

# **Structure-Based Design of a Blood Coagulation Factor XIII Blocker**

## **Dissertation**

zur Erlangung des Doktorgrades  
der Naturwissenschaften  
(Dr. rer. nat.)

dem  
Fachbereich Pharmazie  
der Philipps-Universität Marburg  
vorgelegt von

Diplom-Chemiker

**Martin Stieler**

aus Marburg

Marburg/Lahn 2017

Dem Fachbereich Pharmazie der Philipps-Universität Marburg  
als Dissertation eingereicht am: 08.01.2017

Erstgutachter: Prof. Dr. Gerhard Klebe

Zweitgutachter: Prof. Dr. Peter Kolb

Tag der mündlichen Prüfung: 23.02.2017

Hochschulkennziffer: 1180

Die Untersuchungen zur vorliegenden Arbeit wurden auf Anregung von Herrn Prof. Dr. Gerhard Klebe am Institut für Pharmazeutische Chemie des Fachbereiches Pharmazie der Philipps-Universität Marburg in der Zeit von November 2011 bis April 2016 durchgeführt.





# Table of Contents

<b>1</b>	<b>Introduction .....</b>	<b>10</b>
1.1	Introductory Remarks.....	10
1.2	Transglutaminases.....	10
1.3	Blood Coagulation Factor XIII .....	14
1.4	Structure-Based Drug Design .....	16
<b>2</b>	<b>Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants .....</b>	<b>18</b>
2.1	Introductory Remarks.....	18
2.2	Introduction.....	18
2.3	Results & Discussion.....	23
2.3.1	General Structural Features and Overall Structural Rearrangements .....	23
2.3.2	Structural Rearrangements Triggered by Calcium Binding .....	28
2.3.3	Binding Mode of ZED1301.....	33
2.3.4	Modeling of a Co-substrate Mimetic.....	36
2.3.5	Catalytic Dyad.....	37
2.4	Summary & Conclusion .....	39
2.5	Experimental Part.....	40
2.5.1	Methods Protein Expression, Purification and Inhibition .....	40
2.5.2	Crystallization, Data Collection and Structure Determination .....	40
2.5.3	Molecular Modelling .....	41
2.5.4	Synthesis of ZED1301 .....	42
2.6	Appendix.....	43
2.6.1	Alignment of the Catalytic Domain of Human Transglutaminases .....	43
2.6.2	Crystallographic Table .....	45

---

<b>3</b>	<b>Structure-Based Design of FXIIIa-Blockers: Addressing the Hydrophobic Pocket of the “<math>\alpha</math>-Space”</b>	<b>48</b>
3.1	Introduction.....	48
3.2	Results & Discussion.....	52
3.2.1	ZED1630.....	52
3.2.2	ZED2360.....	55
3.2.3	ZED2369.....	58
3.2.4	Indole Ring of the Hydrophobic Tunnel Affected by Calcium Binding .....	60
3.3	Summary & Conclusion .....	66
3.4	Experimental Part.....	68
3.4.1	Data Collection and Structure Determination.....	68
3.5	Appendix.....	69
3.5.1	Crystallographic Table .....	69
<b>4</b>	<b>Blood Coagulation Factor XIII – An Enzyme with a Highly Adaptive Binding Site .....</b>	<b>73</b>
4.1	Introduction.....	73
4.2	Results & Discussion.....	81
4.2.1	Binding Mode of the Inhibitor Mi0621.....	81
4.2.2	The Three-Stranded $\beta$ -Sheet Adopts Different Conformations .....	89
4.2.3	Consensus sequence for FXIII substrates .....	92
4.3	Summary & Conclusion .....	95
4.4	Experimental Part.....	97
4.4.1	Crystallization, Data Collection and Structure Determination .....	97
4.5	Appendix.....	98
4.5.1	Crystallographic Tables.....	98
<b>5</b>	<b>Does TG2 Adopt Two Different Active States? .....</b>	<b>104</b>
5.1	Introduction.....	104

---

5.2	Comparison of FXIIIa <sup>o</sup> and TG2a* .....	107
5.2.1	Overall Structure .....	107
5.2.2	Calcium Binding Sites.....	109
5.2.3	Active Site .....	111
5.3	Comparison of FXIIIa <sup>o</sup> and TG3a <sup>c</sup> .....	115
5.3.1	Overall Structure .....	115
5.3.2	Calcium Binding Sites.....	116
5.3.3	Active Site .....	118
5.4	Comparison of FXIIIa <sup>o</sup> and a Homology Model of TG2a Based on the Structure of FXIIIa <sup>o</sup> . ..	121
5.4.1	Overall Structure .....	121
5.4.2	Calcium Binding Sites.....	122
5.4.3	Active Site .....	124
5.5	Summary & Conclusion .....	131
5.6	Experimental Part.....	133
5.6.1	Homology Model Building of TG2a .....	133
5.7	Appendix.....	134
<b>6</b>	<b>Crystallographic Fragment Screening Using the Example of the Surrogate Protease Endothiapepsin.....</b>	<b>136</b>
6.1	Introduction.....	136
6.2	Results & Discussion.....	139
6.2.1	Overview.....	139
6.2.2	Fragments Addressing the S <sub>1</sub> '-S <sub>3</sub> ' Subpockets.....	141
6.2.3	Fragments Bind Remotely from the Active Site .....	145
6.3	Summary & Conclusion .....	152
6.4	Experimental Part.....	154
6.4.1	Processing and Refinement.....	154

6.5	Appendix.....	155
6.5.1	Difference Electron Density Map (2.5 sigma).....	155
6.5.2	Crystallographic Table .....	157
<b>7</b>	<b>Summary and Outlook (Transglutaminases).....</b>	<b>163</b>
<b>8</b>	<b>Zusammenfassung und Ausblick (Transglutaminasen).....</b>	<b>171</b>
	<b>References .....</b>	<b>180</b>
	<b>Abbreviations.....</b>	<b>187</b>
	<b>One- and three-letter code of proteinogenic amino acids .....</b>	<b>190</b>
	<b>Publications and Patents .....</b>	<b>191</b>
	<b>Conference Presentations.....</b>	<b>192</b>
	<b>Danksagung .....</b>	<b>193</b>
	<b>Lebenslauf .....</b>	<b>195</b>
	<b>Eidesstattliche Erklärung .....</b>	<b>196</b>



# 1 Introduction

## 1.1 Introductory Remarks

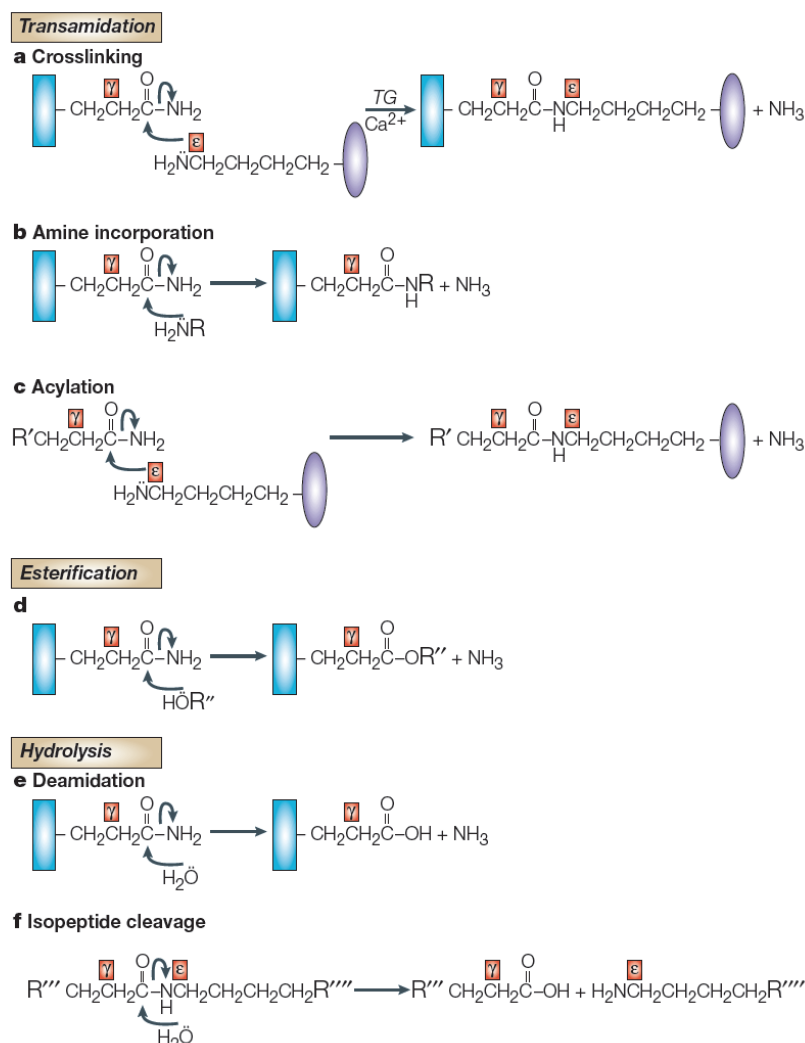
This thesis deals predominantly with the structure-based design of inhibitors of blood coagulation factor XIII (FXIII) which appears to be an auspicious target for an anticoagulant with a lower bleeding risk compared to current anticoagulants. A more detailed elucidation about the relevance of FXIII as a target for novel anticoagulants is given in the introduction of *Chapter 2* and *Chapter 3*. Information detailing the chemical environment of the active site, provided by the crystal structures of FXIII in complex with different inhibitors, was used in collaboration with the biotech company Zedira GmbH to improve the affinity and size of the lead compound (*Chapter 2-4*).

Moreover, the crystal structures of FXIII in complex with different inhibitors reveal new detailed information about mechanistic features of FXIII and its isozymes (transglutaminases), especially, concerning the involvement of calcium in the activation process (*Chapter 5*). This in turn gives valuable information for the development of drugs addressing transglutaminases, as well.

In *Chapter 6*, an academic collaboration projects dealing with a crystallographic fragment screening will be described.

## 1.2 Transglutaminases

Transglutaminases (TGs) represent an enzyme class catalyzing the covalent linkage of proteins/peptides by transamidation between the  $\gamma$ -carboxamide group of glutamine and the  $\epsilon$ -amino group of lysine (*Fig. 1.1a*).<sup>[1]</sup> As substrates also other amines and carboxamides can serve with an appropriate distance to the functional group. Thus, TGs also catalyze the incorporation of amines into proteins (*Fig. 1.1b*) and the acylation of proteins (*Fig. 1.1c*). Furthermore, TGs catalyze the esterification of carboxamides (*Fig. 1.1d*) and the hydrolysis of carboxamides (*Fig. 1.1e*) and isopeptide bonds (*Fig. 1.1f*).



**Fig. 1.1:** Reactions catalyzed by transglutaminases. Image taken from Lorand and Graham.<sup>[1]</sup>

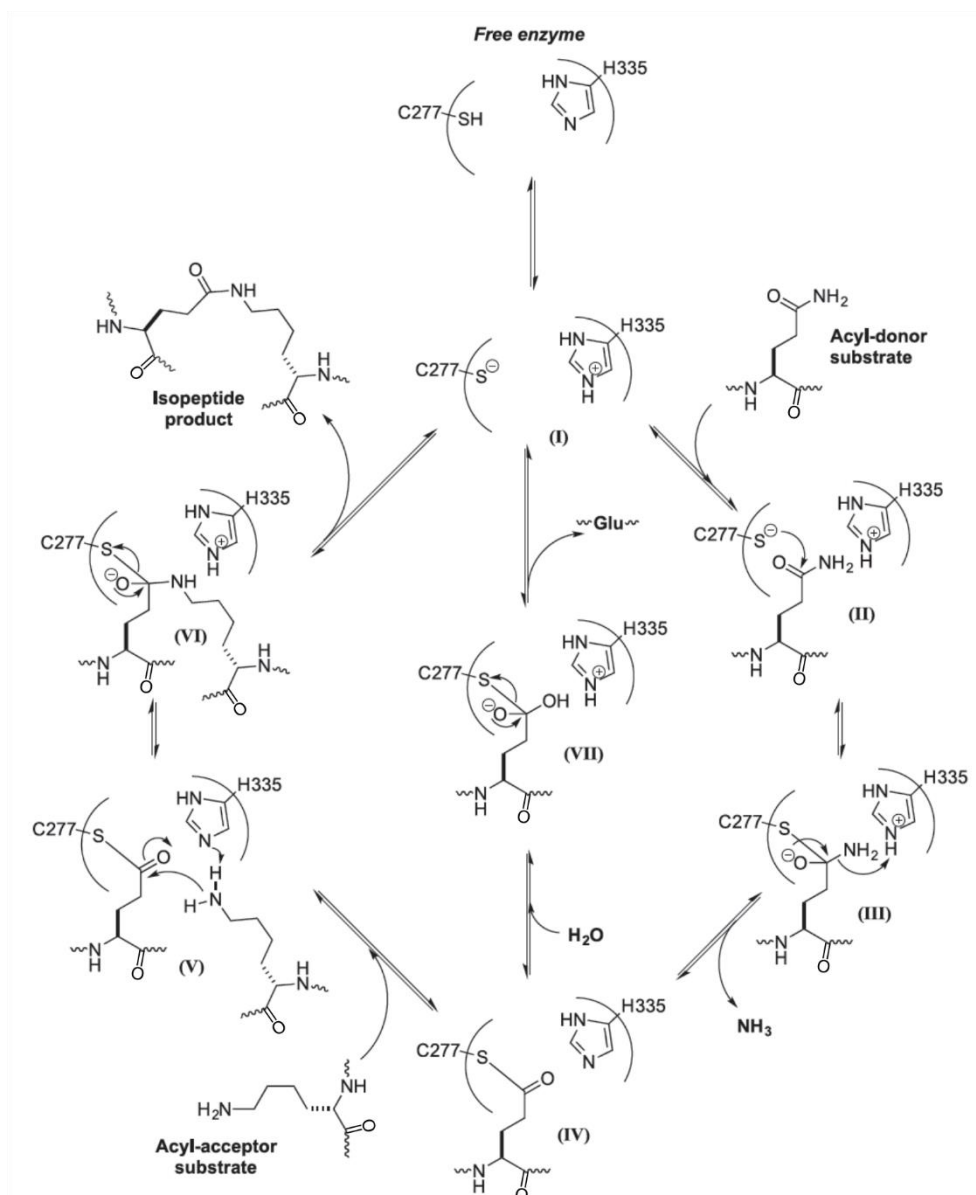
In total, there are eight human catalytically active transglutaminases (TG1-TG7 and FXIII). The prevalent physiological function of TGs is to improve the stability of several components of the organism by protein cross-linking.<sup>[1]</sup> For example, TG2 is involved in the assembly and remodeling of the extracellular matrix<sup>[2]</sup>, TG1, TG3 and TG5 are involved in the formation of the cornified cell envelope<sup>[3-5]</sup> and FXIII stabilizes blood clots.<sup>[6]</sup> Moreover, with their ability to modify proteins post-translationally (Fig. 1.1), TGs exhibit a regulative function as well<sup>[1, 7]</sup>, whereas TG2 can even act as a G-protein<sup>[8]</sup> and kinase.<sup>[9, 10]</sup> Finally, TG2 and FXIII possess also protein disulfide isomerase activity.<sup>[11, 12]</sup> Owing to their versatile reactivity and ubiquity (intracellular and extracellular), particularly, TG2 is expressed in almost all cell types<sup>[7]</sup>, the

activity of TGs is highly regulated. The most important regulator is calcium. In the absence of this ion, the enzyme is not active. Additionally, TG2 is allosterically regulated by GTP/GDP<sup>[13, 14]</sup> and the redox condition of the local environment (formation of a disulfide bond).<sup>[15, 16]</sup>

In 1994, with blood coagulation factor XIII the first crystal structure of a transglutaminase was published by Yee and co-workers and gave firstly structural insights into the enzyme family.<sup>[17]</sup> Transglutaminases consist of a  $\beta$ -sandwich domain, a catalytic domain and two  $\beta$ -barrel domains (described in more detail in *Chapter 2 and Chapter 5*). The catalytic domain carries a catalytic triad comprising a cysteine, histidine and an aspartate residue. Approximately 10 years later, the crystal structure of transglutaminase 3, enzymatically cleaved between the catalytic domain and the  $\beta$ -barrel 1 domain, was published revealing three calcium binding sites.<sup>[18]</sup> However, the enzyme was crystallized in the absence of an inhibitor. Finally, in 2007 the crystal structure of transglutaminase 2 in complex with a substrate-derived inhibitor was published.<sup>[19]</sup> Bizarrely, despite of the calcium dependency of transglutaminases, the calcium binding sites were not established in contrast to transglutaminase 3.

The detailed mechanism of transglutaminases is shown in *Fig. 1.2*.<sup>[20]</sup> The catalytic cycle starts with the entrance of the glutamine-bearing substrate (acyl-donor substrate or Q-substrate) into the active site (I), followed by a nucleophilic attack of the thiolate sulfur of Cys 277 (TG2 nomenclature, FXIII: Cys 314) of the uncomplexed enzyme at the carbonyl carbon of the glutamine side chain of the Q-substrate (II). The tetrahedral intermediate (III) decomposes to the acyl-enzyme (IV) under release of  $\text{NH}_3$  whereas an imidazolium proton is abstracted from His 335 (TG2 nomenclature, FXIII: His 373). Subsequently, the lysine bearing co-substrate (acyl acceptor substrate or K-substrate) enter into the active site and the side chain amine nitrogen of the K-substrate attacks the carbonyl carbon of the thioester nucleophilically (V). Decomposition of the tetrahedral intermediate releases the iso-peptide product and the regenerated enzyme (VI). Alternatively, the thioester (IV) can be hydrolyzed. Here, water acts as a nucleophile instead of the amine whereas the imidazole nitrogen of His 335 accepts a proton of the water molecule (VII). The tetrahedral intermediate decomposes to the deamidated glutamine (glutamate) and the regenerated catalytic site.

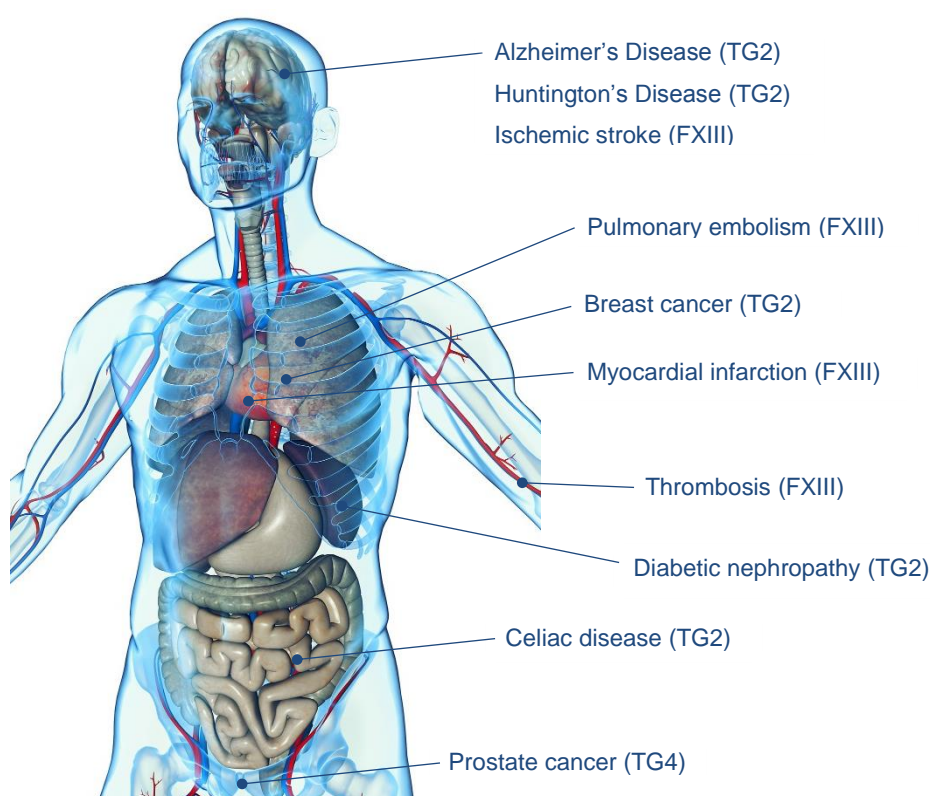




**Fig. 1.2:** Catalytic cycle of transglutaminases with involvement of Cys 277 (TG2 nomenclature, FXIII: Cys 314) and His 335 (TG2 nomenclature, FXIII: His 373) of the catalytic triad.<sup>[20]</sup> Image taken from Keillor (slightly modified).

Due to a progressive knowledge concerning physiological functions of transglutaminases and their involvement in a variety of disease processes (Fig. 1.3), this enzyme class has increasingly evolved pharmacological interest in the last decades.<sup>[21]</sup> Blood coagulation factor XIII represents a promising target in thrombolytic therapy and for preventive treatment of thromboembolic events.<sup>[22-24]</sup> TG2 is traded as a promising target for treatment of auto-immune

diseases like celiac disease<sup>[7, 24, 25]</sup>, neurodegenerative diseases like Alzheimer's and Huntington's Diseases<sup>[26-28]</sup> and fibrotic disorders such as diabetic nephropathy<sup>[29]</sup>. Additionally, TG2 is involved in inflammation and in the regulation of cell growth and apoptosis.<sup>[30]</sup> Thus, TG2 has also become of interest as a pharmacological target for anti-inflammatory drugs<sup>[21, 31]</sup> and in cancer therapy.<sup>[32]</sup> Aside from FXIII and TG2, other human transglutaminases are also involved in several disease processes.<sup>[33-38]</sup> For example, TG4 plays a crucial role in the pathogenesis of prostate cancer.<sup>[39]</sup>

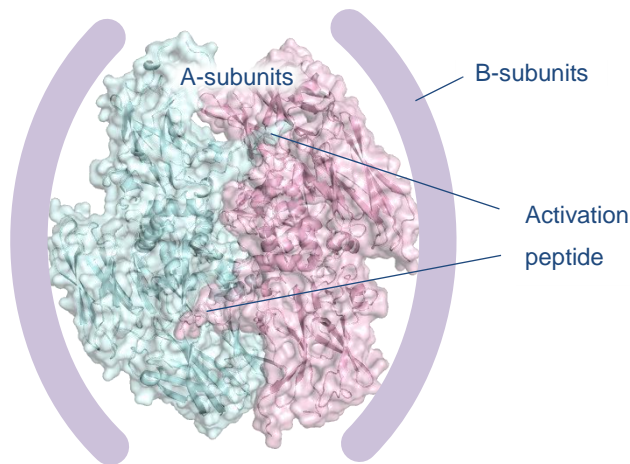


**Fig. 1.3:** Most important human diseases with involvement of transglutaminases.<sup>[40]</sup>

### 1.3 Blood Coagulation Factor XIII

Blood coagulation factor XIII, also designated as fibrin-stabilizing factor or after its discoverer Laki-Lorand factor<sup>[6]</sup>, plays a crucial role in hemostasis and constitutes a promising target for safer anticoagulants (described in more detail in the upcoming chapters).

In the blood plasma, the enzyme exists in a tetrameric form (FXIII-A<sub>2</sub>B<sub>2</sub>) consisting of two identical catalytic active subunits (FXIII-A) and two identical regulatory subunits (FXIII-B).<sup>[41]</sup> The two A-subunits form a homodimer and the two B-subunits are attached to the latter (Fig. 1.4). Each A-subunit is furnished with an activation peptide on the N-terminus that extends into the other monomer and reinforce the interaction between both monomers.



**Fig. 1.4:** Plasmatic form of blood coagulation factor XIII. In the blood plasma FXIII adopts a heterotetrameric structure (A<sub>2</sub>B<sub>2</sub>).

The activation peptide is unique for FXIII within the transglutaminase family and is most likely an important regulatory element required to involve a transglutaminase into the coagulation cascade. Plasmatic FXIII (pFXIII; FXIII-A<sub>2</sub>B<sub>2</sub>) is activated by thrombin that cleaves the activation peptide weakening the interaction between both A-subunits.<sup>[42, 43]</sup> In the presence of calcium the B-subunit dissociates from the A-subunit and the enzyme becomes active.<sup>[44, 45]</sup> Moreover, the presence of polymerized fibrin accelerates the activation by two orders of magnitude.<sup>[45-47]</sup> Interestingly, pFXIII can be also activated by a high calcium concentration without prior proteolytic cleavage.

FXIII is also present in the cytoplasm of certain cells, especially in platelets and monocytes/macrophages.<sup>[48, 49]</sup> In the cellular form (cFXIII), the enzyme exists as a homodimer (FXIII-A<sub>2</sub>)<sup>[17, 41]</sup> and is activated by an elevation of the intracellular calcium level without prior cleavage of the activation peptide.<sup>[50, 51]</sup>

The main function of FXIII concerns the enhancement of the mechanical stiffness of fibrin.<sup>[52]</sup> Here, FXIII cross-links fibrin fibers by forming isopeptide bonds between glutamine and lysine side chains. In addition to fibrin-cross-linking, FXIII attaches  $\alpha$ 2-antiplasmin to fibrin increasing the resistance towards fibrin degradation (fibrinolysis) by plasmin. Notably, aside from fibrin and  $\alpha$ 2-antiplasmin, 145 further substrates of FXIII in the blood plasma have been identified whereas 48 substrates are incorporated into the blood clot.<sup>[53]</sup> FXIII is also involved in additional biological processes like pregnancy<sup>[54]</sup>, wound healing<sup>[55]</sup> and angiogenesis<sup>[56]</sup>.

