

[112] M. M. Hassan, O. O. Olaoye, *Molecules* **2020**, *25*, 2285.

[113] T.-A. Nguyen, M. Cigler, K. Lang, *Angewandte Chemie* **2018**, *130*, 14548–14559.

[114] J. Brunner, H. Senn, F. Richards, *Journal of Biological Chemistry* **1980**, *255*, 3313–3318.

[115] H. C. Kolb, M. G. Finn, K. B. Sharpless, *Angewandte Chemie International Edition* **2001**, *40*, 2004–2021.

[116] R. Huisgen, *Angewandte Chemie International Edition in English* **1963**, *2*, 565–598.

[117] D. Clarke, R. W. Mares, H. McNab, *Journal of the Chemical Society Perkin Transactions 1* **1997**, 1799–1804.

[118] V. V. Rostovtsev, L. G. Green, V. V. Fokin, K. B. Sharpless, *Angewandte Chemie* **2002**, *114*, 2708–2711.

[119] C. W. Tornøe, C. Christensen, M. Meldal, *The Journal of Organic Chemistry* **2002**, *67*, 3057–3064.

[120] L. Zhang, X. Chen, P. Xue, H. H. Y. Sun, I. D. Williams, K. B. Sharpless, V. V. Fokin, G. Jia, *Journal of the American Chemical Society* **2005**, *127*, 15998–15999.

[121] C. R. Becer, R. Hoogenboom, U. S. Schubert, *Angewandte Chemie International Edition* **2009**, *48*, 4900–4908.

[122] N. J. Agard, J. A. Prescher, C. R. Bertozzi, *Journal of the American Chemical Society* **2004**, *126*, 15046–15047.

[123] J. Zaia, *Analytical and Bioanalytical Chemistry* **2023**, *415*, 527–532.

[124] E. L. Rodriguez, S. Poddar, S. Iftekhar, K. Suh, A. G. Woolfork, S. Ovbude, A. Pekarek, M. Walters, S. Lott, D. S. Hage, *Journal of Chromatography B* **2020**, *1157*, 122332.

[125] S. Ziegler, V. Pries, C. Hedberg, H. Waldmann, *Angewandte Chemie International Edition* **2013**, *52*, 2744–2792.

[126] P. Cuatrecasas, M. Wilchek, C. B. Anfinsen, *Proceedings of the National Academy of Sciences* **1968**, *61*, 636–643.

[127] S. Magdeldin, *Affinity chromatography*, BoD–Books on Demand, **2012**.

[128] T. Kline, *Handbook of affinity chromatography*, CRC Press, **1993**.

[129] G. Leriche, L. Chisholm, A. Wagner, *Bioorganic & medicinal chemistry* **2012**, *20*, 571–582.

[130] H. A. Beard, D. Korovesis, S. Chen, S. H. L. Verhelst, *Molecular Omics* **2021**, *17*, 197–209.

[131] P. K. A. Jagtap, S. Asami, C. Sippel, V. R. Kaila, F. Hausch, M. Sattler, *Angewandte Chemie* **2019**, *131*, 9529–9533.

[132] J. A. Lerma Romero, C. Meyners, A. Christmann, L. M. Reinbold, A. Charalampidou, F. Hausch, H. Kolmar, *Frontiers in Molecular Biosciences* **2022**, 1245.

[133] V. Gabelica, E. Marklund, *Current opinion in chemical biology* **2018**, *42*, 51–59.

[134] S. D. Pringle, K. Giles, J. L. Wildgoose, J. P. Williams, S. E. Slade, K. Thalassinos, R. H. Bateman, M. T. Bowers, J. H. Scrivens, *International Journal of Mass Spectrometry* **2007**, *261*, 1–12.

[135] A. A. Shvartsburg, R. D. Smith, *Analytical chemistry* **2008**, *80*, 9689–9699.

[136] B. T. Ruotolo, J. L. P. Benesch, A. M. Sandercock, S.-J. Hyung, C. V. Robinson, *Nature protocols* **2008**, *3*, 1139–1152.

[137] J. W. Chin, A. B. Martin, D. S. King, L. Wang, P. G. Schultz, *Proceedings of the National Academy of Sciences* **2002**, *99*, 11020–11024.

[138] S. I. Presolski, V. P. Hong, M. G. Finn, *Current protocols in chemical biology* **2011**, *3*, 153–162.

[139] V. Hong, S. I. Presolski, C. Ma, M. â. G. Finn, *Angewandte Chemie* **2009**, *121*, 10063–10067.

[140] L. Wang, J. Xie, P. G. Schultz, *Annu. Rev. Biophys. Biomol. Struct.* **2006**, *35*, 225–249.

[141] W. Gao, N. Bu, Y. Lu, *Methods and protocols* **2019**, *2*, 16.

[142] A. J. Mijalis, D. A. Thomas III, M. D. Simon, A. Adamo, R. Beaumont, K. F. Jensen, B. L. Pentelute, *Nature chemical biology* **2017**, *13*, 464–466.

[143] N. Hartrampf, A. Saebi, M. Poskus, Z. P. Gates, A. J. Callahan, A. E. Cowfer, S. Hanna, S. Antilla, C. K. Schissel, A. J. Quartararo u. a., *Science* **2020**, *368*, 980–987.

[144] J. E. Syka, J. J. Coon, M. J. Schroeder, J. Shabanowitz, D. F. Hunt, *Proceedings of the National Academy of Sciences* **2004**, *101*, 9528–9533.