## 1.4 Structure-Based Drug Design

The average human life time has been increased drastically in the last century by medical science. The development of therapeutic substances is an elaborate process and requires the entire knowledge of natural science and engineering.

The common procedure to discover first leads for a drug development project in industry represents high-throughput screening (HTS). Here, the target protein is exposed to hundreds of thousands of molecules in a test assay system to detect functional modulation of the target protein. Substances with sufficient affinity serve as starting point for a subsequent drug design process where the lead molecule is improved in terms of affinity concerning the target protein by chemical modification. This procedure is costly in terms of time and financial resources because there is no information available about how the molecule has to be exactly modified. This issue was overcome with the development of X-ray crystallography for large protein systems as a routine tool in the 1990s providing information about the target protein on a molecular level. Ligands can be crystallized in complex with their target protein and knowledge of the chemical environment of the active site can be used for specific chemical modifications. Additionally, computer programs can be used to estimate how modified ligands can bind to the protein (docking).

The common approach in industry to develop new drugs is to combine HTS for lead discovery and X-ray crystallography to optimize the lead molecule based on the crystal structure. However, a screening with hundreds of thousands of molecules can be costly. Thus, an *in silico*

search for potential lead structures based on the crystal structure of the protein (virtual ligand screening) is discussed as a powerful alternative to HTS.<sup>[57, 58]</sup>

Nevertheless, a clear disadvantage of HTS is to start with large molecules that can only be optimized by a considerable synthetic effort, often by simultaneously reducing them in their original size. This issue can be overcome by using a fragment-based approach where a search is conducted based on small molecules (fragments) binding to different sub-pockets of the active site. Subsequently, potential fragments can be chemically connected obtaining high-affinity inhibitors of the corresponding target. The fragment-based development of drugs has already turned out to be a powerful alternative to HTS.<sup>[59]</sup> Additionally, the computer-aided approach enables the adjustment of the compound library on the respective target. This flexibility is a clear advantage over HTS because the success of a screening campaign depends predominantly on the quality of the preselected substance library.

## **2 Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants**

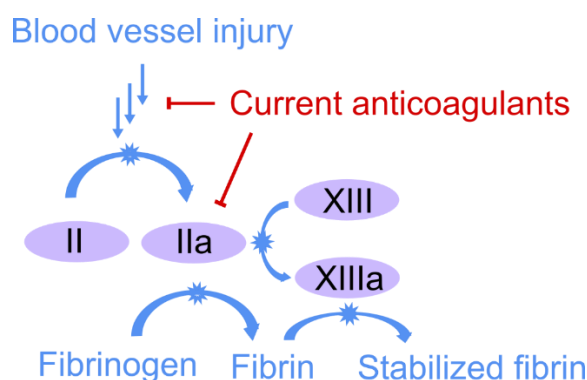
### **2.1 Introductory Remarks**

The following chapter represents a slightly modified and supplemented version of an article published in *Angewandte Chemie*.<sup>[60]</sup> The inhibited protein for crystallization was provided by Zedira GmbH. The crystallization was performed by the author of this thesis and Zedira GmbH. The minimization of a co-substrate mimetic was done by Prof. Dr. Peter Kolb. The structure determination, functional analysis and draft of the paper was performed by the author of this thesis.

### **2.2 Introduction**

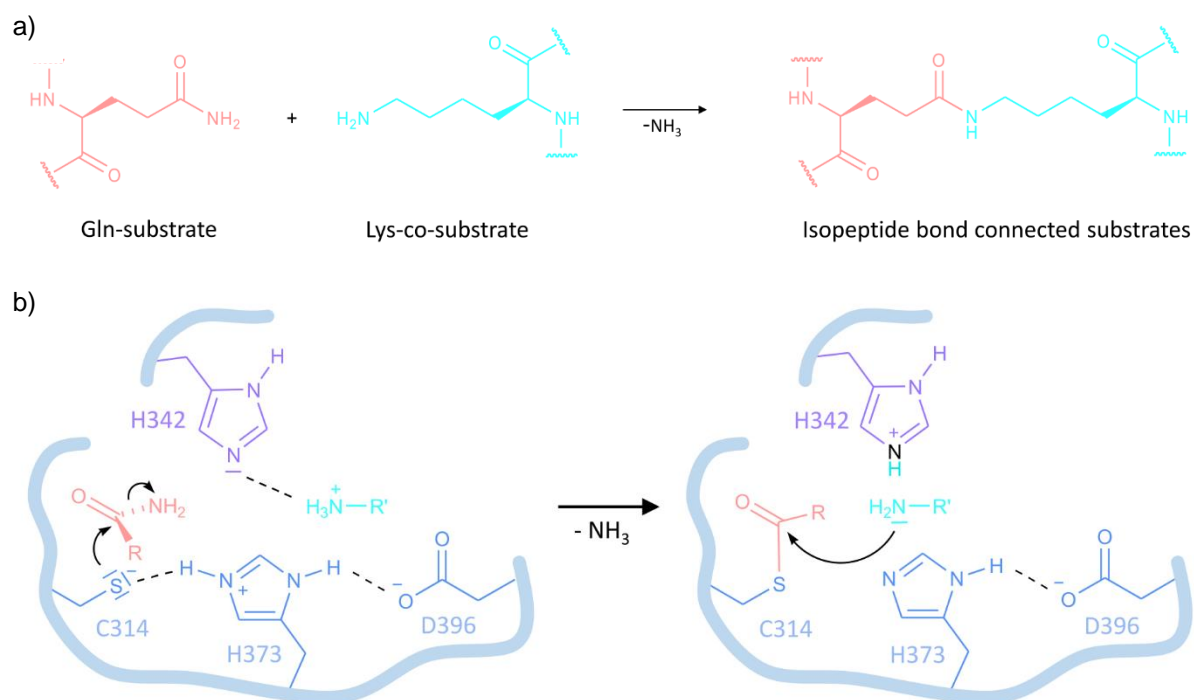
Remarkable effort has been devoted to develop drugs that target coagulation factors or platelet activation to prevent and treat thrombosis, pulmonary embolism, and acute coronary syndromes, or for reducing the risk of stroke in patients with atrial fibrillation. For decades, vitamin K antagonists were the only available anticoagulants that could be administered orally. Only recently, a new generation of direct-acting oral anticoagulants that block thrombin directly or indirectly via upstream factor Xa has become available.<sup>[61]</sup> However, depending on its activation state, thrombin can either promote or prevent blood clotting. Interference with thrombin activity by the currently available anticoagulants is characterized by an enhanced bleeding risk, thus excluding many patients from beneficial treatment.<sup>[62]</sup>

The currently targeted enzymes of the blood-clotting cascade, namely factor Xa and thrombin, belong to the family of serine proteases. Remarkably, the final enzyme in the blood clotting cascade, coagulation factor XIII, shows a markedly different mode-of-action (*Fig. 2.1*).



**Fig. 2.1:** In case of a blood vessel injury it comes to a cascade of zymogen activations that finally leads to the formation of thrombin (FIIa). Thrombin is the central key enzyme of the clotting cascade activating fibrinogen to fibrin that forms together with platelets and other blood cells a so called “soft” thrombus. Concomitantly, fibrin is stabilized by FXIIIa achieved by cross-linking fibrin chains via iso-peptide bonds and furthermore decorating fibrin with  $\alpha$ 2-antiplasmin that prevents fibrinolysis catalyzed by plasmin. All current anticoagulants directly or indirectly inhibit thrombin leading to reduced blood coagulation. In any case, inhibition of FXIIIa will not interfere with the coagulation cascade nor affect fibrin formation providing a unique mode of action.

FXIII is a member of the transglutaminase family, and catalyzes the covalent cross-linking of protein chains bearing susceptible glutamine and lysine residues through iso-peptide bonds (Fig. 2.2).<sup>[1, 41, 63, 64]</sup> In a blood clotting event, FXIII recognizes fibrin as the substrate and triggers clot maturation and accretion. In the case of congenital FXIII deficiency, delayed bleeding has been reported as a clinical manifestation. Taking these considerations into account, FXIII is regarded as a prospective target to achieve potentially safer and more efficient thrombolysis at a lower dosage of clot-dissolving agents.<sup>[23]</sup> In fact, a FXIII inhibitor may even prevent thrombus formation altogether.<sup>[22]</sup>



**Fig. 2.2:** General reaction and mechanism catalyzed by transglutaminases. The carboxamide group of a glutamine containing substrate (orange) is crosslinked with the amino group of a lysine containing substrate (cyan) by formation of a iso-peptide bond (a). The carboxamide side chain of a peptide- or protein-bound glutamine (orange) is activated leading to the acyl-enzyme intermediate while ammonia is released. Subsequently, the thioester reacts with the  $\epsilon$  amino group of lysine in the co-substrate (cyan) yielding the stable and characteristic iso-peptide bond (b). The catalytic site exhibits a catalytic triad composed of Cys 314, His 373 and Asp 396 (light blue) and a catalytic His 342 (violet). Once the Gln-substrate (orange) is accommodated, Cys 314 attacks nucleophilically its carboxamide group, His 373 binds the proton of the sulfhydryl group and is electrostatically stabilized by the adjacent Asp 396. The formed tetrahedral intermediate interacts with the oxyanion hole formed by the indole NH of Trp 279 and backbone NH of Cys 314 (not shown). Subsequently, the Lys-co-substrate (cyan) enters the catalytic center and attacks nucleophilically via its amino group the acyl-enzyme complex. Its nucleophilicity is enhanced by the neighboring His 342.

FXIII acts downstream of thrombin, effectively determining the mechanical stability, half-life, and lysis rate of clots. Although discussed as a promising target to interfere with coagulation, no suitable drug candidates are available to explore the pharmacological potential of FXIII inhibition. This situation might also result from the fact that a protein structure representing a

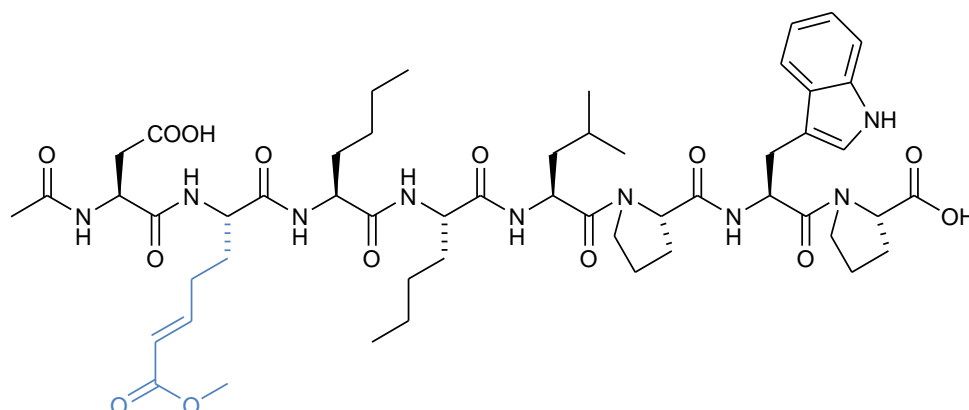


relevant active state of FXIII was previously unknown, but is necessary to embark on structure-based drug design. Only crystal structures of inactive, homodimeric FXIII have been reported so far, where the active site is completely buried and any access to the catalytic center is obscured. Such structures are unsuitable for drug design. In the structure described herein, three calcium ions recruit polar functional groups of the protein and establish local metal ion coordination sites, which induce the rearrangement of two domains along with local adaptations of the catalytic domain to expose the enzyme in an active state. The observed transformations establish the substrate and co-substrate binding sites for the formation of an iso-peptide bond and suggest involvement of a catalytic triad and a newly identified dyad in the enzyme mechanism which is considered valid for the entire transglutaminase family.

Until now, very few FXIII inhibitors have been described. Finney et al.<sup>[65]</sup> reported that the 66 amino acid peptide tridegin from the salivary gland of the giant Amazon leech *Haementeria ghilianii* is a potent inhibitor. In the late 1980s, a series of small molecules that irreversibly inhibit FXIII were explored in animal models of thrombosis. As a result of the lack of selectivity along with the short plasma half-life of only a few minutes, these inhibitors were solely considered as pharmacological tools to elucidate the physiological role of FXIII and not as prospective drug candidates.<sup>[66]</sup> For preclinical development of direct FXIII blockers it is important to have potent and selective compounds at hand to study the benefit-to-risk ratio relative to existing anticoagulants.

Factor XIII is composed, as are all human transglutaminases, of a  $\beta$ -sandwich domain, a catalytic domain, and two  $\beta$ -barrel domains. In several features, however, FXIII differs remarkably from all other members of this enzyme family. The enzymatically inactive pro-transglutaminase carries a unique 37-residue-long activation peptide which is excised by thrombin in the final phase of a coagulation event. Moreover, FXIII is the only dimeric transglutaminase composed of two zymogen A subunits (cellular FXIII-A<sub>2</sub> found, for example, in platelets and megakaryocytes) or complexed in plasma with two B-carrier subunits, thereby forming a heterotetramer (plasma FXIII-A<sub>2</sub>B<sub>2</sub>).

Recombinant human cellular factor XIII (FXIII-A<sub>2</sub>) was activated using a high concentration of calcium ions rather than by applying proteolytic activation with thrombin.<sup>[67]</sup> Activated FXIIIa<sup>o</sup> was subsequently inhibited by the irreversible peptide inhibitor ZED1301, which possesses a Michael acceptor warhead that mimics the substrate's glutamine residue (*Fig. 2.3*).

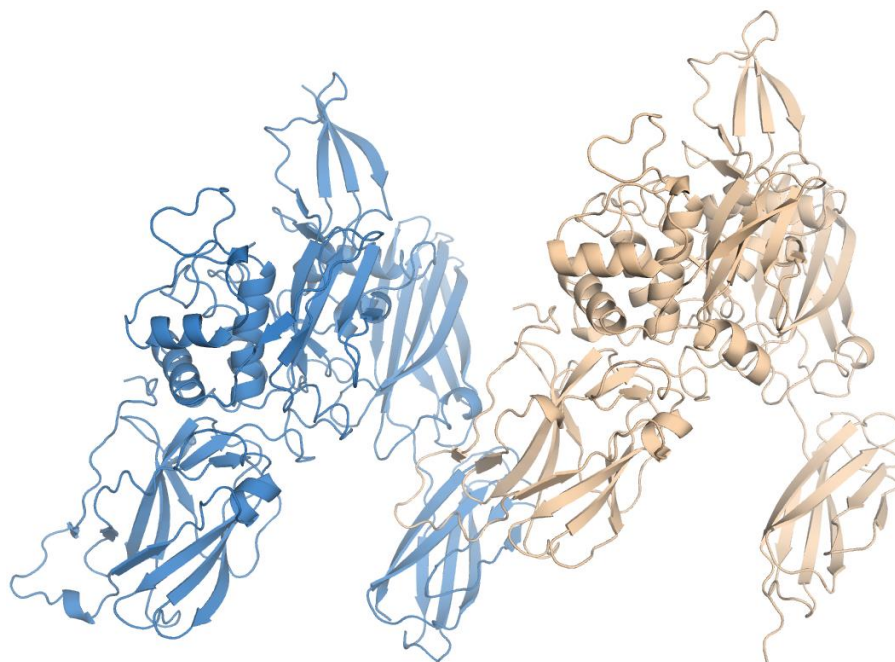


**Fig. 2.3:** Inhibitor ZED1301 used for inhibition and subsequent crystallization of FXIIIa°. The inhibitor was derived based on a phage display screening<sup>[68]</sup> and bears instead of the substrate glutamine an  $\alpha,\beta$ -unsaturated methyl ester (blue) for covalent linkage to Cys 314.

## 2.3 Results & Discussion

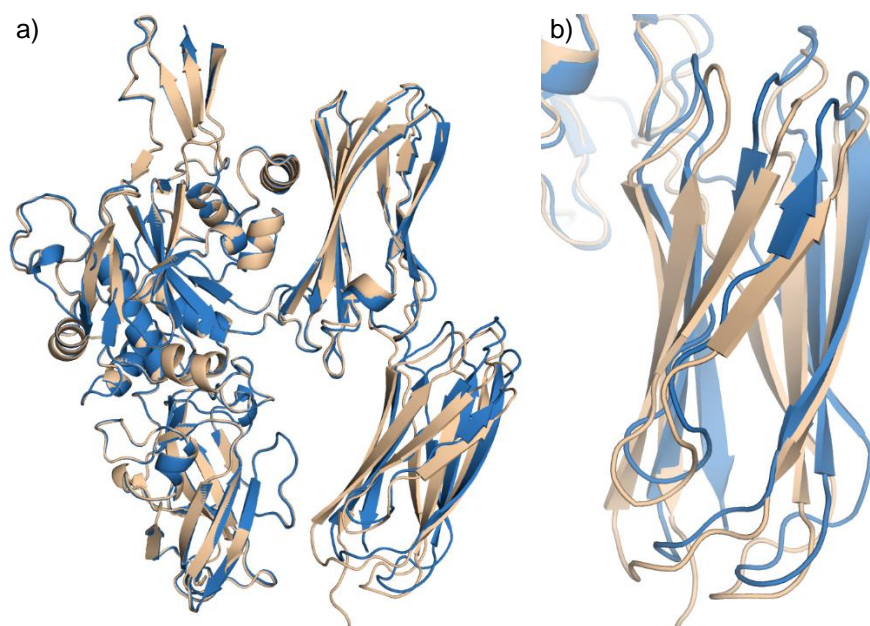
### 2.3.1 General Structural Features and Overall Structural Rearrangements

Here we report the first high-resolution crystal structure (1.98 Å) of FXIII in an active state (termed FXIIIa°, as suggested by Muszbek et al.<sup>[69]</sup> for Ca<sup>2+</sup> activation without proteolysis) in complex with an irreversibly bound inhibitor. The complex crystallized in the space group P1 with two independent FXIIIa° molecules per asymmetric unit (*Fig. 2.4*).



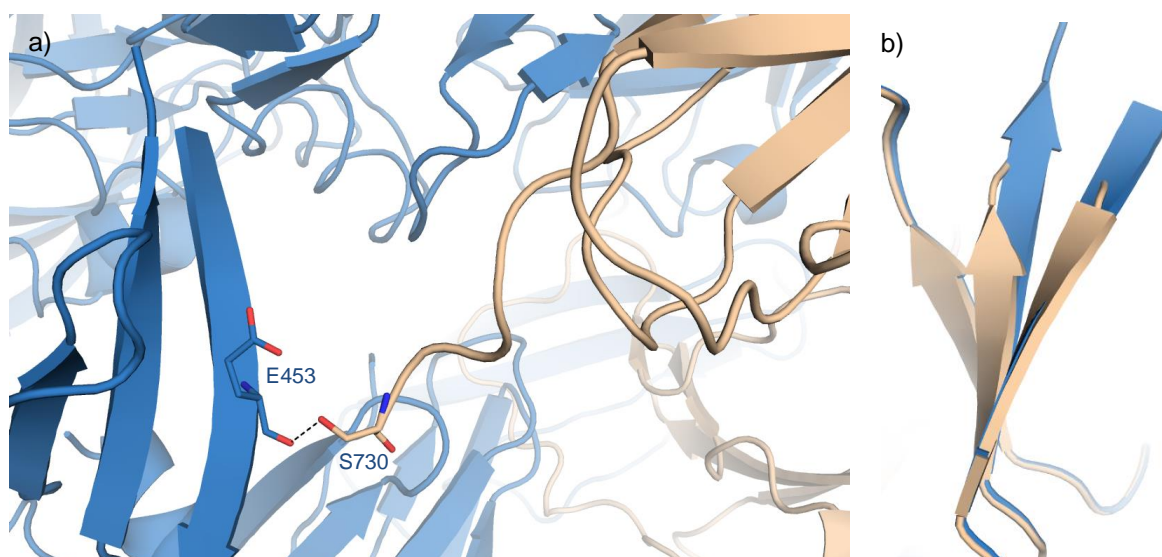
**Fig. 2.4:** Molecule 1 (blue) and molecule 2 (beige) in the asymmetric unit of FXIIIa° (PDB ID: 4KTY).

A structural superposition of both molecules of the asymmetric unit shows that there are only minor structural differences (*Fig. 2.5a*). The rmsd of both molecules of the asymmetric unit amounts 0.816 Å. The most significant one concerns the  $\beta$ -barrel 2 domain whereby both domains are translocated to each other by around 4 Å and show an rmsd of 0.877 Å (*Fig. 2.5b*).



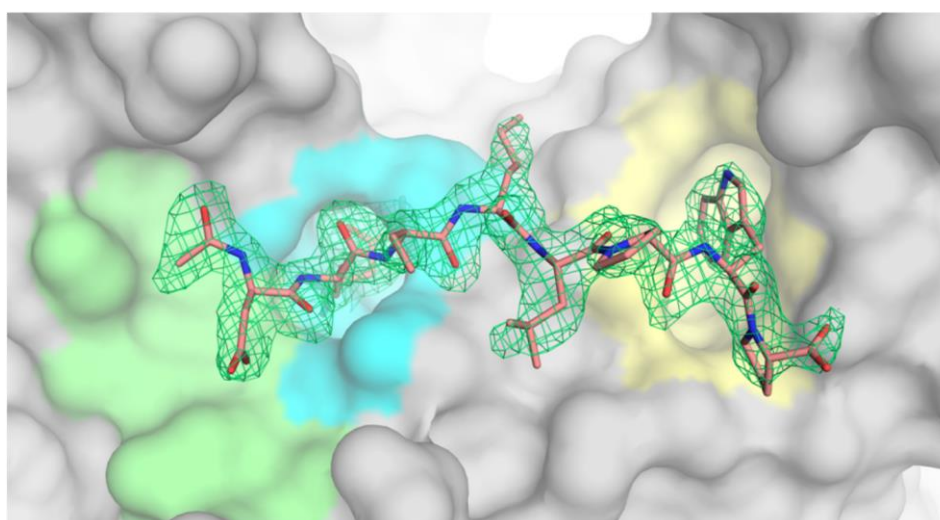
**Fig. 2.5:** a) Structural superposition of both molecules of the asymmetric unit of FXIIIa<sup>o</sup> (PDB ID: 4KTY) on the  $\beta$ -sandwich and catalytic domain. b) Magnified representation of the  $\beta$ -barrel 2 domains of both molecules of the asymmetric unit.

Regarding the crystal packing, an interaction between the three-stranded  $\beta$ -sheet of molecule 1 and the C-terminus of molecule 2 occurs. In detail, the backbone oxygen of Glu 543 accepts a hydrogen bond of the hydroxyl group of Ser 730. However, comparing both three-stranded  $\beta$ -sheet of molecule 1 and 2, the  $\beta$ -sheet of molecule 1 adopts the same conformation as the  $\beta$ -sheet of molecule 2 that exhibits no interactions.



**Fig. 2.6:** a) Hydrogen bond between Glu 453 of the three-stranded  $\beta$ -sheet and Ser 730 of the C-terminus between two symmetry-related molecules. b) Interaction has no effect on the conformation of the three-stranded  $\beta$ -sheet.

The inhibitor ZED1301 binds at the surface of the catalytic domain of FXIIIa<sup>o</sup>. This region can be split into three main interaction sites, the catalytic site, the adjacent proximal “ $\alpha$ -space”, and the more distant hydrophobic pocket (Fig. 2.7).

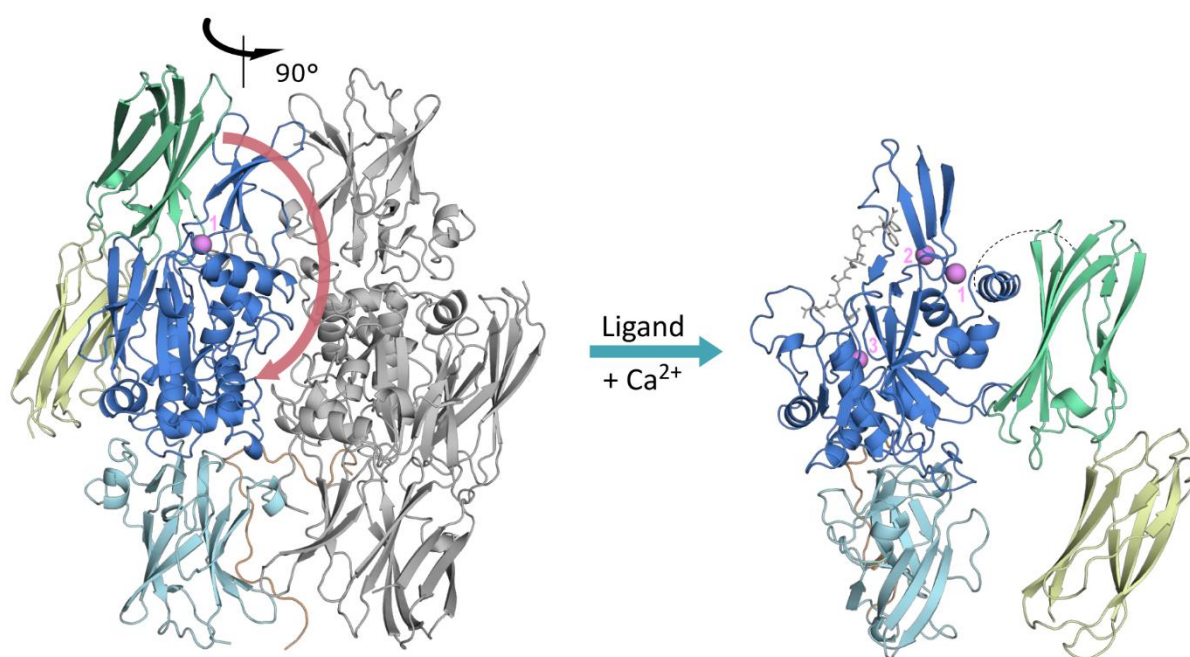


**Fig. 2.7:** The ligand ZED1301 (orange,  $F_o-F_c$  difference electron density contoured at  $3\sigma$ ) binds covalently to Cys 314 in the catalytic site. Next to the latter site, the interaction region, called  $\alpha$ -space,



is formed (green). At its far end, the ligand exposes an indole moiety to the hydrophobic pocket (yellow). The surface of FXIIIa° is depicted in gray.

In the previously determined inactive state, the catalytic site is covered by  $\beta$ -barrel 1, where Tyr 560OH forms a hydrogen bond to S<sub>γ</sub> of Cys 314. Here we show that binding of ZED1301 in the presence of calcium ions induces a large conformational transition, whereby the two  $\beta$ -barrel domains swing aside to expose the catalytic center (Fig. 2.8).



**Fig. 2.8:** In the inactive state FXIII (PDB ID: 1GGU) exists as a dimer (one monomer colored by domain, second in gray). Upon substrate binding, FXIII dissociates and the positions of the  $\beta$ -sandwich (turquoise) and catalytic domain (blue) remain unchanged, whereas  $\beta$ -barrel 1 (green) and  $\beta$ -barrel 2 (yellow) undergo a remarkably large movement (red arrow). The activation peptide (orange) stabilizes the dimer contact in the inactive state and is still bound to the protein and partly exposed to solvent. The active site is located in the upper left part of the catalytic domain, where the inhibitor ZED1301 (gray) is covalently attached to the sulfur atom of Cys 314. The catalytic domain of the active form (PDB ID: 4KTY) contains three calcium ions (pink) instead of only one calcium ion in the case of the inactive form. The loop connecting the  $\beta$ -sandwich domain and  $\beta$ -barrel 1 (dashed line) is not defined by electron density.

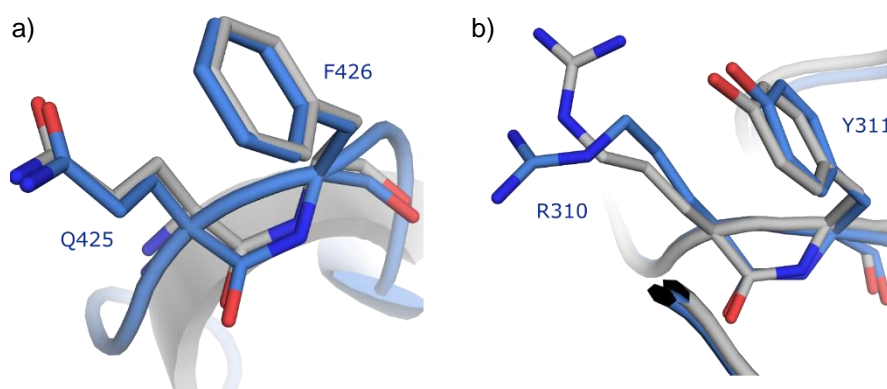
The complex crystallizes as a monomer, in agreement with our biochemical data determined from size-exclusion chromatography (*Tab. 2.1*). The apparent molecular weight of the FXIII-ZED1301 complex was determined using size exclusion chromatography. Beside standard globular calibration proteins FXIII-A<sub>2</sub> (dimeric, calculated 166 kDa) and TG2-ZED1301 complex (monomeric, calculated 79 kDa) were analyzed. The data revealed an apparent molecular weight of FXIIIa°-ZED1301 complex of 47.7 kDa matching to the monomeric tissue transglutaminase inhibitor complex (TG2-ZED1301; 41.1 kDa) while displaying about half the apparent molecular weight of dimeric FXIII-A<sub>2</sub> (93.5 kDa).

**Tab. 2.1:** Size exclusion chromatography (performed by Zedira GmbH).

Protein/ Complex	V <sub>e</sub> [ml]	V <sub>e</sub> /V <sub>0</sub>	Apparent MW [kDa]
FXIII-A <sub>2</sub>	8.77	1.589	93.5
FXIII-ZED1301	9.26	1.678	47.7
TG2-ZED1301	9.37	1.697	41.1

We propose that the adopted conformation represents an active FXIII structure, relevant for drug design, as functionally important and previously hidden sites are exposed. These findings are in agreement with mass spectrometry data obtained for hydrogen/deuterium exchange in solution.<sup>[70, 71]</sup> Structural rearrangements upon activation occur only within the catalytic domain, while the  $\beta$ -sandwich and the two  $\beta$ -barrel domains conserve their overall fold. The electron density of the loop connecting the catalytic and  $\beta$ -barrel 1 domain (residues 502-515) is not visible, possibly because of the pronounced flexibility of this portion. This observation agrees with biochemical evidence<sup>[72]</sup> that the Lys 513-Ser 514 peptide bond is easily accessible in this flexible loop and can be cleaved by thrombin. Furthermore, in our structure, the uncleaved pro-peptide remains at the N terminus of the  $\beta$ -sandwich domain. We cannot fully exclude an effect of the flexible and distant pro-peptide on the conformation of the catalytic subunit; however, we presume its relevance for drug design should be negligible. Nonetheless, most of the contacts formed in the inactive state are lost due to rearrangements upon activation and dissociation of the dimer in the monomeric state. Twenty-three amino acids of the N-terminal propeptide chain are visible in the electron density.

The active state of factor XIII contains two cis peptide bonds that are also present in the inactive form (*Fig. 2.9*). Weiss et al. suggested that a cis-trans isomerisation would parallel the transition from inactive to active form and would produce part of the energy required for such a tremendous conformational change.<sup>[73]</sup> However, their assumption is not confirmed by our structure.

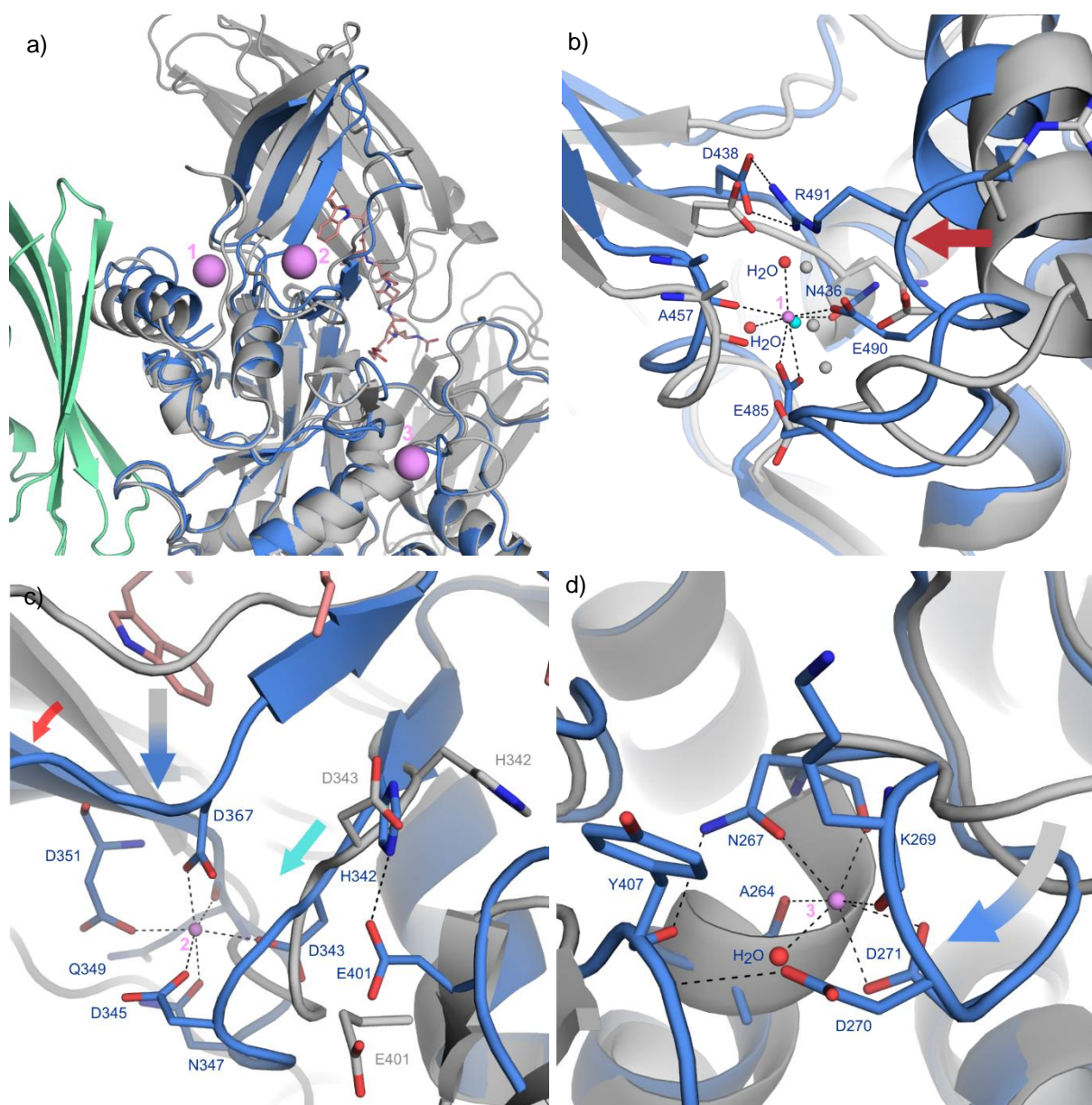


**Fig. 2.9:** Both inactive (grey) and active form (blue) of FXIII contains cis-peptide bonds. One is formed by the amino acids Gln 425 and Phe 426 (a). The other one is formed by Arg 310 and Tyr 311 (b).

### 2.3.2 Structural Rearrangements Triggered by Calcium Binding

There has been much speculation concerning the driving forces that induce rearrangement upon activation. Since the discovery of FXIII by Laki and Lorand in the 1940s, the importance of calcium ions for FXIII activation, as for all transglutaminases, has now become well established.<sup>[6]</sup> Thus, it was surprising that crystals grown with inactive FXIII at high calcium concentrations, or crystals of inactive FXIII exposed to solutions containing large amounts of calcium ions, indicated only a single calcium-binding site per monomer (*Fig. 2.10b*).<sup>[17, 74]</sup> In our FXIIIa<sup>o</sup> structure, we observe three calcium ion sites (*Fig. 2.10a*) that are essential for driving the enzyme into an active state. Presumably, this form can only be captured when a peptide inhibitor is bound simultaneously with the three calcium ions and stabilizes the protein in this state.



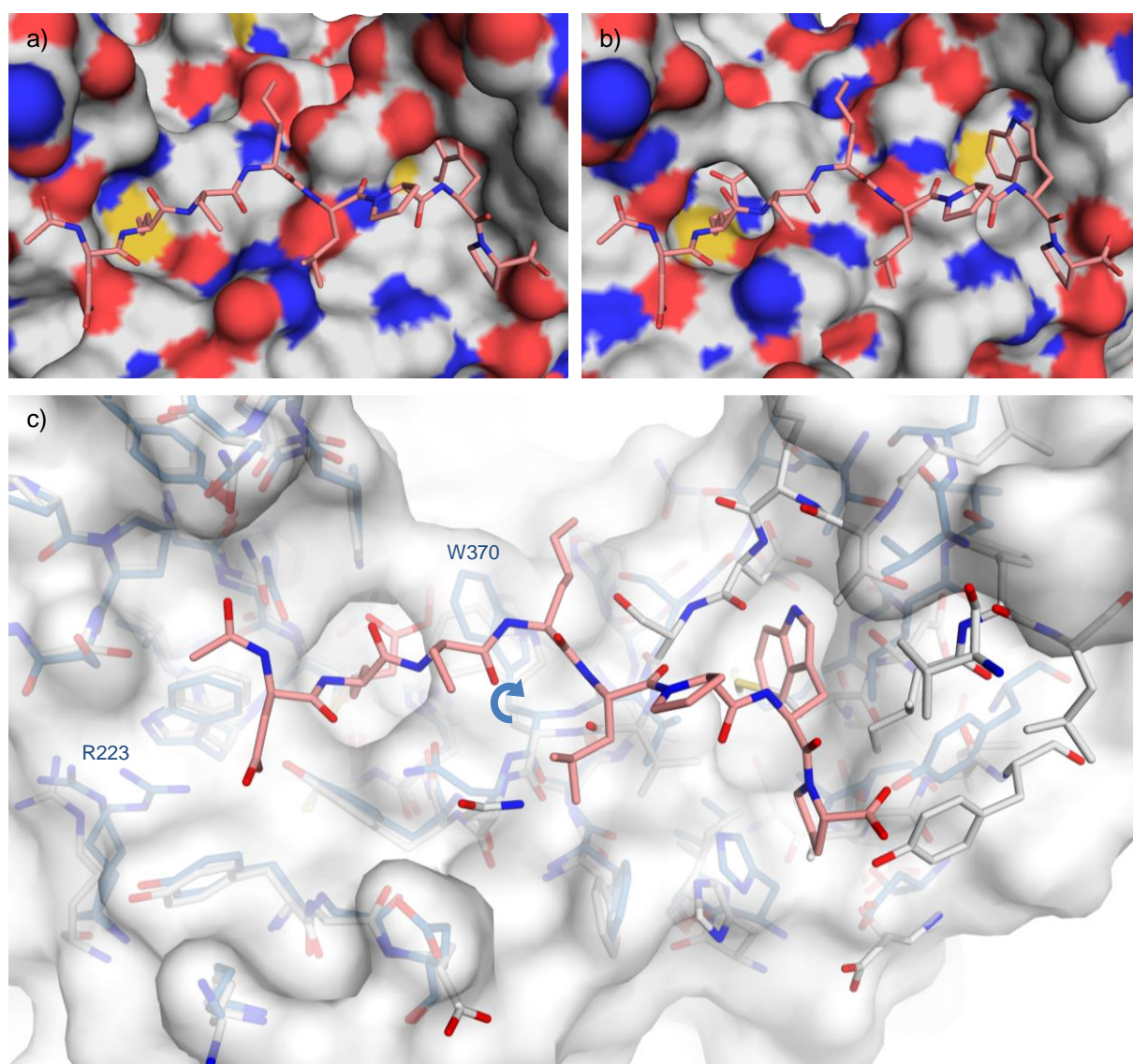


**Fig. 2.10:** a) In the active state, three calcium ions (pink) bind to the catalytic domain (blue). They induce structural changes from the inactive to the active form (gray, site 1: 1GGU; sites 2 and 3: 1F13). Dashed lines indicate coordinative, hydrogen bonds or salt bridges. ZED1301 is colored in orange. b) At site 1,  $\text{Ca}^{2+}$  binding results in a shift of an  $\alpha$ -helix in the active state of FXIII enabling a strong interaction between Arg 491 and Asp 438. The inactive form also binds calcium (turquoise sphere), but only with one direct interaction to Ala 457CO. The remaining coordination sphere is completed by water molecules (gray spheres). c) At site 2, the calcium ion binding drags the loop containing Asp 367 downwards and supports rotational movement of the three-stranded  $\beta$ -sheet by calcium coordination of Asp 351 (red arrow). Both interactions lead to the formation of the hydrophobic pocket next to the active site (gray-blue arrow), which is important for substrate accommodation. Site 2 is also involved in the

*formation of the catalytic dyad. Coordination of Asp 343 to  $\text{Ca}^{2+}$  drags a loop downwards (turquoise arrow) and enables the spatial proximity of His 342 and Glu 401 required for activity. d) Population of site 3 results in a loop rearrangement that contains residues which are part of the K-substrate binding region such as Lys 269 and Tyr 407.*

From a comparison of the structures of active and inactive FXIII, the following activation mechanism can be proposed. Calcium binding site 1 is already populated in the inactive state. However, only Ala 457CO coordinates directly to the calcium ion and the remaining coordination sites are filled with water molecules. Upon activation, the calcium ion recruits, besides Ala 457CO, the side chains of Asn 436, Glu 485, and Glu 490 to accomplish its coordination polyhedron. Glu 490 belongs to the  $\alpha$ -helix and is mainly responsible for a shift of this helix (*Fig. 2.10b*). This in turn facilitates Arg 491 to form a strong salt bridge with Asp 438 between the guanidinium moiety and the carboxy group, respectively. Importantly, both residues are highly conserved between the human transglutaminase family and Asp 438 belongs to the three-stranded  $\beta$ -sheet.

The development of the second calcium binding site translocates part of the catalytic domain into a geometry that enables formation of a hydrophobic pocket next to the catalytic center (*Fig. 2.11b*). This movement is initiated by the coordination of Asp 367 to  $\text{Ca}^{2+}$ , and the adjacent  $\beta$  strand is dragged towards the central ion (*Fig. 2.10c*). Importantly, the pocket thus created accommodates, and is stabilized at the same time by, the C-terminal Trp residue of our inhibitor. Furthermore, the  $\beta$  strand comprises Trp 390, whose indole moiety is flipped by almost  $90^\circ$  to form a hydrophobic tunnel next to the catalytic site together with the side chain of Trp 279 (*Fig. 2.11c*). This tunnel can subsequently host the putative K-substrate to form an iso-peptide bond. Such a tunnel has similarly been observed for TG2 in the crystal structure assumed to correspond to an active state.<sup>[17, 19, 75]</sup>

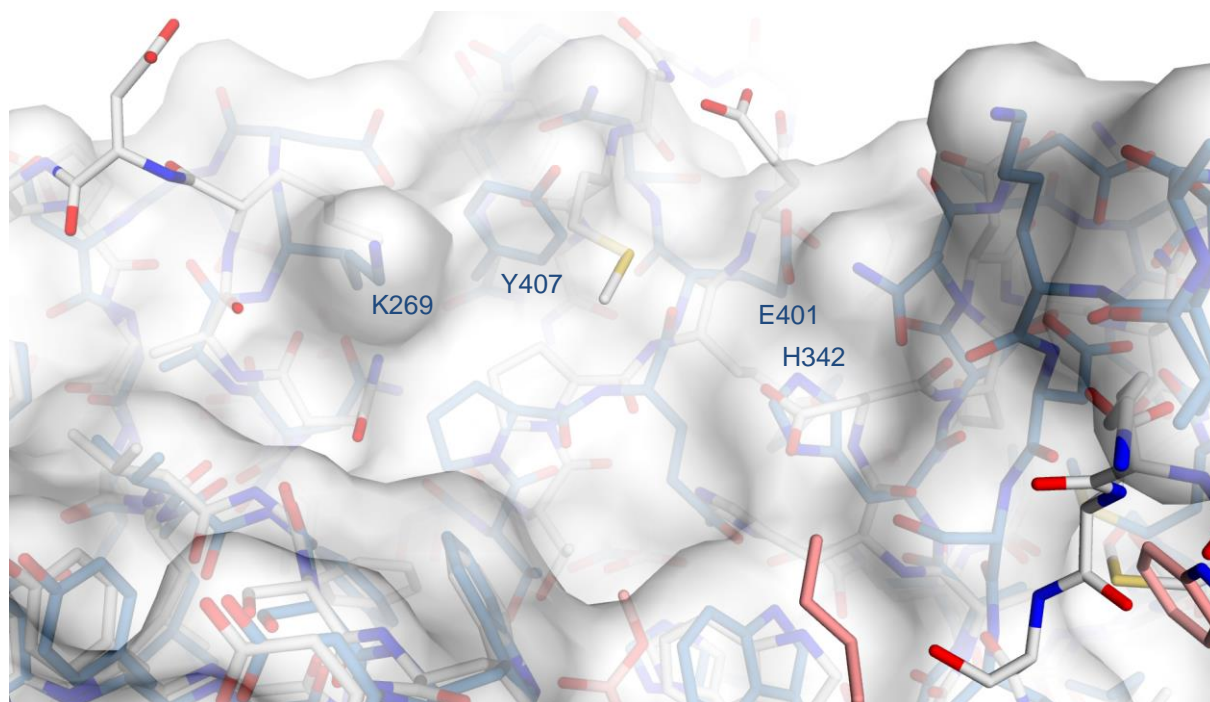


**Fig. 2.11:** Comparison of the substrate binding site of FXIII adopted in the inactive (PDB ID: 1F13) and the active state (PDB ID: 4KTY). The shape of the surface of FXIIIa° (b), which hosts the inhibitor ZED1301 (red), differs significantly from the geometry adopted in the inactive state (a), particularly with respect to the hydrophobic pocket induced upon calcium binding and occupied by the indole ring of the ZED1301. The inhibitor of the superimposed FXIIIa° structure is also shown for simplifying the orientation. c) Superposition of the inactive state (grey sticks) and the active state (blue sticks) reveals that the main reorientation of both states occurs in the area of the hydrophobic pocket. At the  $\alpha$ -space, only Arg 223 adopts a further conformation. At the catalytic site, the indole ring of Trp 370 rotates aside and thus enables the formation of the hydrophobic tunnel.

Moreover, population of the second  $\text{Ca}^{2+}$  site affects the orientation of a three-stranded antiparallel  $\beta$ -sheet in the upper part of the catalytic domain that is also involved in the formation of the above-mentioned hydrophobic pocket. The involvement of Asp 351 in  $\text{Ca}^{2+}$  coordination moves the adjacent  $\beta$  strand downwards. In addition, on the following loop, the side chain and the backbone carbonyl oxygen atom of Gln 349 together with Asn 347, Asp 345, and Asp 343 shape the second  $\text{Ca}^{2+}$  site by wrapping around the metal ion. This site has a significant impact on the structural establishment of a catalytic dyad formed by His 342 and Glu 401 needed for activation of the K-substrate (*Fig. 2.10c*, *Fig. 2.12*). In the inactive conformation, the  $\text{N}_\epsilon$  imidazole nitrogen atom of His 342 and the carboxylate oxygen atom of Glu 401 are too distant for any productive interaction (about 10 Å, *Fig. 2.10c*). The coordination of Asp 343 to  $\text{Ca}^{2+}$  drags the loop towards the calcium ion, thereby resulting in a shift of His 342 into a position now capable of forming a hydrogen bond with Glu 401, which is located on a flexible loop and can therefore easily be shifted towards His 342.

The third calcium binding site further affects the binding of the lysine-containing co-substrate that attacks the thioester intermediate from the opposite side of the above-mentioned hydrophobic tunnel. The interaction of Asp 271 with  $\text{Ca}^{2+}$  results in a reorientation of a loop near the protein surface (*Fig. 2.10d*). Moreover, Asn 267 and Asp 270 are part of this loop and interact with Tyr 407. The incorporated calcium ions act as a template and induce, through the involvement of polar side chain functions in  $\text{Ca}^{2+}$  coordination, structural changes essential for the enzyme to shift from its active-site-shielded, inactive state to a functionally active state.





**Fig. 2.12:** The structural rearrangement of the co-substrate binding site during transition from the inactive state (grey, PDB ID: 1F13) to the active state (blue, PDB ID: 4KTY) is even more extensive compared to the substrate binding site (as shown in Fig. 2.11). The surface is depicted of FXIIIa°. Both calcium binding site 2 and 3 affects the shape of the co-substrate binding site. The formation of the catalytic dyad (His 342, Glu 401) is induced by site 2. Calcium coordination at site 3 leads to loop rearrangements (Lys 269 and Tyr 407 are shown as surrogates for the structural change).