[145] P. ROEPSTORFE, *Biomed. Mass Spectrom.* **1984**, *11*, 601–605.

[146] K. Biemann, *Biomedical & environmental mass spectrometry* **1988**, *16*, 99–111.

[147] G. C. McAlister, W. T. Berggren, J. Griep-Raming, S. Horning, A. Makarov, D. Phanstiel, G. Stafford, D. L. Swaney, J. E. Syka, V. Zabrouskov u. a., *Journal of proteome research* **2008**, *7*, 3127–3136.

[148] C. Kozany, A. März, C. Kress, F. Hausch, *ChemBioChem* **2009**, *10*, 1402–1410.

[149] S. N. Mthembu, A. Sharma, F. Albericio, B. G. de La Torre, *ChemBioChem* **2020**, *21*, 1947–1954.

[150] F. H. Knaup, C. Meyners, W. O. Sugiarto, S. Wedel, M. Springer, C. Walz, T. M. Geiger, M. Schmidt, M. Sisignano, F. Hausch, *Journal of Medicinal Chemistry* **2023**, *66*, 5965–5980.

[151] K. L. Rundlett, D. W. Armstrong, *Analytical chemistry* **1996**, *68*, 3493–3497.

[152] D. Botelho, M. J. Wall, D. B. Vieira, S. Fitzsimmons, F. Liu, A. Doucette, *Journal of proteome research* **2010**, *9*, 2863–2870.

[153] B. J. Peiffer, L. Qi, A. R. Ahmadi, Y. Wang, Z. Guo, H. Peng, Z. Sun, J. O. Liu, *Cell chemical biology* **2019**, *26*, 652–661.

[154] O. Edelheit, A. Hanukoglu, I. Hanukoglu, *BMC biotechnology* **2009**, *9*, 1–8.

[155] Z.-X. Wang, N. R. Kumar, D. K. Srivastava, *Analytical biochemistry* **1992**, *206*, 376–381.

[156] Z.-X. Wang, *FEBS letters* **1995**, *360*, 111–114.

[157] Z. Wu, D. S. Roberts, J. A. Melby, K. Wenger, M. Wetzel, Y. Gu, S. G. Ramanathan, E. F. Bayne, X. Liu, R. Sun, *Journal of proteome research* **2020**, *19*, 3867–3876.

[158] D. A. Polasky, S. M. Dixit, S. M. Fantin, B. T. Ruotolo, *Analytical chemistry* **2019**, *91*, 3147–3155.

[159] M. Gerlach, U. Mueller, M. S. Weiss, *Journal of large-scale research facilities JLSRF* **2016**, *2*, A47–A47.

[160] L. Potterton, J. Agirre, C. Ballard, K. Cowtan, E. Dodson, P. R. Evans, H. T. Jenkins, R. Keegan, E. Krissinel, K. Stevenson, *Acta Crystallographica Section D: Structural Biology* **2018**, *74*, 68–84.

[161] E. Potterton, P. Briggs, M. Turkenburg, E. Dodson, *Acta Crystallographica Section D: Biological Crystallography* **2003**, *59*, 1131–1137.

[162] C. C. Project, *Acta Crystallographica Section D: Biological Crystallography* **1994**, *50*, 760–763.

[163] M. D. Winn, C. C. Ballard, K. D. Cowtan, E. J. Dodson, P. Emsley, P. R. Evans, R. M. Keegan, E. B. Krissinel, A. G. W. Leslie, A. McCoy, *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 235–242.

[164] P. R. Evans, *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 282–292.

[165] P. R. Evans, G. N. Murshudov, *Acta crystallographica. Section D Biological crystallography* **2013**, *69*, 1204–1214.

[166] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Journal of applied crystallography* **2007**, *40*, 658–674.

[167] A. A. Vagin, R. A. Steiner, A. A. Lebedev, L. Potterton, S. McNicholas, F. Long, G. N. Murshudov, *Acta Crystallographica Section D: Biological Crystallography* **2004**, *60*, 2184–2195.

[168] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Acta Crystallographica Section D: Biological Crystallography* **2010**, *66*, 486–501.

[169] G. N. Murshudov, P. Skubák, A. A. Lebedev, N. S. Pannu, R. A. Steiner, R. A. Nicholls, M. D. Winn, F. Long, A. A. Vagin, *Acta Crystallographica Section D: Biological Crystallography* **2011**, *67*, 355–367.

[170] G. N. Murshudov, A. A. Vagin, E. J. Dodson, *Acta Crystallographica Section D: Biological Crystallography* **1997**, *53*, 240–255.

[171] R. A. Nicholls, F. Long, G. N. Murshudov, *Acta Crystallographica Section D: Biological Crystallography* **2012**, *68*, 404–417.

[172] M. D. Winn, G. N. Murshudov, M. Z. Papiz in *Methods in enzymology*, *Bd. 374*, Elsevier, **2003**, S. 300–321.

[173] D. M. Van Aalten, R. Bywater, J. B. Findlay, M. Hendlich, R. W. Hooft, G. Vriend, *Journal of computer-aided molecular design* **1996**, *10*, 255–262.

[174] F. Long, R. A. Nicholls, P. Emsley, S. Gražulis, A. Merkys, A. Vaitkus, G. N. Murshudov, *Acta Crystallographica Section D: Structural Biology* **2017**, *73*, 112–122.