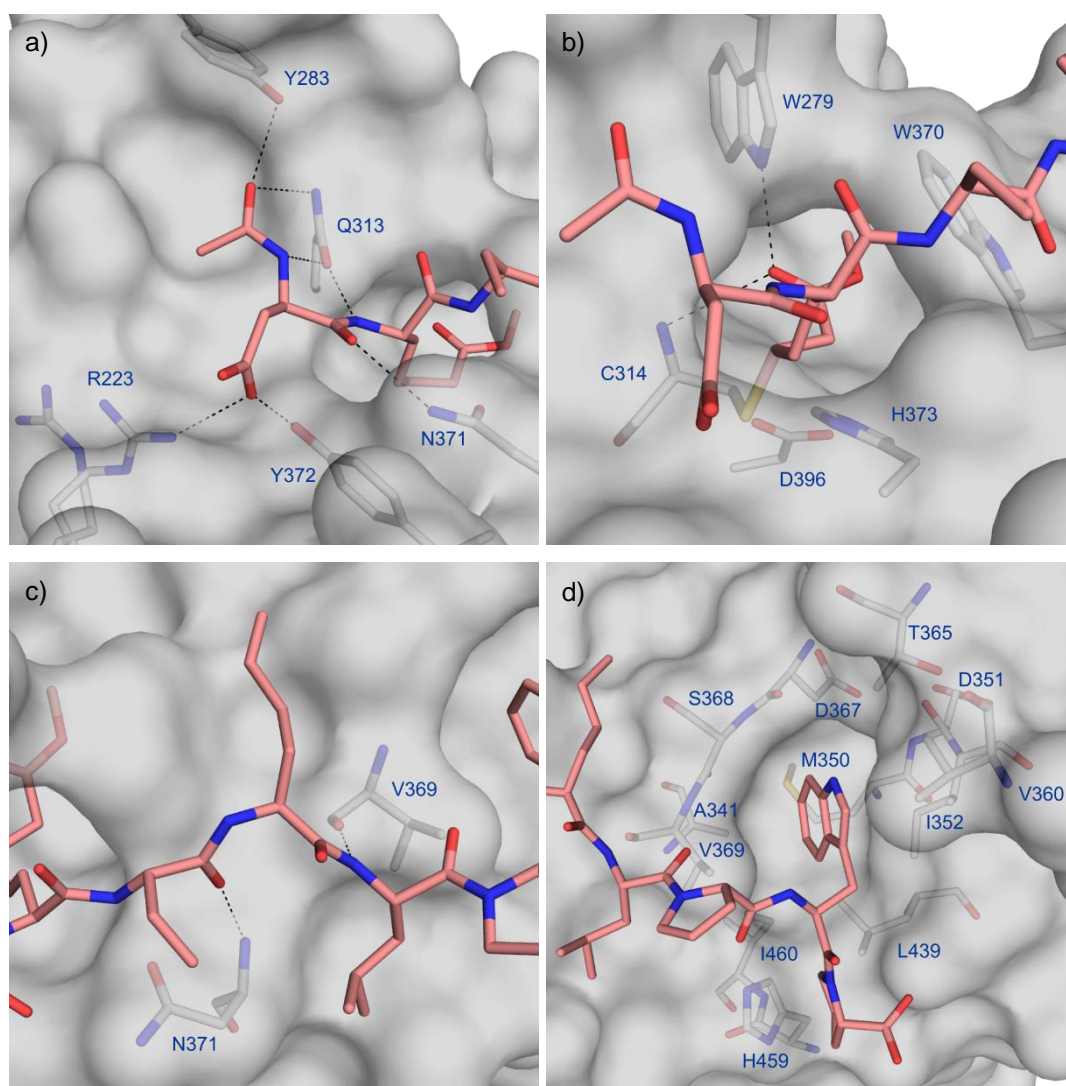
### 2.3.3 Binding Mode of ZED1301

In the  $\alpha$ -space the inhibitor interacts via seven hydrogen bonds with the FXIIIa° (Fig. 2.13a). The acetyl oxygen of the ligand forms hydrogen bonds to the oxygen of Tyr 283 and carboxamide nitrogen of Gln 313. The backbone NH of the N-terminal aspartate of the ligand interacts via a hydrogen bond with the carboxamide oxygen of Gln 313 that forms a further hydrogen bond to the adjacent backbone nitrogen of the ligand. The backbone oxygen of the ligand's N-terminal aspartate is hydrogen-bonded to the terminal nitrogen of Asn 371. Furthermore, the hydroxyl group of the N-terminal aspartate of the ligand forms H-bonds to both the hydroxyl oxygen of Tyr 372 and the guanidinium NH of Arg 223 that adopts two conformations.

Within the active site of FXIIIa<sup>o</sup>, the catalytic triad most likely adopts a geometry similar to that required for substrate binding and subsequent transformation into the thioester form. The Michael acceptor  $\beta$ -carbon atom of ZED1301 forms a covalent bond with Cys 314 (*Fig. 2.13b*). Moreover, the carbonyl oxygen atom of the terminal methyl ester establishes hydrogen bonds to the oxyanion hole formed by the indole nitrogen atom of Trp 279 and the backbone NH group of Cys 314. The adjacent hydrophobic tunnel is formed by Trp 279 and Trp 370 and prevents hydrolysis of the intermediately formed thioester.

The central bridging part of the ligand interacts via two H-bonds with the protein (*Fig. 2.13c*). The backbone nitrogen of the amino acid, that comprises the war head, forms a hydrogen bond to carboxamide oxygen of Asn 371. Asn 371 interacts with the backbone oxygen of the adjacent N-terminal norleucine of the ligand. The second H-bond is formed between the ligand's backbone NH of valine and the backbone carbonyl oxygen of Val 369.

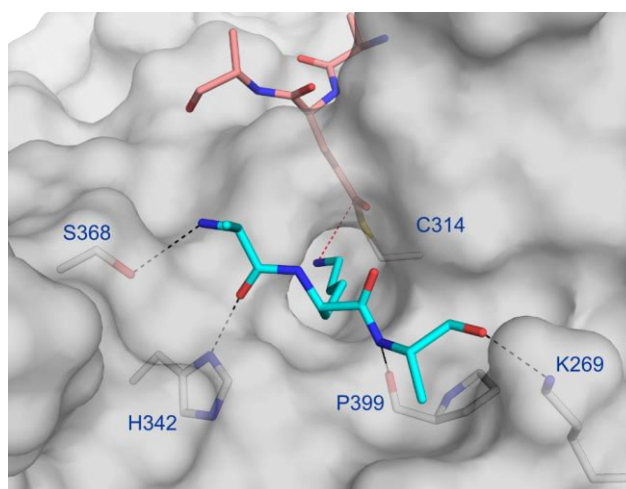
The indole ring of the ligand's tryptophan addresses the hydrophobic pocket of FXIIIa<sup>o</sup> that consists predominantly of apolar amino acids residues (*Fig. 2.13d*).



**Fig. 2.13:** a) Multiple hydrogen bonds formed between the ligand and FXIIIa<sup>o</sup> in the  $\alpha$ -space. b) The ligand forms a covalent bond to Cys 314 in the catalytic site and interacts through its Michael acceptor carbonyl oxygen atom with the NH groups of Trp 279 and Cys 314, which comprise the oxyanion hole. A hydrophobic tunnel is formed, flanked by Trp 279 and Trp 370, preventing hydrolysis of the intermediately formed thioester. c) Two hydrogen bonds are formed between the peptidic backbone of the inhibitor and the region between the catalytic site and the hydrophobic pocket. d) Enlargement of the hydrophobic pocket which is predominantly comprised of hydrophobic residues.

### 2.3.4 Modeling of a Co-substrate Mimetic

The tunnel adopts a shape similar to that in TG2,<sup>[19]</sup> but in the active state of FXIIIa° the terminal “exit” that serves as an entrance port for the side chain of the reacting lysine residue of the K-substrate is much better established. This helps to orient the K-substrate in the correct geometry required to react with the thioester intermediate. This assumption is corroborated by docking an artificial Ala-Lys-Ala tripeptide that mimics the orientation of the K-substrate in this site (*Fig. 2.14*). The lysine side chain fits perfectly into the hydrophobic tunnel and places the lone pair of electrons on its terminal  $\epsilon$ -amino group at an appropriate distance to the thioester carbon atom. This location is compatible with the requirements of a nucleophilic attack of Lys on the thioester-enzyme intermediate. Furthermore, potential hydrogen bonds are readily identifiable between the tripeptide and protein, firmly fixing the co-substrate in position (*Fig. 2.14*).

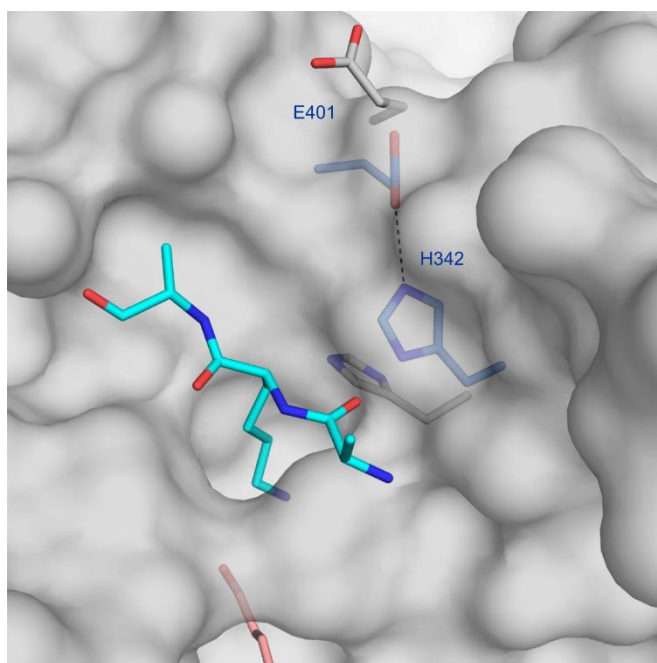


**Fig. 2.14:** The K-substrate modeled as an Ala-Lys-Ala tripeptide (cyan) fits perfectly into the back side of the hydrophobic pocket. Hydrogen bonds of the minimized tripeptide are shown as dashed lines. The original ligand was remodeled into a Q-substrate (Ala-Gln-Ala tripeptide, orange) to simulate the binding of the two substrates.



### 2.3.5 Catalytic Dyad

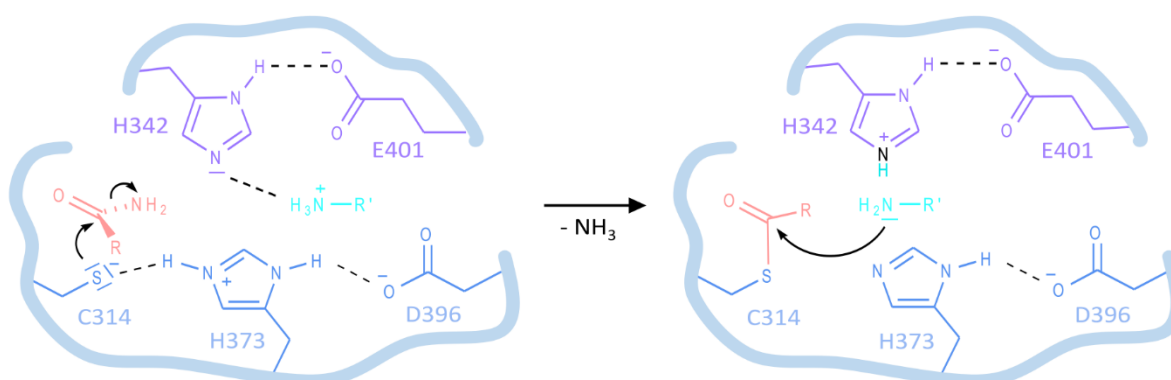
In addition to the well-accepted catalytic triad, we propose the existence of a catalytic dyad also exposed to the catalytic center to facilitate nucleophilic attack of the K-substrate in FXIIIa<sup>o</sup> (Fig. 2.15). This dyad is composed of His 342 and Glu 401 and attains a productive arrangement upon calcium activation (Fig. 2.10c). Both His 342 and Glu 401 are conserved across all human transglutaminases except TG7, which contains Gln instead of Glu. About 20 years ago, Yee et al. attributed a catalytic function to His 342 and suggested assistance by Glu 434 and not Glu 401.<sup>[17]</sup>



**Fig. 2.15:** The newly established catalytic dyad consisting of His 342 and Glu 401 is located next to the entrance of the hydrophobic tunnel at the binding site of the K-substrate and is only formed in the active state (blue sticks).

Mutation of His 342 to Ala resulted in a strikingly reduced activity with respect to the formation of the iso-peptide bond, but only a minor change in the affinity of the K-substrate towards the enzyme.<sup>[76]</sup> In the recently determined TG2 structure, assumed to represent the active state, the corresponding His 305 residue is oriented towards the glutamine-containing substrate-binding region.<sup>[19]</sup> Thus, one could argue that the latter structure does not yet correspond to the fully

active state of TG2. Admittedly, our structure also represents only a different frozen situation in the solid state, reflecting a geometry the protein is, in principle, able to adopt. Nonetheless, we propose that this structure approaches the situation in the active state more closely as it can easily accommodate both substrates along with full population of all three calcium binding sites, establishment of the hydrophobic tunnel, and formation of the additional catalytic dyad. The change in the protonation states of the substrates during formation of the iso-peptide bond appears puzzling. Under physiological conditions, the lysine residue of the K-substrate should bear a charged ammonium group. Yet, entering the protein binding site could induce a substantial  $pK_a$  shift towards lower basicity, thus resulting in an uncharged state of lysine. This effect might be supported by the adjacent His 373 of the catalytic triad, which could carry a positive charge. Upon formation of the amide bond, the catalytic dyad could operate reversibly, similar to the well-known mechanism of the catalytic triad in serine proteases, and pick up the second proton from the amino head group of the reacting lysine residue. Finally, the catalytic site of FXIII and also of other human transglutaminases can be extended by Glu 401 (*Fig. 2.16*).



**Fig. 2.16:** General mechanism for FXIII and other human transglutaminases completed by Glu 401.

## 2.4 Summary & Conclusion

Blood coagulation factor XIII represents a promising target for the development of anticoagulants with a lower bleeding risk compared to current anticoagulants. We were able to crystallize calcium activated FXIII in complex with a covalently attached inhibitor. The yielded crystal structure adopts an open conformation whereby the active site is not covered by the  $\beta$ -barrel 1 domain anymore as in the inactive state.

Aside the usage of the crystal structure for structure based development of a blood coagulation factor XIII blocker, the structure of FXIIIa<sup>o</sup> reveals the involvement of calcium in the activation process. Calcium binding affects both the shape of the substrate and co-substrate binding site. Furthermore, a catalytic dyad could be identified induced by calcium binding as well.

## 2.5 Experimental Part

### 2.5.1 Methods Protein Expression, Purification and Inhibition

The functional expression of recombinant FXIII-A<sub>2</sub> (A-subunits) in insect cells, its respective purification and the synthesis of Michael acceptor (MA) transglutaminase blockers has recently been described<sup>[77]</sup> and will be shortly summarized below. Based on the results of phage display screening for preferred substrates<sup>[68]</sup> we designed the lead ZED1301 Ac-Asp-MA-Nle-Nle-Leu-Pro-Trp-Pro-OH (Nle: norleucine) as a potent inhibitor. Activation of blood coagulation factor XIII was achieved by high calcium ion concentrations instead of exposure to thrombin.<sup>[78]</sup> A solution of purified recombinant FXIII-A<sub>2</sub> (0.2 mg/mL) in 10 mM Tris-HCl, pH 8.0, 300 mM NaCl, 15% (v/v) glycerol containing 100 mM Ca<sup>2+</sup> was mixed with the inhibitor at a molar ratio of 1:25. The mixture was stirred and incubated at room temperature for 1 h. The inhibited complex was concentrated and subsequently applied to a Superdex75pg HiLoad gel filtration column equilibrated with 10 mM Tris-HCl pH 8.0, 300 mM NaCl, 15% (v/v) glycerol containing 3 mM CaCl<sub>2</sub>.

### 2.5.2 Crystallization, Data Collection and Structure Determination

The eluted protein was concentrated to 9.7 mg/mL and protein crystals were grown at ambient temperature using the hanging drop vapor diffusion method. Crystallization was achieved by mixing 5 µL protein solutions with 5 µL precipitation solution (170 mM ammonium sulphate, 85 mM sodium cacodylate pH 6.5, 25.5% (w/v) PEG 8,000, 15% (v/v) glycerol). A suitable single crystal was obtained after removal of small satellites from its surface. The crystal was flash-frozen in liquid nitrogen and a dataset of 400 images with 0.5 degree rotation at a wavelength of 0.91841 Å was collected at the BL14.2, BESSY II electron storage ring.<sup>[79]</sup> Data were processed to a resolution of 1.98 Å in space group P1 using the program HKL2000.<sup>[80]</sup> Higher C2 symmetry could be excluded by checking for alternative settings with XPREP.<sup>[81]</sup> The structure was determined by molecular replacement with the CCP4<sup>[82]</sup> program Phaser.<sup>[83]</sup> The molecular replacement model had to be split into the four subunits. The β-sandwich domain

and the two  $\beta$ -barrel domains were used as determined for the inactive form of FXIII (PDB ID: 1F13). For the catalytic domain, a homology model based on the active form of TG2 (PDB ID: 2Q3Z) was generated using Modeller.<sup>[84]</sup> The appropriate sequence alignment was performed taking into account secondary structural motifs with PROMALS.<sup>[85]</sup> Subsequently, significant improvement of the starting model could be achieved by automated improvement of a molecular replacement model using ARP/wARP.<sup>[86]</sup> The model was manually improved with Coot<sup>[87]</sup> and refined with Phenix.<sup>[88]</sup> At the beginning, a simulated annealing step was performed. The structure was refined with coordinates, individual B-factors, occupancies, TLS groups, optimized X-ray/stereochemistry weight and optimized X-ray/ADP weight. The ligand was optimized with the Tripos Force Field in Sybyl.<sup>[89]</sup> The ligand restraints were created with the Monomer Library Sketcher implemented in the CCP4 Program Suite. According to Procheck,<sup>[90]</sup> 1107 residues (90 %) are found in the most favored, 120 residues (9.8 %) in allowed, and 3 residues (0.2 %) occur in disallowed regions (Asp 270 of molecule 1 and 2 and Asp 139 of molecule 1). Asp 270 is located between two coordinating residues of  $\text{Ca}^{2+}$  at site 3, enforcing a tight loop turn. Asp 139 of molecule 1 is located in a tight  $\beta$ -turn and is part of a three-stranded  $\beta$ -sheet in the  $\beta$ -sandwich domain. In the final difference electron density map residual density could be identified next to the amino acids Phe 60 and Lys 61. The peak is in H-bonding distance to the latter residues and approximates tetrahedral geometry. Several tetrahedral species were tested and a sulphate ion, possibly picked up from the crystallization buffer, gave the best fit. However, as some unexplained residual density remained after refinement we did not assign any molecular species to this density.

### 2.5.3 Molecular Modelling

The K- and Q-substrates were modeled as tripeptides Ala-Xxx-Ala (Xxx: Lys, Gln) based on the coordinates of the inhibitor in the crystal structure and placed such that the  $\epsilon$ -amino group of the lysine residue was within  $\sim 7 \text{ \AA}$  of  $\text{S}_\gamma$  of the catalytic cysteine. In the following HARMm22 force field minimization,<sup>[91]</sup> the protein structure was kept rigid, while full flexibility was allowed for both substrates. No restraints were included, such that the final position represents an unbiased minimum of the potential energy.

## 2.5.4 Synthesis of ZED1301

The peptidic blocker ZED1301 is a modified octapeptide derived from phage display screening<sup>[68]</sup> containing a Michael Acceptor as substrate glutamine mimetic. The C-terminal peptide sequence “NleNleLPWP” was synthesized by standard Fmoc solid-phase peptide chemistry (reactions in DMF, coupling with TBTU/HOBt/DIPEA, deprotection with piperidine) using H-Pro-2-ClTrt-resin as starting material. Coupling (in DMF, using HATU/DIPEA) with the Boc-protected Michael Acceptor amino acid was followed by cleavage from the resin (using 95 % TFA / 2.5 % water / 2.5 % triisopropylsilane). The resulting peptide “MA-NleNleLPWP-OH” was subsequently modified by coupling with Ac-Asp(OtBu)-OH (in DMF, using HATU/DIPEA). Finally, the tBu-protecting group was removed by TFA (in DCM) yielding Ac-Asp-MA-NleNleLPWP-OH. Purification was performed by reverse phase HPLC. Identity was confirmed by massspectrometry, purity was >97 %, as analyzed by HPLC at 214 nm.

## 2.6 Appendix

### 2.6.1 Alignment of the Catalytic Domain of Human Transglutaminases

An alignment of the human transglutaminase family shows that the calcium ion binding sites (site 1: yellow, site 2: green, site 3: cyan) are well conserved with exception of these residues interacting by their amide oxygen with the calcium ion. Also Arg 491 and Asp 438 involved in the interacting between the shifted  $\alpha$ -helix and the three-stranded  $\beta$ -sheet are highly conserved (grey). A high conservation is also presented in case of the catalytic site (red) consisting of a catalytic triad and a hydrophobic tunnel. The alignment indicates a fundamental function of the catalytic dyad (violet) within the entire transglutaminase family due to a high conservation of His 342 and Glu 401.

	190	200	210	220	230	240
	.	.	.	.	.	.
<b>F13A</b>	NPWCEDDAVYLDNEKERE	EYVLNDIGVIFYGEVNDI	KTRSWSYGQFEDGILDT	CLYVMDRA	.....	
TG1	NPWCPEDIVYVDHEDWR	QEYVLNESGRIYYGTEA	QIGERTWNYGQFDHGVLD	ACLYILDRR	.....	
TG2	NAWCPADAVYLDSEEE	RQEYVLTQQGFYQGS	AKFIKNIPWNFGQFED	GILDICLILLDVNPK	FLKNA	
TG3	NPWLNVDVFMGNHAERE	EYVQEDAGIIFVGSTNR	IGMIGWNFGQFEEDIL	SICLSILDRSLN	FRRDA	
TG4	NPWCKEDMVMPDEDER	KEYILNDTGCHYVGAAR	SIKCKPWNFGQFEKNV	LDCCLSLTES	.....	
TG5	NPWCPEDAVYLDSEPQR	QEYVMNDYGFYQGSKN	WIRPCPWNFGQFEDKI	IDICLKLLDKSL	HFQTD	
TG6	NPWCAEDDVFLASEEE	RQEYVLSDSGIIFRGV	EKHIRAQGWNYGQFEED	ILNICSILDRSPGH	QNNP	
TG7	NPWSPEDDVYLPSEILL	QEYIMRDYGFVYKGHER	FITSWPWNFGQFEEDI	IDICFEILNKSLY	HLKNP	

	250	260	270	280	290	300	310		
	.	.	.	.	.	.	.		
F13A	MDLSGRGNPIKVS	RVGSA	AMVNAKDD	EGVLVGSWDNI	YAYGVPPSA	WTGSVDILLEY	RSSE.NPVR	RYGQ	
TG1	MPYGGRGDPVNV	SRVIS	AMVNSLDD	NGVLIGNWSGD	YSRGTNP	SAWVGSVEILL	SYLRTG.Y	SVPY	GQ
TG2	RDCSRRSSPVYV	GRVVS	GMVNCNDD	QGVLLGRWDNN	YGDGVSPMS	WIGSVDILRR	WKNHGCQR	VKY	GQ
TG3	TDVASRNDPKYV	GRVLS	AMINSNDD	NGVLAGNWSGT	YTGGRDPR	SWNGSVEIL	KNWKKSGF	SPVR	YGQ
TG4	LKPTDRRDPVLV	CRAMC	AMMSFEKG	QGVLIGNWTGD	YEGGTAPY	KWTGSAPILQ	QYYNTK.Q	AVCF	GQ
TG5	TDCALRGSPVYV	SRVVC	AMINSNDD	NGVLNGNWS	ENYTDGAN	PAEWTGSVAIL	KQWNATGC	QPV	YGQ
TG6	TDVSCRHNPIYV	TRVIS	AMVNSNND	RGVVQGGWQ	GKYGGGT	SPLHWRGSVAIL	QKWLKGRY	KPV	YGQ
TG7	KDCSQRNDVVYV	CRVVS	AMINSNDD	NGVLQGNWGED	YSKGVSPLEW	KGSVAILQQ	WSARGGQ	VPK	YGQ
	320	330	340	350	360	370	380		
	.	.	.	.	.	.	.		
F13A	CWVFAGVFNTFL	RCLGIPARIV	TNYFSAHND	DANLQMDIF	LEEDGNVNS	KLTKDSVW	NYHCWNEA	WMT	
TG1	CWVFAGVTTTVL	RCLGLATRTV	TNFNSAHD	TDTSLTMDI	YFDENMKP	LEHLNHD	SVWNFHV	WNDCW	MK
TG2	CWVFAAVACTVL	RCLGIPTRVV	TNYSAN	QNSNLLIEY	FRNEFGEI	QGD.KSEMI	WNFHCW	VESW	MT
TG3	CWVFAGTLNTAL	RSLGIPSRVIT	NFNSAHD	TDRNLSVD	VYDPMGN	PLDK.GSD	SVWNFHV	WNEGW	FV
TG4	CWVFAGILTTVL	RALGIPARSVT	GFD	SAHDTERNLTV	DITYVNE	NGEKITSM	THDSVWN	FHVWTD	AWMK
TG5	CWVFAAVMCTVM	RCLGIPTRVIT	NFDSGH	DTDGNLTI	DEYDNTGR	ILGNKKK	DTIWNF	HVWNEC	WMA
TG6	CWVFAGVLCTVL	RCLGIATR	VSNFNSAHD	TDQNLSD	KYVDSFG	RTLLED	TESMWN	FHVWNES	WFA
TG7	CWVFASVMCTVM	RCLGVPTRV	VSNFRSAH	NDVDRNLTI	IDTYYDR	NAEMLSTQ	KRDKIWN	FHVWNEC	WMI
	390	400	410	420	430	440			
	.	.	.	.	.	.			
F13A	RPDLPGVFGGWQ	AVDSTPQ	ENS	DGMYRCGPAS	VQAIKHGHV	CFQFDAPFV	FAEVNSD	LIYITAK	KD..
TG1	RPDLPSGFDGWQ	VVDATPQ	ETSS	GIFCCGPCSV	ESIKNGLVY	MKYDTPFIF	AEVNSD	KVYWQR	QDD..
TG2	RPDLQPGYEGWQ	ALDPTPQ	EKSE	GYCCGPVPV	RAIKEGDL	STKYDAPFV	FAEVNAD	VVDWIQ	QDD..
TG3	RSDLGPSYGGWQ	VLDATPQ	ERSQ	GVFQCGPAS	VIGVREGD	VQLNFDMP	FIFAEV	NADRITW	LYDNT..
TG4	RPDLPGYDGWQ	AVDATPQ	ERSQ	GVFCCGPS	PLTAIRKGD	IFIVYDTR	FVSEVNG	DRLIWL	VKVMVNG
TG5	RKDLPPAYGGWQ	VLDATPQ	EMSN	GVYCCGPAS	VRAIKEGE	VDLNYDTP	PFVFSMV	NADCMS	WLQGG..
TG6	RQDLGPSYNGWQ	VLDATPQ	ESE	GVFRCGPAS	VTAREGDV	HLAHDGPF	VFAEVN	ADYITW	LWHED..
TG7	RKDLPPGYNGWQ	VLDPTPQ	ETSS	GLFCCGPAS	VKAIREGDV	HLAYDTPF	VYAEV	NADEVI	WLLGDG..



	450	460	470	480	490	500
	.	.	.	.	.	.
<b>F13A</b>	..GTHVV.ENVDA	THIGKLIVTKQIGGDGMM	DITDTYKFQ	EGQEEER	LALETALMYGAKKPLNTE...	
<b>TG1</b>	..GSFKI.VYVE	EKAIGTLIVTKAISSNM	REDITYLYKHP	EGSDAER	KAVETAAAHGSKPNVYA....	
<b>TG2</b>	..GSVHK.SINRS	LIVGLKISTKSVGRDER	EDITHYKYP	EGSSEER	EAFTRANHLNKL.....	
<b>TG3</b>	..TGKQWKNSVNS	HTIGRYISTKAVGSNARM	DVTDKYKYP	EGSDQER	QVFQKALGKLPNTPFAAT..	
<b>TG4</b>	GQEELHV.ISME	TTSIGKNISTKAVGQDR	RRDITYEYKYP	EGSSEER	QVMDHAFLLSSEREHRR...	
<b>TG5</b>	..KEQK..LHQDT	SSVGNFISTKSIQSDER	DDITENYKYE	EGSLQER	QVFLKALQKLKARSFHGSQ	RG
<b>TG6</b>	..ESRER.VYSNT	KKIGRCISTKAVGSDSR	VDITDLYKYP	EGSRKER	QVYSKAVNRLFGVEASGR	RIW
<b>TG7</b>	..QAQEI.LAHNT	TSSIGKEISTKMVGSDQ	RQSITSSYKYP	EGSPEER	AVFMKASRKMLGP.....	QRA
			510			
		.				
<b>F13A</b>	.....	G..	VMKSR			
<b>TG1</b>	.....	NRGSA				
<b>TG2</b>	.....	AEKEE				
<b>TG3</b>	.....	SSMGLETEEQE				
<b>TG4</b>	.....	PVKE				
<b>TG5</b>	AELQPSRPTSL	SQDSPRSLHTPS.LR	PSDV			
<b>TG6</b>	I.....	RRAGGRCLWRDD	LLEPATK			
<b>TG7</b>	SL.....	PF.....	LDLLESGG.LRD...			

## 2.6.2 Crystallographic Table

**Tab. 2.2:** Data collection and refinement statistics of inhibitor ZED1301.

<b>Inhibitor</b>	ZED1301
<b>PDB Entry</b>	4KTY
<b>Data collection and processing</b>	
Wavelength [Å]	0.91841
Beamline	BESSY – 14.2

Space group	P1
<b>Unit cell parameters</b>	
a, b, c [Å]	56.8, 80.5, 102.8
$\alpha$ , $\beta$ , $\gamma$ [°]	88.1, 76.7, 82.0
<b>Diffraction data</b>	
Resolution range [Å]*	50.00 – 1.98 (2.03 – 1.98)
Unique reflections	119138 (7864)
R(I) <sub>sym</sub> [%]	8.4 (36.3)
Completeness [%]	97.1 (95.8)
Redundancy	2.2 (2.1)
I/ $\sigma$ (I)	10.0 (2.0)
<b>Refinement</b>	
Resolution range [Å]	20.80 – 1.98
Reflections used in refinement (total)	119041
Reflections used in refinement (work/free)	113061/ 5980
Final R values for all reflections (work/free) [%]	16.9/ 20.6
Protein residues (A/B)	695/ 698
Atoms Inhibitor (A/B)	74/ 74
Water molecules	920
Other ligand atoms	
<b>RMSDs from ideality</b>	
Bond lengths [Å]	0.010
Bond angles [°]	1.3
<b>Ramachandran plot</b>	
Residues in most favored regions [%]	90.0
Residues in additional allowed regions [%]	9.7

Residues in generously allowed regions [%]	0.1
Residues in disallowed regions [%]	0.2
<hr/>	
<b>Mean B-factor [<math>\text{\AA}^2</math>]</b>	
<hr/>	
Protein total (A/B)	27.3/ 27.9
Protein main chain (A/B)	26.4/26.8
Protein side chain (A/B)	28.2/29.1
Inhibitor (A/B)	35.3/ 38.9
Calcium ions	21.9
Other ligand atoms	34.8
Water molecules	29.7

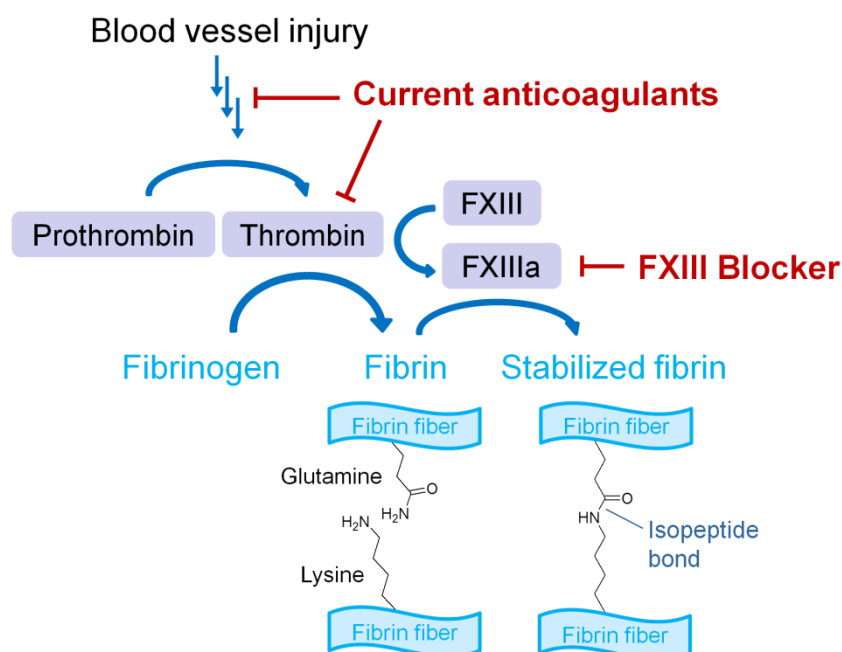
\*Highest resolution shell is shown in parenthesis

### 3 Structure-Based Design of FXIIIa-Blockers: Addressing the Hydrophobic Pocket of the “ $\alpha$ -Space”

#### 3.1 Introduction

Thromboembolic events such as strokes, pulmonary embolism and acute coronary syndrome are among the leading causes of mortality and morbidity.<sup>[92]</sup> The administration of anticoagulants reduces the risk of the occurrence of a thrombus in risk patients but unfortunately enhance the bleeding tendency can occur as an undesired side effect.

Thrombin represents the central enzyme of the blood coagulation cascade. Accordingly, all current anticoagulants lower thrombin activity, whether directly or indirectly (*Fig. 3.1*). Thrombin activates fibrinogen to fibrin that forms along with other blood cells a weak blood clot. Factor XIII, also activated by thrombin, finally stabilizes the clot by forming iso-peptide bonds between the fibrin fibers.<sup>[1, 41, 63]</sup> Accordingly, FXIII is the only enzyme in the coagulation cascade acting downstream of thrombin. Consequently, inhibition of FXIII would still allow the formation of a “soft” blood clot but would not interfere with e.g. platelet activation. These features rendering FXIII to a promising target for the development of anticoagulants with a lower bleeding risk compared to current anticoagulants. In addition to its involvement in blood coagulation, FXIII also decorates the clot with  $\alpha 2$ -antiplasmin, lowering fibrinolysis of the clot by plasmin.<sup>[93]</sup> As a result, inhibition of FXIII would lower both, half-life and weight, accelerating the degradation by patient’s own fibrinolytic systems.



**Fig. 3.1:** Mode-of-Action of FXIII blockers compared to current anticoagulants illustrated on a simplified representation of the blood coagulation cascade. Factor XIII is the last enzyme of the cascade and catalyzes the formation of an iso-peptide bond between a glutamine and a lysine residue. All current anticoagulants reduce the activity of thrombin. As a result, the coagulation and platelet activation is disturbed. In contrast, inhibition of FXIII would still allow the fibrin formation (that forms with blood cells a “soft” blood clot and does not interfere with platelet activation). Consequently, FXIII is a promising target for anticoagulants with an expected lower bleeding risk.

The occurrence of severe bleeding events in patients with congenital FXIII deficiency should be kept in mind.<sup>[94, 95]</sup> However, heterozygous patients show only a mild or even none bleeding phenotype. Ultimately, clinical trials have to show whether there is an improved benefit-to-risk ratio compared to current anticoagulants or not. Looking at the unique mode-of-action it seems to be realistic that a direct-acting FXIIIa blocker could provide an additional treatment option at least for certain patients. To prove this hypothesis druggable molecules are needed to perform pharmacological studies first in relevant animal models and – if successful – in future clinical trials on humans.

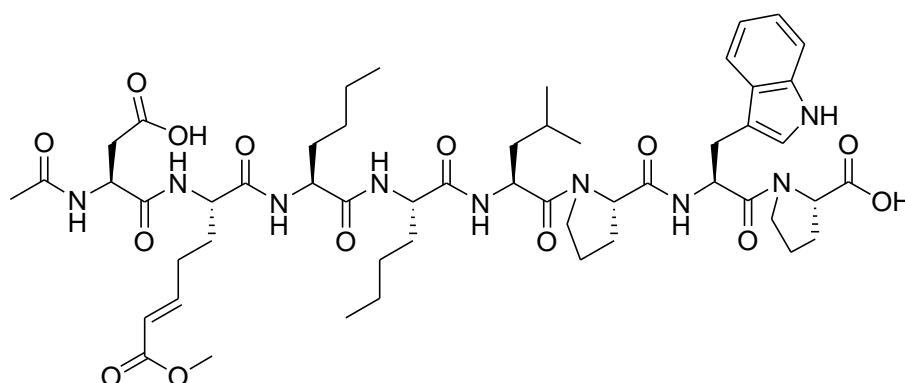
A promising strategy to obtain such druggable compounds follows the approach of structure-based drug design. So far, all attempts to develop a FXIII blocker failed due to the lack of a suitable crystal structure. The crystal structure of factor XIII in the inactive state, published in

1994, was an important academic milestone for the transglutaminase community. Unfortunately, the structure is not useful for structure-based design of a FXIII blocker because the active site is buried by the  $\beta$ -barrel 1 domain and the binding mode of substrates and inhibitors remained unclear.<sup>[17]</sup>

In 2013, we published a high-resolution crystals structure of FXIII in the active conformation by co-crystallization of FXIII with a covalently-attached ligand.<sup>[60]</sup> Herein, the  $\beta$ -barrel 1 domain of the protein does not covers the active site any longer but is located laterally of the catalytic domain.

Despite of an excellent binding affinity of our lead structure ZED1301, the observed plasma half-life is short. This attributes primarily to the peptidic structure rendering the molecule prone to digestion by proteases. Furthermore, the molecular weight of 1,063 g/mol exceeds by far the Lipinski rule of five (500 g/mol) indicating properties favorable for oral application. Thus, we aimed to reduce the size of our inhibitors and to introduce a more non-peptidic character. We decided firstly to optimize the affinity in the “ $\alpha$ -space”. As  $\alpha$ -space the region next to the catalytic tunnel is defined.

ZED1301 is an octapeptide consisting of an N-terminal acetylated aspartate followed by the warhead replacing in the substrate a glutamine residue. Next, three hydrophobic amino acids (two norleucines and a leucine) are present bridging the distance to the primary binding pocket which is addressed by a proline-tryptophan dipeptide. Another proline residue completes the peptidic lead structure (*Fig. 3.2*). Since the N-terminal aspartate displays most interactions within the active site, we decided to focus on the replacement of this amino acid by alternative residues.

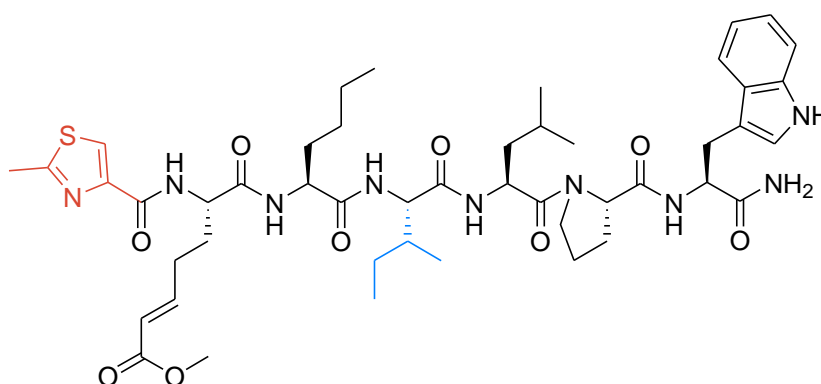


**Fig. 3.2:** The peptidic lead structure ZED1301 possesses an acetylated aspartate at the N-terminus addressing the “ $\alpha$ -space”.

## 3.2 Results & Discussion

### 3.2.1 ZED1630

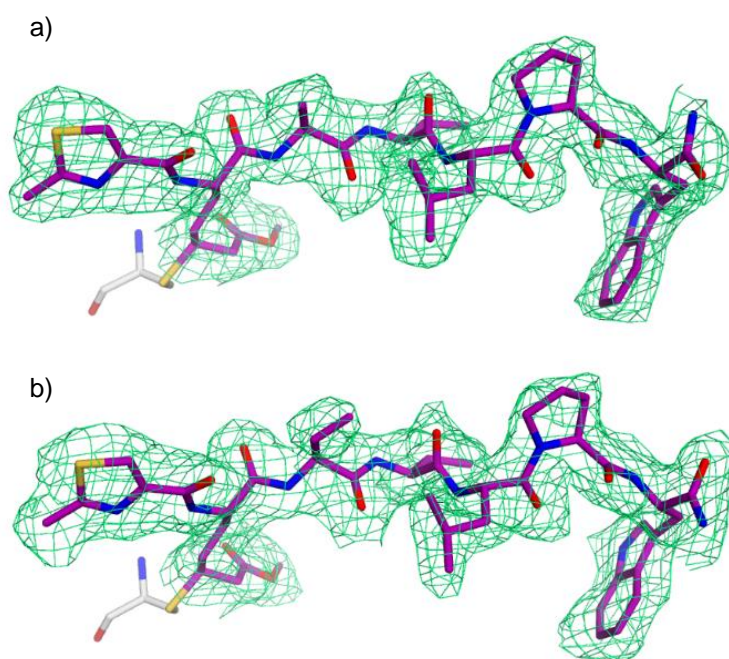
The lower-affinity inhibitor ZED1630 ( $IC_{50}$ : 139 nm) possesses a methylated thiazole heterocycle at the N-terminus. Additionally, ZED1630 differs from ZED1301 in that way the second norleucine is substituted by an isoleucine and the terminal proline was dispensed.



**Fig. 3.3:**  $\alpha$ -Space modification (red) of ZED1630 and further modifications (blue) compared to the peptidic lead structure ZED1301.

As shown in Fig. 3.4, the inhibitor ZED1630 is clearly defined in the electron density of both symmetry-related molecules (resolution: 2.10 Å). Only for the first norleucine there remain three and two carbon atoms in molecule 1 and 2 of the asymmetric unit undefined in the electron density, respectively.

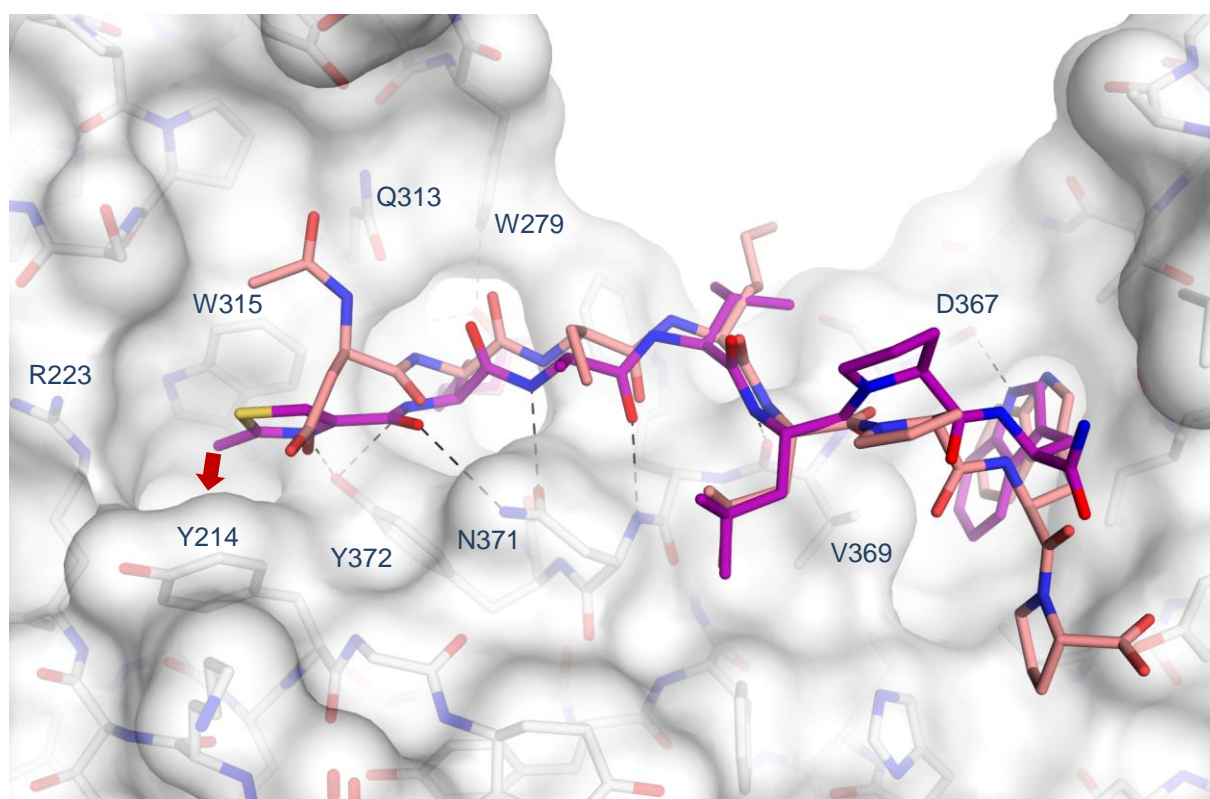




**Fig. 3.4:**  $F_o - F_c$  difference electron density contoured at  $3\sigma$  for the inhibitor ZED1630 (purple) in molecule 1 (a) and 2 (b) of the asymmetric unit. The inhibitor is covalently attached to the catalytic site Cys 314 (grey).

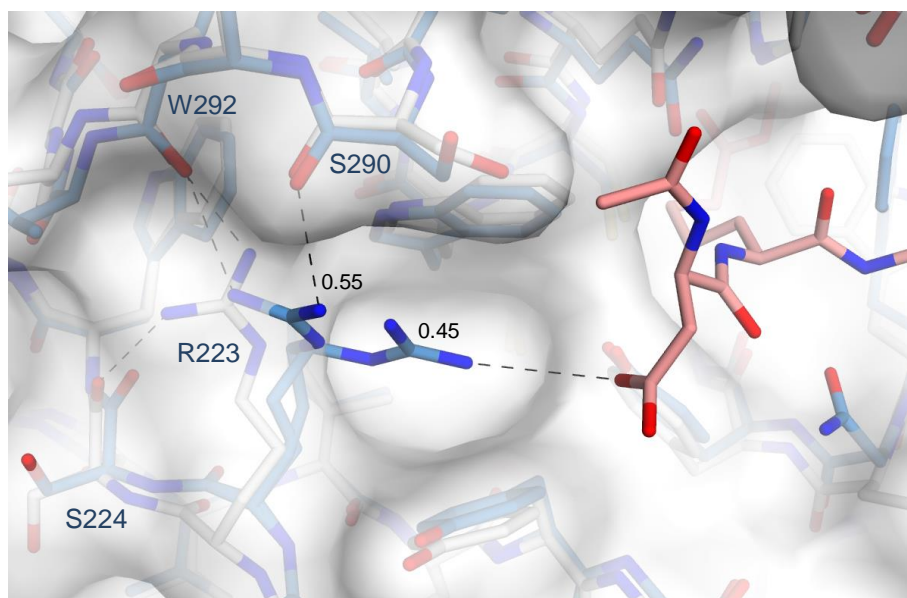
Regarding the crystal structure, the thiazole is located as expected in the  $\alpha$ -space (Fig. 3.5). Notably, the methyl group is exactly positioned above a hydrophobic pocket in the  $\alpha$ -space. Addressing this hydrophobic pocket can result in an increase of the binding affinity. Furthermore, in the  $\alpha$ -space ZED1630 binds significantly closer to the protein surface compared to ZED1301 (difference of the distances between  $C_\alpha$  of the N-terminal aspartate and the corresponding atom of ZED1630 and the hydroxyl oxygen Tyr 372 amounts 1.3 Å), suggesting for an improved binding geometry.

Finally, the indole ring of ZED1630 binds deeper into the hydrophobic pocket (1 Å) and the indole nitrogen forms a comparatively strong hydrogen bond to the carbonyl oxygen of Asp 367 with a distance of 2.9 Å. It should be mentioned that Asp 367 belongs to the residues coordinating calcium at binding site 2 (Chapter 2.3.2) and Asp 367 adopts this position only after calcium has bound to the enzyme. Also the proline next to the tryptophan adopts a different spatial orientation in the binding modes of both inhibitors ( $C_\alpha$ -distance: 1.4 Å). These conformational deviations of ZED1630 and ZED1301 may result from the lack of a terminal proline in ZED1630.



**Fig. 3.5:** Interaction pattern of ZED1630 (magenta, PDB ID: 5MHM) with FXIIIa<sup>o</sup> compared to ZED1301 (light red, PDB ID: 4KTY). The methyl group of a thiazole ring located in the  $\alpha$ -space can serve as starting point for addressing the hydrophobic pocket (red arrow). The indole ring reaches deeper into the hydrophobic pocket due to the lacking terminal proline.

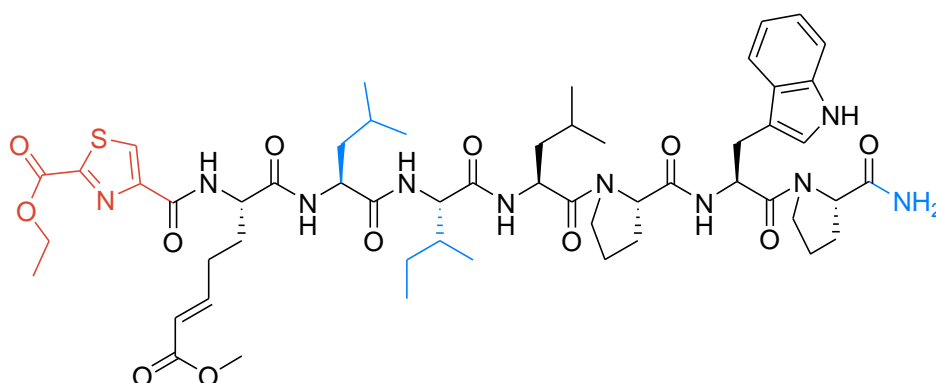
Indeed, the appearance of the hydrophobic pocket in the  $\alpha$ -space is not surprising after considering the fact that upon enzyme activation amino acids of the  $\alpha$ -space remain at their position (Fig. 3.6). The inhibitor, described in Chapter 2.3.3, interacts with the side chain of Arg 223, resulting in a sealing of the hydrophobic pocket. However, the corresponding H-bond is according to its length, rather weak (3.6 Å) and the new conformation of Arg 223 has to compete with hydrogen bonds formed to the backbone oxygen of Ser 224 and Trp 292. Consequently, the side chain of Arg 223 adopts two conformations with a similar occupancy. Interestingly, in crystal structures of FXIIIa<sup>o</sup> where the hydrophobic pocket is established (ZED1630, ZED2360, ZED2369), Arg 223 adopts the same conformation as the higher populated conformation of Arg 223 in FXIIIa<sup>o</sup>-ZED1301 (not shown).



**Fig. 3.6:** The hydrophobic pocket in the  $\alpha$ -space is already formed in the inactive state of FXIII (grey, PDB ID: 1F13). In the crystal structure of FXIIIa° in complex with ZED1301 (light red, PDB ID: 4KTY) colored in blue, a weak H-bond (3.6 Å) between the aspartate oxygen of ZED1301 and the guanidinium nitrogen of Arg 223 triggers a closure of the hydrophobic pocket. The side chain of Arg 223 adopts two conformations. One interacts with the backbone oxygen of Ser 290 and Trp 292, the other one with ZED1301, whereas the latter has a slightly lower occupancy of 0.45. Notably, compared to the crystal structure of FXIIIa°-ZED1301 in the inactive state the side chain of Arg 223 donates its hydrogen bond to the backbone oxygen of Ser 224 instead of Ser 290.

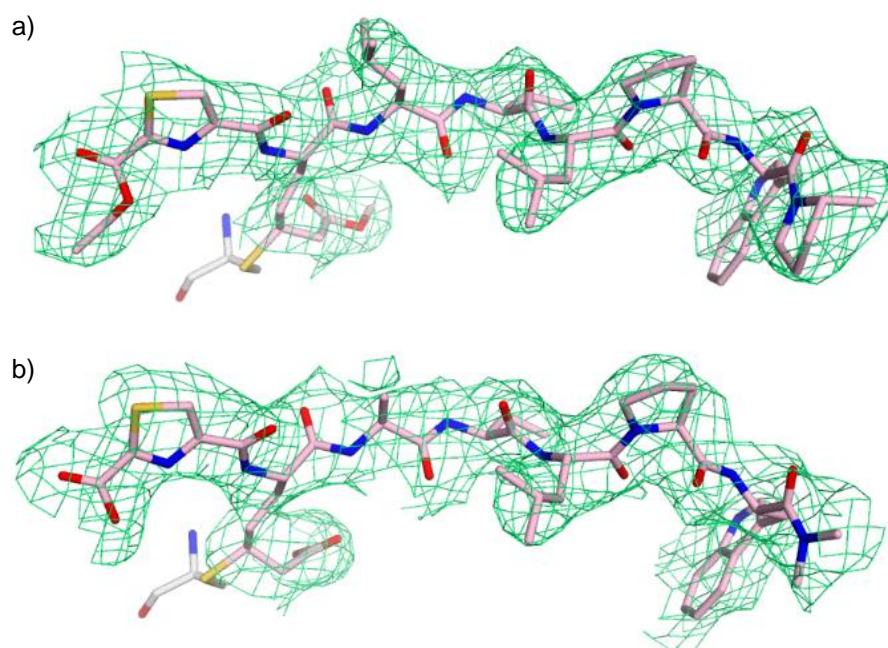
### 3.2.2 ZED2360

The inhibitor ZED2360 was synthesized to address the hydrophobic pocket in the  $\alpha$ -space. This was attempted by enhancing the methylated thiazole by attaching an ethyl ester (*Fig. 3.7*). Furthermore, when compared to ZED1301, ZED2360 bears a leucine and an isoleucine instead of the two norleucines. Finally, the C-terminus is a carboxamide.



**Fig. 3.7:**  $\alpha$ -space modification (red) of ZED2360 and further modifications (blue) compared to the peptidic lead structure ZED1301.

The collected dataset of the crystallized complex of ZED2360 and FXIIIa<sup>o</sup> shows a moderate resolution of 2.48 Å. Nevertheless, the inhibitor with some exceptions is decisively shaped by the difference electron density (Fig. 3.8). The first inhibitor molecule in the asymmetric unit is almost complete visible except for the terminal carboxamide group and the leucine side chain. In the second molecule of the asymmetric unit, some atoms remain undefined in the difference density (ethyl group of the thiazole ester, methyl group of the warhead ester, terminal proline and some atoms of leucine and isoleucine side chain).

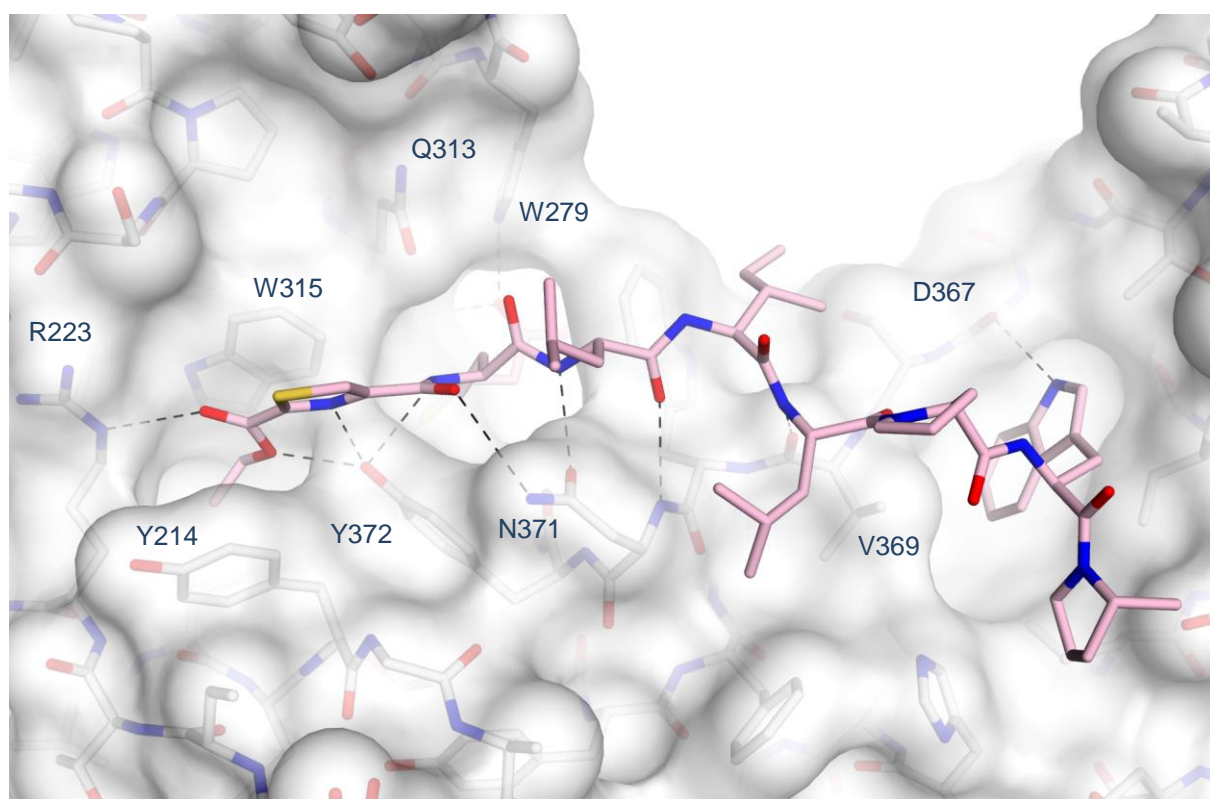


**Fig. 3.8:**  $F_o-F_c$  difference electron density contoured at  $3\sigma$  for the inhibitor ZED2360 (light purple) in molecule 1 (a) and 2 (b) of the asymmetric unit. The inhibitor is covalently attached to the catalytic site Cys 314 (grey).

As anticipated, the crystal structure shows that the ethyl group of the ester introduced at the thiazole moiety addresses the hydrophobic pocket in the  $\alpha$ -space (Fig. 3.9). Moreover, the ester carbonyl oxygen serves as acceptor for the guanidinium NH of Arg 223 and the other ester oxygen accepts a hydrogen bond of the phenolic hydroxyl group of Tyr 372. The other hydrogen bonds are similarly established to ZED1630 (Fig. 3.5).

Considering the geometry and the interaction pattern, the thiazole ester fits perfectly into the complementary binding pocket of FXIIIa<sup>o</sup> within the  $\alpha$ -space. This is also reflected by a significantly improved binding affinity. The  $IC_{50}$  of ZED1630 exceeds with 29 nM that of ZED1301 by almost an order of magnitude.

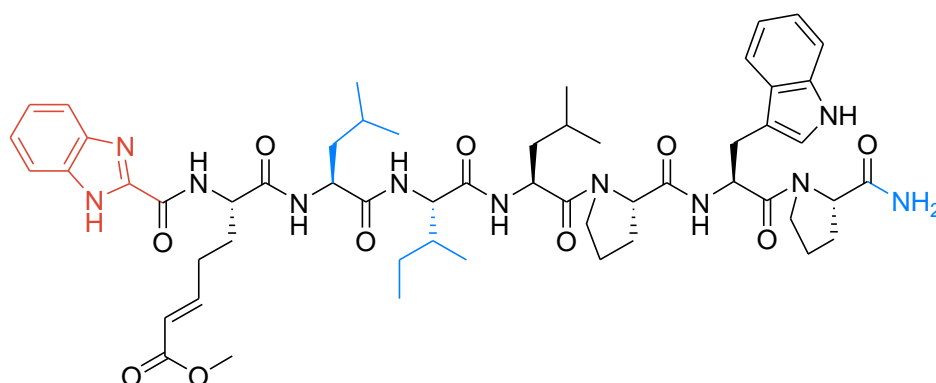




**Fig. 3.9:** Interaction pattern of ZED2360 (cyan, PDB ID: 5MHN) with FXIIIa°. The ethyl group of the thiazole ester occupies the hydrophobic pocket. The ester oxygen atoms interact via H-bonds with Arg 223 and Tyr 372.

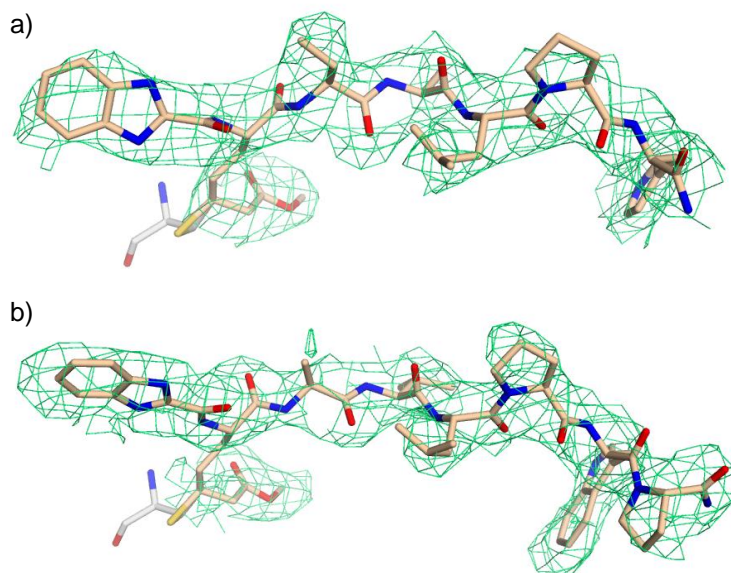
### 3.2.3 ZED2369

A further  $\alpha$ -space modified inhibitor possesses a bulky and hydrophobic benzimidazole ( $IC_{50}$ : 102 nM). The remaining substitution pattern corresponds to that of ZED2360 (Fig. 3.10). Also here, the two norleucines are replaced by leucine and isoleucine, respectively. The C-terminus is equally amidated.



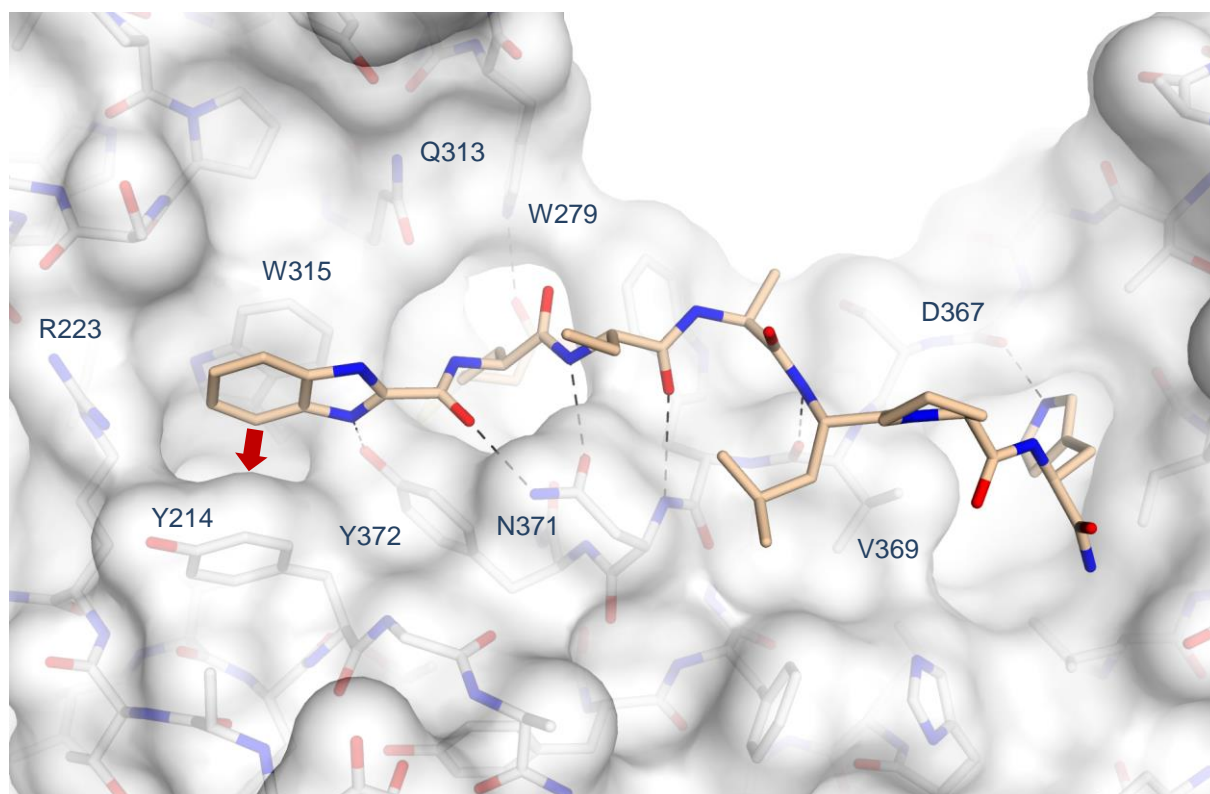
**Fig. 3.10:**  $\alpha$ -space modification (red) of ZED2369 and further modifications (blue) compared to the initial ligand ZED1301.

Owing to the moderate resolution of 2.92 Å, some atoms remain undefined in the density (Fig. 3.11) when compared to the other better resolved structures. This also affects the indole ring of the first molecule in the asymmetric unit that is not clearly defined in the difference electron density map. Furthermore, the benzimidazole adopts two different orientations considering the two symmetry-related molecules, although a reliable assignment of the conformational properties is difficult in that part of the structure due to limited resolution.



**Fig. 3.11:**  $F_o-F_c$  difference electron density contoured at  $3\sigma$  for the inhibitor ZED2369 (beige) in molecule 1 (a) and 2 (b) of the asymmetric unit. The inhibitor is covalently attached to the catalytic site Cys 314 (grey).

However, the crystal structure reveals that the benzimidazole moiety is located next to the  $\alpha$ -space (Fig. 3.12). Importantly, the carbon atom at position 4 is directly positioned above the hydrophobic pocket and can serve as starting point for addressing the hydrophobic pocket, with additional substituents, analogous to ZED1630.



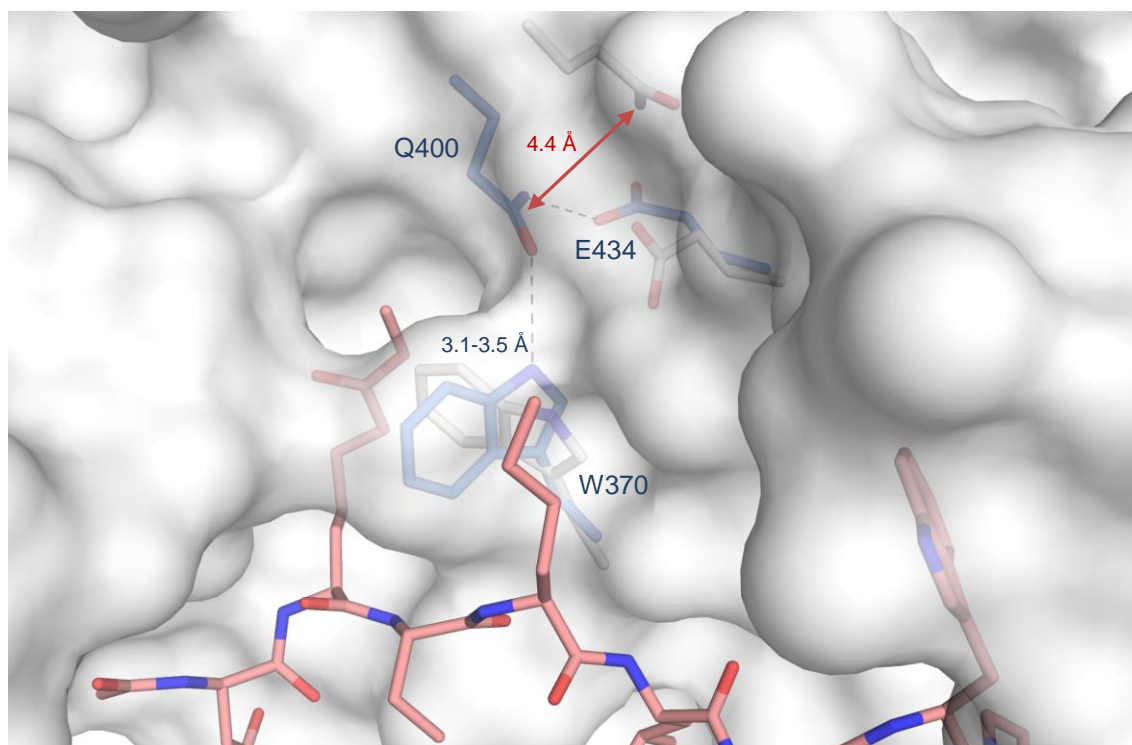
**Fig. 3.12:** Interaction pattern of ZED2369 (cyan, PDB ID: 5MHO) with FXIIIa<sup>o</sup>. The benzimidazole is located in the  $\alpha$ -space and can also be used for occupying the hydrophobic pocket (red arrow).

### 3.2.4 Indole Ring of the Hydrophobic Tunnel Affected by Calcium Binding

Comparing the new crystal structures described above, an H-bond between the indole NH of Trp 370 and the carboxamide oxygen of Gln 400 attracts attention (because of a shorter distance with up to 3.1 Å). This is as compared to the first crystal structure of FXIIIa<sup>o</sup> which showed a rather weak interaction with 3.5 Å (Fig. 3.13). Importantly, the orientation of this indole ring is responsible for the formation of the hydrophobic tunnel, protecting the thioester intermediate from hydrolysis.



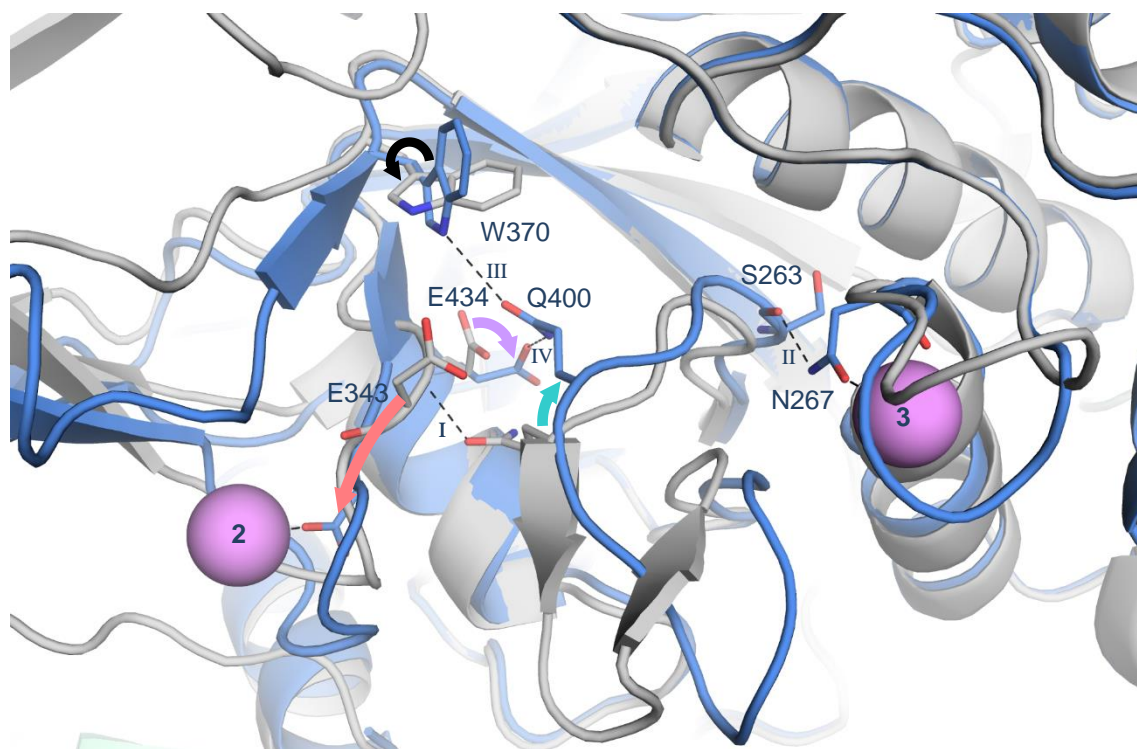
Additionally, Gln 400 also interacts with its carboxamide NH<sub>2</sub> with the carboxyl oxygen of Glu 434, rendering the carboxamide oxygen a stronger hydrogen acceptor. Remarkably, both residues adopt this arrangement only in the active state of FXIII (*Fig. 3.13*), whereas the distance of the carboxamide of Gln 400 increases from the active and inactive state to a value of 4.4 Å.



**Fig. 3.13:** The indole nitrogen of Trp 370 forms an H-bond with the carboxamide oxygen of Gln 400. The superimposed structure of the inactive state (grey sticks, PDB ID: 1F13) shows that this arrangement exists only in the active state (blue sticks, PDB ID: 4KTY). The distance between the carboxamide oxygen of Gln 400 and the indole nitrogen of Trp 370 differs between the different crystal structures from 3.1 to 3.5 Å. During transition from the active to the inactive state, the side chain carboxamide group of Gln 400 is translocated by 4.4 Å (red arrow).

Closer inspection of the crystal structure of FXIIIa reveals that the calcium binding sites 2 and 3 are involved in the formation of this tunnel simultaneously affecting the arrangement of Gln 400 and Glu 434 (*Fig. 3.14*). Coordination of Ca<sup>2+</sup> at binding site 2 drags a loop downwards (red arrow), as already described for the formation of the catalytic dyad (*Chapter 2.3.2*), thus disrupting an H-bond between the backbone NH of Glu 343 and the carboxamide oxygen of

Gln 400. Along with a loop rearrangement induced by the calcium coordination at site 3, the adjacent loop moves upwards (turquoise arrow) and enables the interaction of Gln 400 and the indole nitrogen of Trp 370 to be formed. This newly adopted position of the side chain of Gln 400 also leads to a conformational rearrangement of the side chain of Glu 434 (purple arrow).



**Fig. 3.14:** Superposition of the inactive state (grey) and the active state (blue) of FXIII reveals that  $\text{Ca}^{2+}$  (purple spheres) coordination at binding sites 2 and 3 triggers the H-bond formation between Gln 400 and the indole moiety of Trp 370. During transition from the inactive to the active state the indole ring rotates by around  $70^\circ$  (black arrow). The interaction of Glu 343 with calcium (red arrow) breaks a hydrogen bond (I) between Gln 400 and the backbone nitrogen of Glu 343. This in turn enables along with calcium coordination at site 3 (H-bond II) a loop rearrangement (turquoise arrow) and brings Gln 400 and the indole of Trp 370 in an appropriate H-bond interaction distance (III). Finally, the side chain of Glu 434 adapts to the new chemical environment by a side chain flip (purple arrow) and forms a corresponding hydrogen bond (IV).

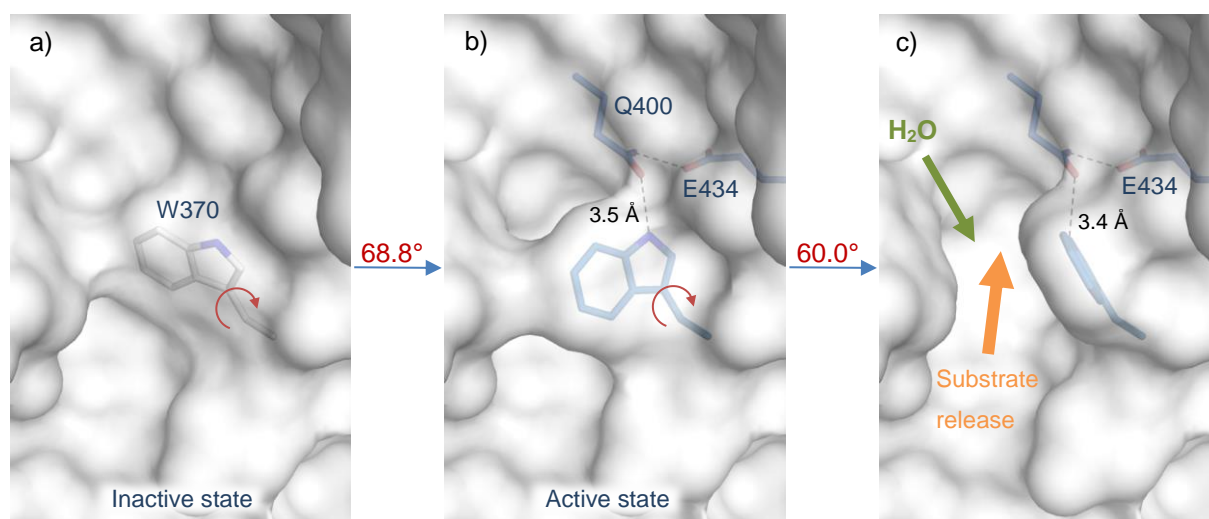
Regarding the amino acid sequences of human transglutaminases, both residues (Gln 400 and Glu 434) are highly conserved (*Fig. 3.15*). Consequently, the formation of this dyad and the interaction with the indole ring of the hydrophobic tunnel should have a special function with respect to the catalytic cycle of transglutaminases. An opening of the hydrophobic tunnel prior to the glutamine side chain penetration of the substrate can be excluded because in the calcium bound crystal structure of transglutaminase 3, where Gln 400 and Glu 434 adopts the same spatial arrangement as in FXIIIa<sup>o</sup>, the indole ring is identically oriented as in the inactive state of FXIII.<sup>[18, 75, 96]</sup>

	390	400	410	420	430
<b>F13A</b>	GGWQAVDSTPQ	ENS	DGMYRCGPASVQA	IKHGHVCFQ	FDAPFVFAEVNS
TG1	DGWQVVDATPQ	ETSSGIFCCG	PCSVESIKNGLV	MYKYDTPFI	FAEVNS
TG2	EGWQALDPTPQ	EKSEGTGCCG	PVPVRAIKEGDL	STKYDAPFV	FVAEVNA
TG3	GGWQVLDATPQ	ERSQGVFQCG	PASVIGVREGDV	QLNFDMPFI	FAEVNA
TG4	DGWQAVDATPQ	ERSQGVFCCG	PSPLTAIRKGDIF	IVYDTRFVF	SEVNG
TG5	GGWQVLDATPQ	EMSNGVYCCG	PASVRAIKEGEVD	LNVDTPFVF	FSMVNA
TG6	NGWQVLDATPQ	EESQGVFRCG	PASVTAIREGDV	HLAHDGPFV	FVAEVNA
TG7	NGWQVLDPTPQ	QTSSGLFCCG	PASVKAIREGDV	HLAYDTPFV	YAEVNA

**Fig. 3.15:** Section of an alignment of human transglutaminases. The amino acids Gln 400 and Glu 434 are conserved (highlighted in cyan) with exception of TG5.

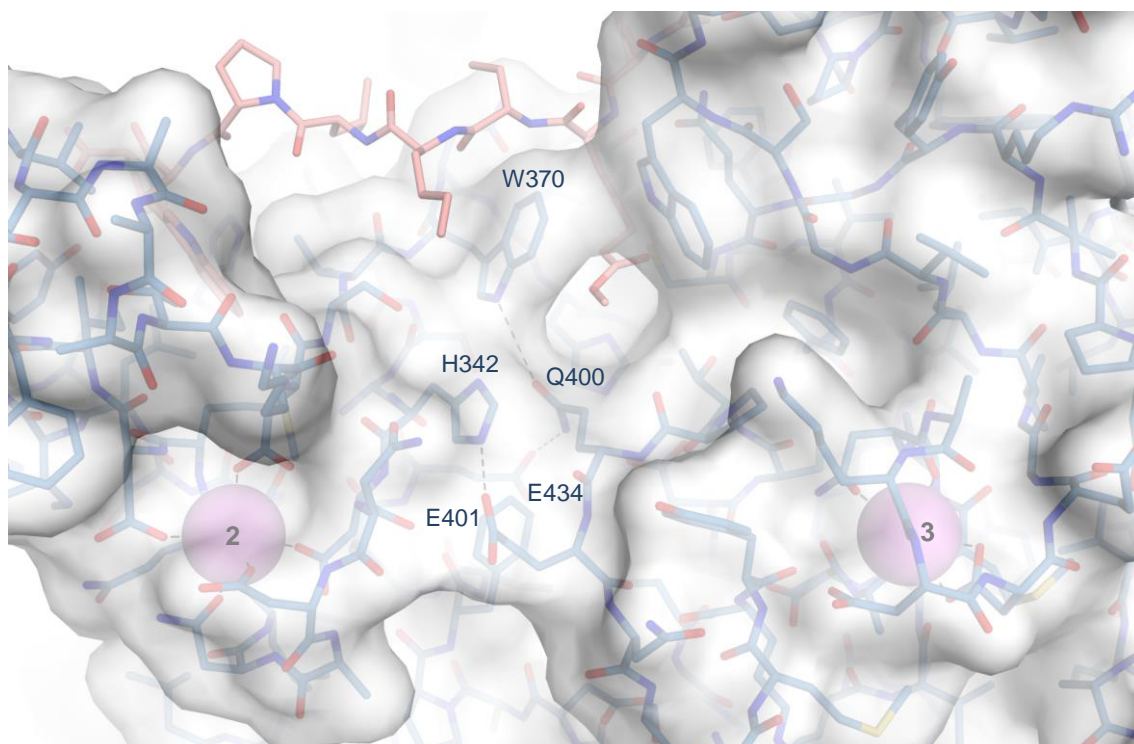
Alternatively, the formation of the hydrophobic tunnel could take place in a concerted process along with substrate binding, whereas the H-bond between Gln 400 and Glu 434 facilitates the opening of the tunnel.

As illustrated in *Fig. 3.16*, a further role of the established H-bond could be anticipated to prevent further rotation of the indole ring leading to an extended opening of the hydrophobic tunnel which might result in the hydrolysis of the thioester intermediate. However, rotation of the indole ring by 60° would result in a shortening of the hydrogen bond distance (3.4 Å).



**Fig. 3.16:** Indole rotamers of Trp 370. a) In the inactive state of FXIII (PDB ID: 1F13), the catalytic site is buried by the indole ring of Trp 370 (shown: FXIIIa° with the rotamer of Trp 370 of the inactive state of FXIII). b) Upon substrate binding the indole ring rotates by 68.8° and establishes formation of the hydrophobic tunnel (PDB ID: 4KTY). Simultaneously, an H-bond between the indole nitrogen and the carboxamide oxygen is formed (3.5 Å). c) Further rotation of the indole ring (here 60°) would open the tunnel for release of the iso-peptide cross-linked product. The hydrogen bond would become slightly stronger during opening of the hydrophobic tunnel (3.4 Å). Opening of the hydrophobic pocket would also enable water to enter the catalytic site and hydrolyzes the thioester intermediate (deamidation).

Transglutaminases are known to catalyze apart from the transamidation also deamidation resulting in the hydrolysis of a glutamine to a glutamate.<sup>[1]</sup> Regarding the crystal structures of all active human transglutaminases, it has been demonstrated that calcium binding predominantly affects the shape of the co-substrate binding site and renders the enzyme prominent to catalyze the transamidation reaction (Fig. 3.17).<sup>[60]</sup> Consequently, in contrast to the K-substrate the Q-substrate should be able to bind to the active site of transglutaminases also in the absence of calcium.



**Fig. 3.17:** Calcium binding site 2 and 3 (purple spheres) affects the shape of the co-substrate binding site along with the formation of the catalytic dyad (His 342, Glu 401) and the H-bond between Gln 400 with the indole nitrogen of Trp 370.

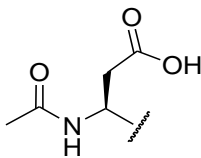
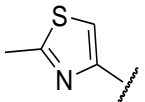
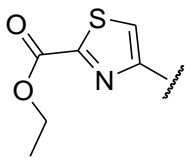
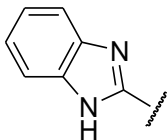
Additional to the not established co-substrate binding site<sup>[60]</sup>, in the absence of calcium the catalytic dyad is not formed as well as the H-bond between the indole of Trp 370 and Gln 400. Thus, the thioester intermediate should be hydrolyzed by water molecules penetrating into the hydrophobic tunnel. Considering all these observations, the calcium dependency could be a regulatory element of transglutaminases to control the preference of transamidation over deamidation. A preferred deamidation with respect to transamidation at a lower pH value has already been described by Fleckenstein et al.<sup>[97]</sup> and corroborates the assumption of a dependency of the transamidation-deamidation-ratio on chemical environmental factors such as the local calcium concentration. Finally, the hypothesis with respect to a preferred deamidation at a lower calcium level has to be evidenced experimentally.



### 3.3 Summary & Conclusion

The initial affinity of inhibitor ZED1301 could be improved by optimizing the occupancy of the  $\alpha$ -space (Tab. 3.1). Substitution of the acetylated aspartate of ZED1301 by a methyl thiazole group (ZED1630) provides a starting point for addressing the hydrophobic pocket of the  $\alpha$ -space by adding an ethyl ester substituent (ZED2360) of which the ethyl group occupies the hydrophobic pocket. As a result, the affinity could be increased to 29 nM. Finally, the attached benzimidazole portion turned out to be useful for addressing the hydrophobic pocket via an alternative scaffold. However, it has to be considered that the inhibitors do not only modified the  $\alpha$ -space. Nevertheless, the added modifications take only minor impact on the affinity with exception of the N-terminal proline (missing in ZED1630).

**Tab. 3.1:** Inhibitors with their modifications in the  $\alpha$ -space, affinity and the resolution of the corresponding crystal structure.

Inhibitor	$\alpha$ -space moiety	Affinity [nM]	Resolution [Å]
ZED1301		110	1.98
ZED1630		139	2.10
ZED2360		29	2.48
ZED2369		102	2.92

Additionally, with reference to the improvement in binding affinity, a highly conserved dyad consisting of Gln 400 and Glu 434 was found. The Gln 400 carboxamide oxygen involves the NH hydrogen of Trp 370 indole into an H-bond in the hydrophobic tunnel. Remarkably, calcium binding is involved in the formation of the dyad. However, the exact function of the hydrogen bond between the indole of Trp 370 and the carboxamide of Gln 400 remains unclear and requires experimental investigations.

## 3.4 Experimental Part

### 3.4.1 Data Collection and Structure Determination

Crystals of FXIIIa° in complex with the inhibitors were flash-frozen in liquid nitrogen and a datasets were collected at the BESSY II electron storage ring.<sup>[79]</sup> All data sets were processed in space group P1 using the programs HKL2000<sup>[80]</sup> (ZED1630) and XDS<sup>[98]</sup> (ZED2360, ZED2369). The initial phases were calculated by Fourier synthesis using the structure (PDB ID: 4KTY) as starting model in the first refinement step with Phenix.<sup>[88]</sup> The model was manually improved with Coot<sup>[87]</sup> and refined with Phenix. At the beginning the structure was refined as rigid body, whereas both molecules of the asymmetric unit were split into the four domains. Subsequently, a simulated annealing was performed. All structures were refined with coordinates, occupancies and individual B-factors (group B-factor for ZED2369 because of the low resolution). Restraints of the ligands were generated by GRADE.<sup>[99]</sup>



## 3.5 Appendix

### 3.5.1 Crystallographic Table

**Table 3.1:** Data collection and refinement statistics of inhibitor ZED1630 and ZED2360.

Inhibitor	ZED1630	ZED2360
PDB Entry	5MHM	5MHN
<b>Data collection and processing</b>		
Wavelength [Å]	0.91841	0.91841
Beamline	BESSY 14.1	BESSY 14.1
Space group	P1	P1
<b>Unit cell parameters</b>		
a, b, c [Å]	56.9, 80.7, 103.2	56.6, 80.5, 103.1
$\alpha$ , $\beta$ , $\gamma$ [°]	88.4, 76.6, 81.8	88.0, 77.0, 82.3
<b>Diffraction data</b>		
Resolution range [Å]*	25.00 – 2.12 (2.17 – 2.12)	48.12 – 2.48 (2.63 – 2.48)
Unique reflections	97508 (6273)	60898 (9684)
R(I)sym [%]	8.3 (40.4)	6.8 (46.9)
Completeness [%]	97.1 (94.1)	97.4 (96.1)
Redundancy	2.2 (2.1)	2.2 (2.2)
I/ $\sigma$ (I)	10.5 (2.0)	9.83 (2.0)
<b>Refinement</b>		
Resolution range [Å]	24.88 – 2.12	48.12 – 2.48
Reflections used in refinement (total)	97449	60829
Reflections used in refinement (work/free)	92550/ 4899	57787/ 3042

Final R values for all reflections (work/free) [%]	16.9/ 21.3	18.0/ 23.8
Protein residues (A/B)	684/ 683	688/ 685
Atoms Inhibitor (A/B)	62/ 63	75/ 65
Water molecules	841	193
Calcium ions	6	6
Other ligand atoms	63	43
<b>RMSDs from ideality</b>		
Bond lengths[Å]	0.009	0.010
Bond angles [°]	1.0	1.1
<b>Ramachandran plot</b>		
Residues in most favored regions [%]	89.6	88.5
Residues in additional allowed regions [%]	10.0	10.9
Residues in generously allowed regions [%]	0.1	0.3
Residues in disallowed regions [%]	0.3	0.2
<b>Mean B-factor [Å<sup>2</sup>]</b>		
Protein total (A/B)	30.6/ 28.7	53.1/ 56.3
Protein main chain (A/B)	29.6/ 27.3	53.1/ 55.8
Protein side chain (A/B)	31.6/ 30.3	53.1/ 56.9
Inhibitor (A/B)	34.9/ 35.6	64.2/ 70.4
Calcium ions	26.0	60.2
Other ligand atoms	36.3	57.7
Water molecules	32.4	44.0

\*Highest resolution shell is shown in parenthesis

**Table 3.2:** Data collection and refinement statistics of inhibitor ZED2369.

<b>Inhibitor</b>	ZED2369
<b>PDB Entry</b>	5MHO
<b>Data collection and processing</b>	
Wavelength [Å]	0.91841
Beamline	BESSY 14.1
Space group	P1
<b>Unit cell parameters</b>	
a, b, c [Å]	56.6, 80.5, 102.9
$\alpha$ , $\beta$ , $\gamma$ [°]	88.2, 76.9, 82.2
<b>Diffraction data</b>	
Resolution range [Å]*	48.11 – 2.92 (3.09 – 2.92)
Unique reflections	37113 (5987)
R(I)sym [%]	8.7 (46.0)
Completeness [%]	97.2 (96.6)
Redundancy	2.2 (2.2)
I/ $\sigma$ (I)	10.7 (2.0)
<b>Refinement</b>	
Resolution range [Å]	17.71 – 2.92
Reflections used in refinement (total)	36458
Reflections used in refinement (work/free)	34683/ 1775
Final R values for all reflections (work/free) [%]	21.7/ 24.8
Protein residues	678/ 676
Atoms Inhibitor	60/ 72
Water molecules	107

Calcium ions	6
Other ligand atoms	15
<b>RMSDs from ideality</b>	
Bond lengths[Å]	0.004
Bond angles [°]	0.7
<b>Ramachandran plot</b>	
Residues in most favored regions [%]	87.6
Residues in additional allowed regions [%]	11.7
Residues in generously allowed regions [%]	0.7
Residues in disallowed regions [%]	0.0
<b>Mean B-factor [Å<sup>2</sup>]</b>	
Protein total (A/B)	46.5/ 47.9
Protein main chain (A/B)	46.0/ 47.1
Protein side chain (A/B)	47.1/ 48.8
Inhibitor (A/B)	53.5/ 65.5
Calcium ions	52.1
Other ligand atoms	52.1
Water molecules	34.6

\*Highest resolution shell is shown in parenthesis

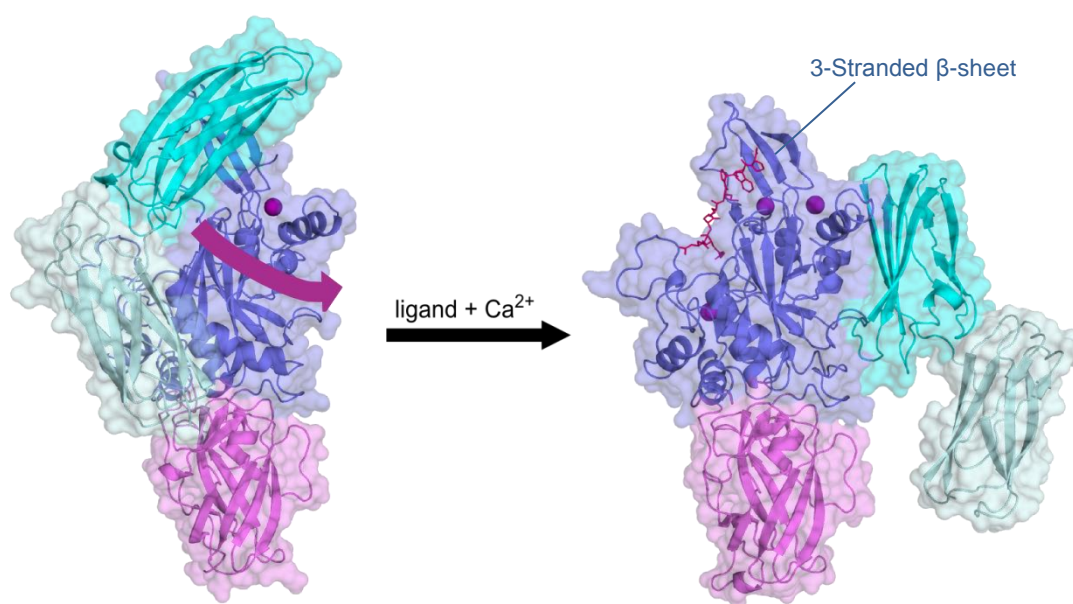
## 4 Blood Coagulation Factor XIII – An Enzyme with a Highly Adaptive Binding Site

### 4.1 Introduction

Blood coagulation factor XIII (FXIII) plays an important role in the coagulation process of vertebrates. Blood vessel damage initiates a cascade of zymogene activations whereby a couple of clotting enzymes are involved. Thrombin, as the central enzyme of the cascade, converts fibrinogen to fibrin that forms together with blood cells like erythrocytes and platelets the initial blood clot. Factor XIII, concomitantly activated by thrombin, stabilizes the clot by forming isopeptide bonds between the fibrin fibers. Furthermore, antifibrinolytic proteins are incorporated rendering the clot stable against premature fibrinolysis. Inhibition of FXIII still allows clot formation and platelet activation highlighting this enzyme as a promising target for the development of safe anticoagulants with a lower bleeding risk compared to current therapeutic compounds.<sup>[66]</sup>

In contrast to all other clotting enzymes, FXIII is not a serine protease but belongs to the transglutaminase family. In total, there are eight human iso-forms in this enzyme class. The characteristic reaction catalyzed by these enzymes is the formation of an isopeptide bond between the carboxamide group of glutamine and the  $\epsilon$ -amino nitrogen of lysine. First, the highly nucleophilic thiolate of the catalytic site residue Cys 314 attacks the carboxamide carbon atom of the protein-bound glutamine substrate forming a thioester. In a second step, the isopeptide bond is formed by a nucleophilic attack of the thioester intermediate on the lysine side chain of the protein co-substrate.

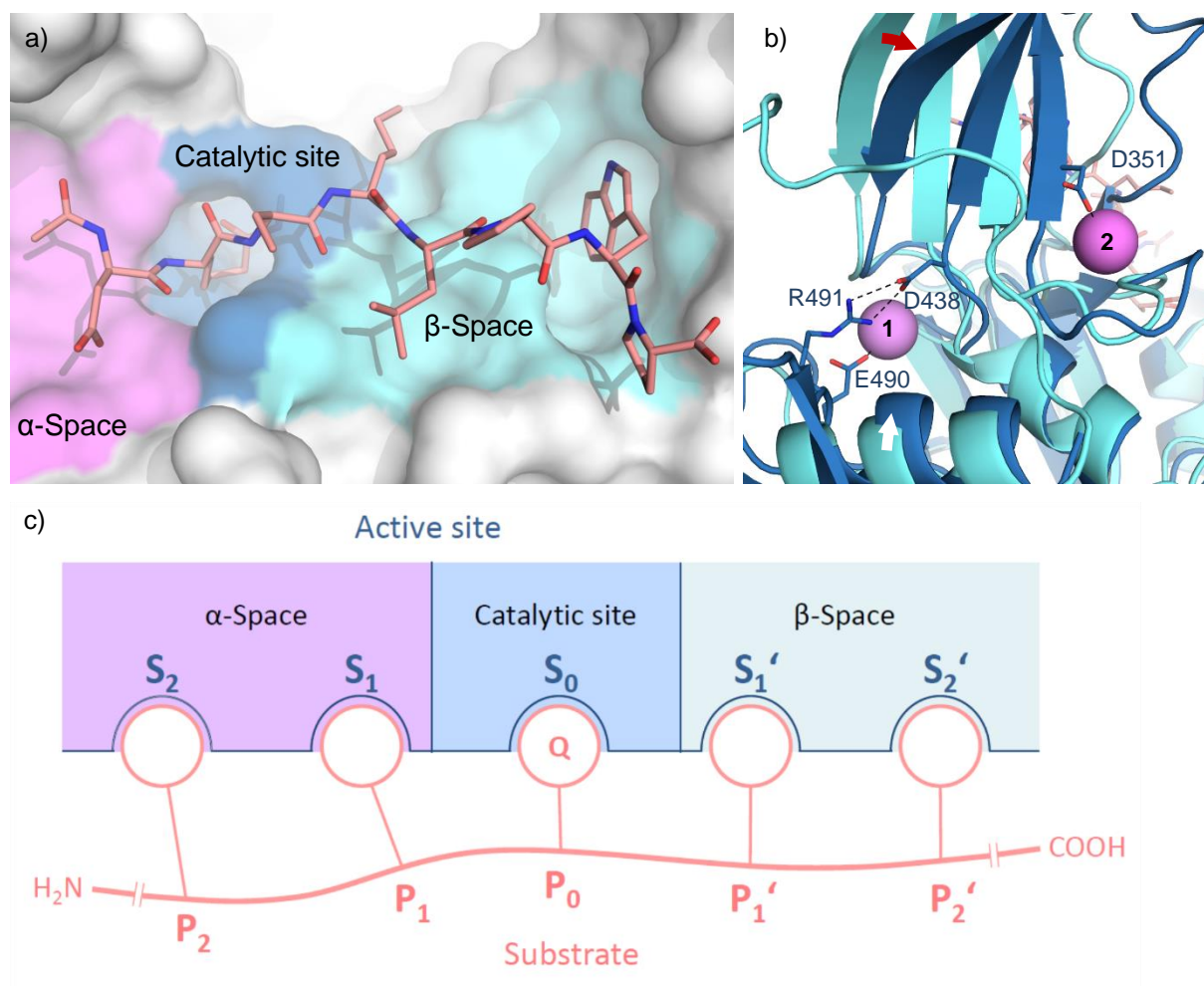
Human transglutaminases like FXIII are composed of a  $\beta$ -sandwich domain, a catalytic domain and two  $\beta$ -barrel domains (*Fig. 4.1*). The crystal structures of FXIII and transglutaminase 2 in complex with substrate-derived inhibitors confirm the early assumption that transglutaminases undergo a drastic conformational change during the catalytic process.<sup>[19, 60, 71]</sup> Functionally and structurally, the three-stranded  $\beta$ -sheet within the core domain appears to be a key element of factor XIII, and due to a high conservation of the involved amino acids, presumably also of the other human transglutaminases.



**Fig. 4.1:** FXIII and other human transglutaminases consist of a  $\beta$ -sandwich domain (purple), a catalytic core domain (blue), a  $\beta$ -barrel 1 domain (cyan) and a  $\beta$ -barrel 2 domain (light blue). In presence of calcium ions and an inhibitor (here: ZED1301) the two  $\beta$ -barrel domains flip aside (arrow) and expose the active site. The homodimeric (here only one monomer is shown) inactive state (left) of FXIII (PDB ID: 1GGU) coordinates one calcium ion (purple sphere). In contrast, the active form (right) has three calcium ion binding sites (PDB ID: 4KTY). ZED1301 (red sticks) binds to the active site located in the upper left part of the catalytic domain. The three-stranded  $\beta$ -sheet represents a key structural element of FXIII.

In our crystal structure of FXIIIa<sup>o</sup> in the active conformation published in 2013, an inhibitor (ZED1301) was used that possesses an  $\alpha,\beta$ -unsaturated carboxylic ester as reactive head group to bind covalently to Cys 314.<sup>[60]</sup> This residue belongs to the catalytic center of factor XIII (colored in blue in Fig. 4.2). The carboxylated aspartate moiety of the ligand binds to a region of the active site of FXIIIa<sup>o</sup> referred to as “ $\alpha$ -space” (purple). The indole ring of the inhibitor occupies a hydrophobic pocket in a second part of the active site referred to as “ $\beta$ -space” (light blue). For discussion and comparison of inhibitors and their binding mode, it is helpful to use a more precise nomenclature. Doiphode et al. suggested using the same nomenclature for proteases introduced by Schechter and Berger.<sup>[100, 101]</sup> However, we suggest designating the substrate glutamine as P<sub>0</sub> and not as P<sub>1</sub> (Fig. 4.2c). This nomenclature was already applied by Nikolajsen and coworkers.<sup>[53]</sup> Residues progressing towards the N-terminus are designated as

$P_1$ ,  $P_2$ , etc. and residues progressing to the C-terminus are designated as  $P_1'$ ,  $P_2'$ , etc. The corresponding interaction pockets to host the substrate side chains are labeled with an S instead of P and they are referred to with the same number as the corresponding substrate side chains. The hydrophobic tunnel of the catalytic site is occupied by the reactive glutamine ( $P_0$ ) or the inhibitor warhead respectively and is referred to as  $S_0$ .

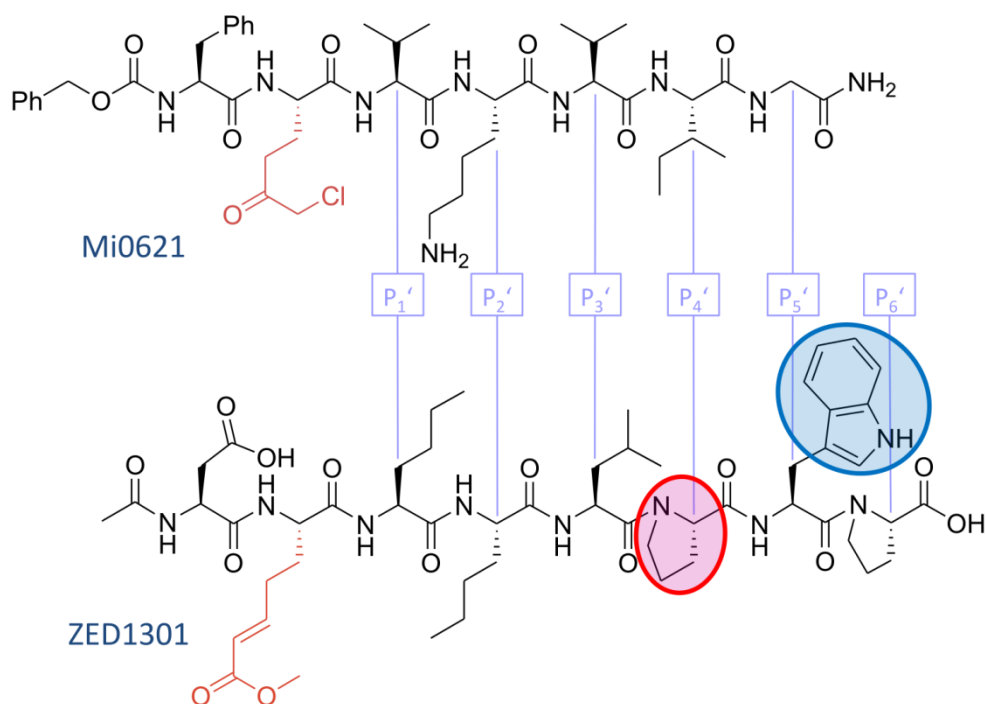


**Fig. 4.2:** a) The inhibitor ZED1301 is covalently attached to Cys 314 (not shown) of the catalytic site colored in blue. The acetylated (Ac protecting group) aspartate binds to the “ $\alpha$ -space” (purple) and the indole ring occupies a hydrophobic pocket in the “ $\beta$ -space” (light blue). b) Regarding the inactive state (cyan) and the active state (blue) of FXIII shows that both calcium binding sites 1 and 2 (purple spheres) promote a rotational movement of the three-stranded  $\beta$ -sheet (red arrow) being part of the substrate binding site. Calcium coordination at site 1 leads to a shift of an  $\alpha$ -helix (white arrow) that brings Arg 491 and Asp 438 in an appropriate distance to form a salt bridge and fixes the three-stranded  $\beta$ -sheet. c) Classification of the substrate and the active site of FXIII.

Transglutaminases are calcium-dependent enzymes. The exact function of calcium could be revealed by comparing the crystal structure of FXIII in the active versus inactive state (*Fig. 4.2b*).<sup>[17, 60, 74]</sup> Calcium binding site 1 is already occupied in the inactive state. Upon activation, the adjacent  $\alpha$ -helix is translocated by about 4 Å. This enables the formation of a strong salt bridge between Arg 491 of the  $\alpha$ -helix and Asp 438 fixing the three-stranded  $\beta$ -sheet at its lower part. Importantly, both amino acids are highly conserved within human transglutaminases.<sup>[63]</sup> Calcium binding site 3 induces conformational changes of the co-substrate binding site. Calcium binding site 2 turns out to be the most important one for the activation of the protein. Full establishment of calcium coordination at this site affects the shape of the  $\beta$ -space and facilitates the formation of the catalytic dyad needed for deprotonation of the amino group of the lysine co-substrate. The interaction between the calcium ion at binding site 2 and Asp 351 promotes a rotational movement of a three-stranded  $\beta$ -sheet located in the upper part of the catalytic domain (*Fig. 4.1* and *Fig. 4.2*) which is also part of the  $\beta$ -space of the substrate binding region.

Here we have crystallized FXIIIa<sup>o</sup> (FXIII activated by 50 mM calcium instead of thrombin) in complex with a  $\beta$ -casein derived inhibitor Mi0621<sup>[102]</sup> that differs remarkably from our former inhibitors (*Fig. 4.3*). Interestingly, Mi0621 possesses neither the central proline (red circle) nor the tryptophan (blue circle). Both amino acids are crucial for the affinity of this ligand prototype. In case one of these amino acids is removed, affinity declines dramatically.



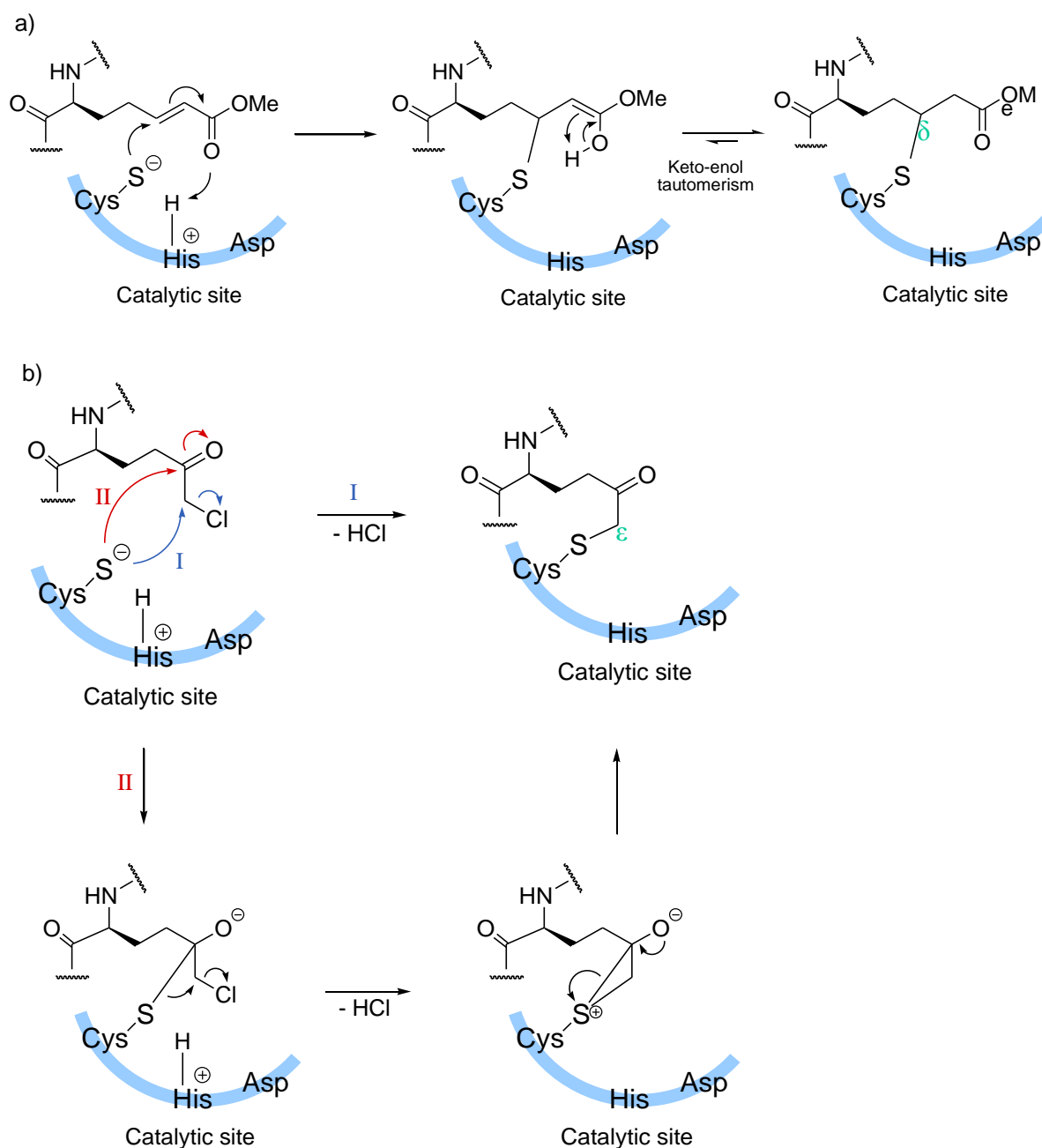


**Fig. 4.3:** Comparison of the inhibitors Mi0621 and ZED1301. Mi0621 lacks both the central proline residue (red circle) at P<sub>4</sub>' position and the P<sub>5</sub>' tryptophan (blue circle), essential for affinity in case of the ligand class ZED1301 belongs to. The inhibitors also differ in the reactive group colored in red. In case of ZED1301, an  $\alpha,\beta$ -unsaturated carboxylic ester was utilized whereas Mi0621, comprises a chloromethyl ketone at the same position.

The inhibitors also differ in both, the reactive group and in the way they bind to the catalytic residue Cys 314 (Fig. 4.4). The  $\alpha,\beta$ -unsaturated carboxylic ester of ZED1301 is attacked nucleophilically by the thiolate of Cys 314 at the  $\beta$ -position of the Michael acceptor facilitated by a hydrogen bond between the catalytic site His 373 and the carbonyl oxygen of the warhead. Subsequently, the intermediately formed methoxy enol tautomerizes to the corresponding ester (Fig. 4.4a). It should be mentioned that this reactive group is also used in ZED1227, a direct acting tissue transglutaminase inhibitor that is currently in Phase Ib clinical trials.<sup>[103]</sup>

In contrast, Mi0621 bears a chloromethyl ketone. Here two mechanisms are discussed in literature.<sup>[104]</sup> Either the thiolate sulfur of Cys 314 nucleophilically attacks the carbon atom between the carbonyl group and the chloro atom followed by chloride elimination (mechanism I in Fig. 4.4b) or the thiolate attacks initially the carbonyl carbon atom (mechanism II in

*Fig. 4.4b*). The thus formed thiohemiketal rearranges by forming a three-membered sulfonium alcoholate intermediate upon release of hydrogen chloride. Subsequently this complex rearranges in a second step to the alkylated cysteine product. Due to its high reactivity compared to the  $\alpha,\beta$ -unsaturated carboxylic ester, the chloromethyl ketone warhead is not considered as part of applicable drug molecules. Furthermore, the cysteine sulfur is connected to the C $_{\epsilon}$  of the warhead-containing amino acid. In contrast, in case of ZED1301, the catalytic site cysteine is covalently attached to the carbon atom at the  $\delta$ -position, whereas Mi0621 reacts at the  $\epsilon$ -position. This offset by one C-C bond length in the latter case deviates from the situation found during the reaction of the natural substrate.



**Fig. 4.4:** Simplified mechanism for the covalent attachment of the  $\alpha,\beta$ -unsaturated carboxylic ester and the chloromethyl ketone as warheads for enzyme inhibition. a) The  $\alpha,\beta$ -unsaturated carboxylic ester is nucleophilically attacked by the sulfur atom of Cys 314. A tautomerization of the intermediately formed enol generates the covalently attached thioester. b) The chloromethyl ketone is nucleophilically attacked by the thiolate of Cys 314 either directly at the carbon atom at the  $\epsilon$ -position followed by elimination of a chloride ion (mechanism I) or at the carbonyl carbon atom (mechanism II). Afterwards the thiohemiketal rearranges to a sulfonium intermediate and subsequently to the final thioether (b).

*Interestingly, the former  $\alpha,\beta$ -unsaturated carboxylic ester is linked via the  $\delta$ -position, the addition of the chloromethyl ketone occurs more remotely in  $\epsilon$ -position.<sup>[104]</sup>*

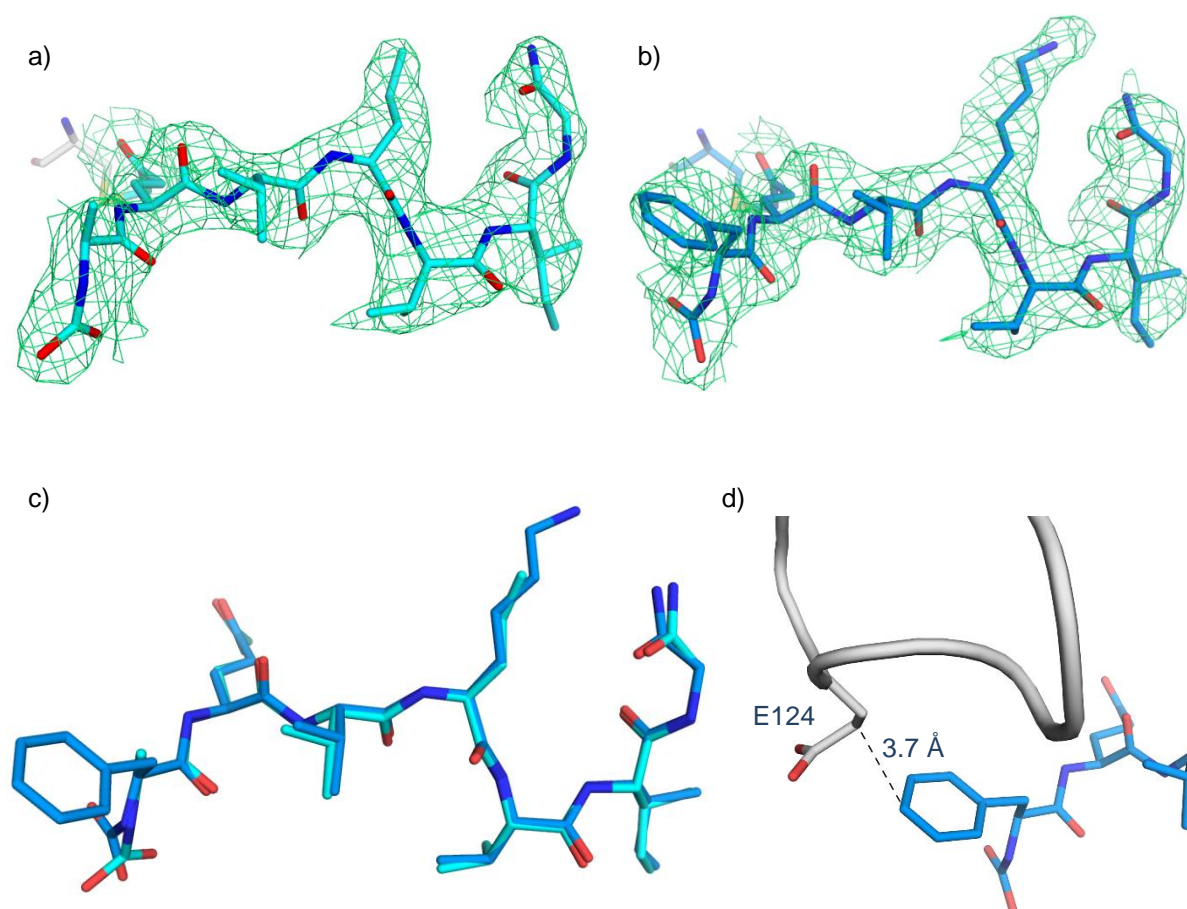
## 4.2 Results & Discussion

### 4.2.1 Binding Mode of the Inhibitor Mi0621

The crystallization of FXIIIa<sup>o</sup> in complex with the inhibitor Mi0621 was found to be challenging due to its lower affinity ( $IC_{50} = 12.5 \mu M$ ) when compared to inhibitor ZED1301 from the other compound class ( $IC_{50} = 110 nM$ ). However, a sufficiently complete dataset was collected with a resolution of  $2.40 \text{ \AA}$  (*Tab. 4.1*).

In the crystal structure, two symmetry independent molecules are found. The first inhibitor molecule in the asymmetric unit is well defined in the electron density contoured at a level of  $3.0 \sigma$  with the exception of the carboxy benzyl protection group, the benzyl group of the N-terminal phenylalanine and the  $\epsilon$ -nitrogen atom of the lysine residue (*Fig. 4.5a*). These side chains seem to be solvent exposed and consequently are not visible in the density due to enhanced scatter over multiple orientations.

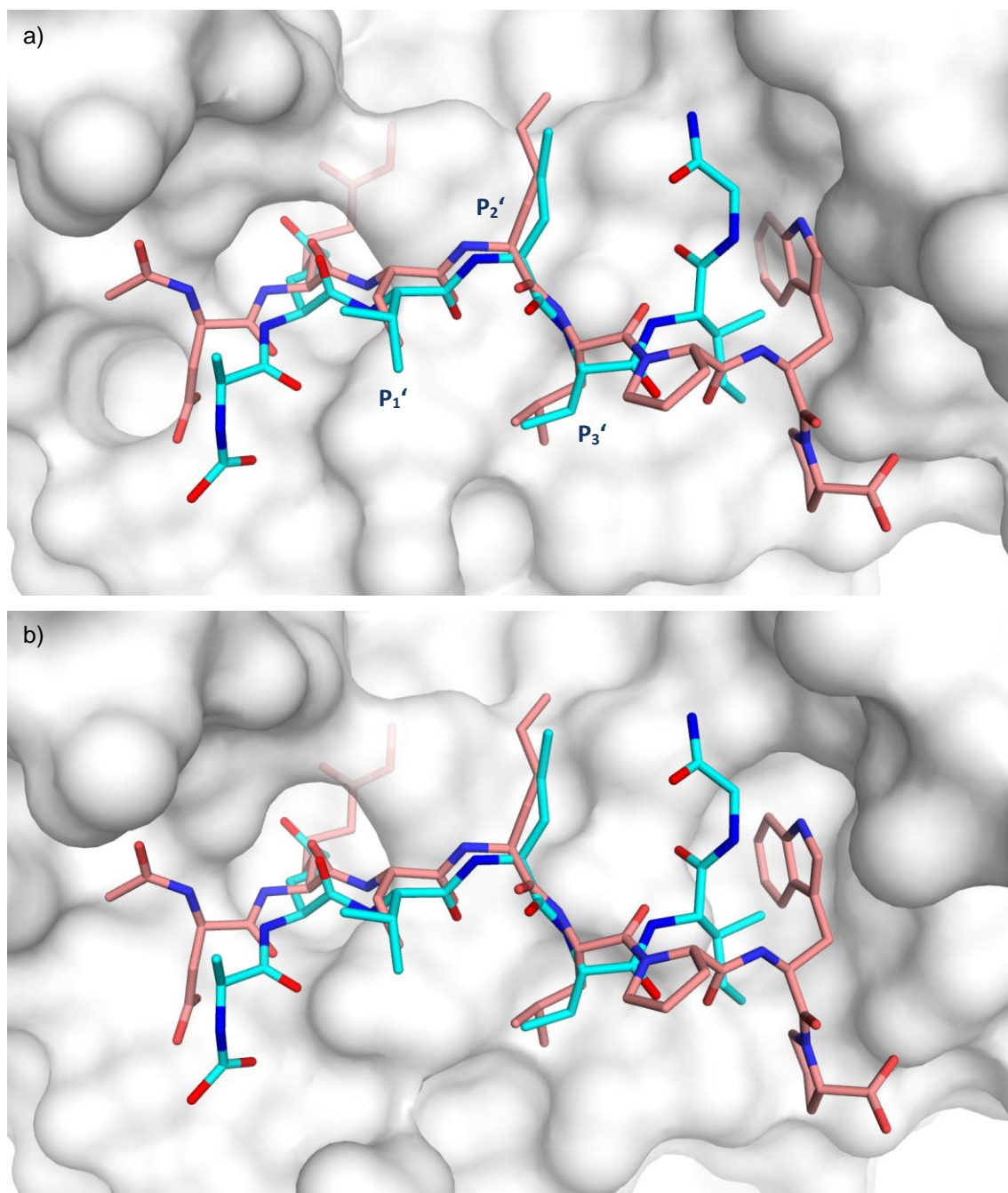
The second molecule in the asymmetric unit is slightly better defined in the electron density at a level of  $3.0 \sigma$  (*Fig. 4.5b*). Here the benzyl group and the lysine  $\epsilon$ -nitrogen are well defined in the difference electron density. However, the  $\epsilon$ -nitrogen atom of the lysine side chain forms no hydrogen bond to the protein. The phenyl ring of the second molecule in the asymmetric unit is located in van der Waals contact distance ( $3.7 \text{ \AA}$ ) to Glu 124 contributed by a symmetry-related protein molecule in the crystal packing (*Fig. 4.5d*). A superposition of the two symmetry independent molecules shows that both inhibitors adopt almost identical conformational arrangements with an rmsd over all nonhydrogen atoms of  $0.46 \text{ \AA}$  (*Fig. 4.5c*).



**Fig. 4.5:** a) The first molecule of inhibitor Mi0621 in the asymmetric unit is well defined by the difference electron density (contour level of  $3.0\sigma$ ) with exception of the carboxy benzyl protection group, the benzyl group of the N-terminal phenylalanine and the lysine side chain. b) In case of molecule 2 in the asymmetric unit, the phenyl ring and the lysine nitrogen atom are additionally well defined in the electron density. c) A superposition of the two symmetry independent molecules shows that both (molecule 1, cyan and molecule 2, blue) are geometrically identical with exception of the lysine and the phenylalanine side chain. d) The phenyl ring of the second inhibitor molecule in the asymmetric unit is in van der Waals distance to  $C_\gamma$  of Glu 124 from a symmetry-related protein molecule.

The crystal structure of FXIIIa<sup>o</sup> in complex with Mi0621 exhibits striking changes in the shape of the ligand binding site compared to the crystal structures of FXIIIa<sup>o</sup> with ZED1301 (Fig. 4.6). As already described in Chapter 3.2.2, the  $\alpha$ -space possesses a hydrophobic pocket that seems to remain unoccupied by the ligand. This hydrophobic pocket is opened owing to

the flexibility of the side chain of Arg 223. More interestingly, the hydrophobic pocket in the  $\beta$ -space is not opened upon ligand accommodation.



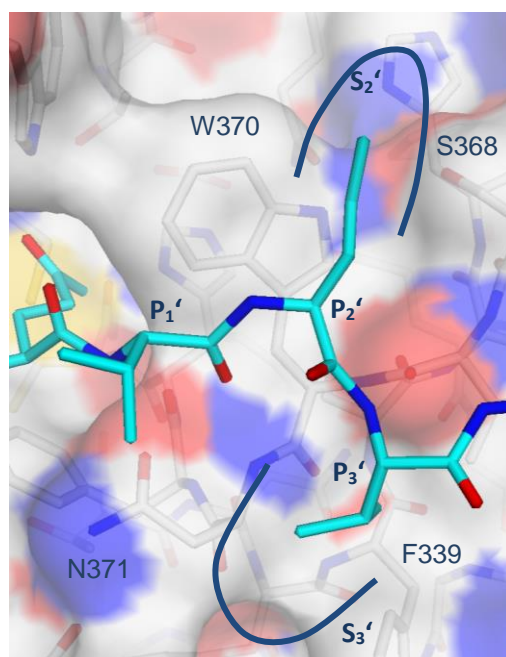
**Fig. 4.6:** Comparing the surface of the active site of FXIIIa° in complex with Mi0621 (cyan, PDB ID: 5MHL) and ZED1301 (light red, PDB ID: 4KTY). The protein surface of FXIIIa° in complex with Mi0621 (a) shows that the hydrophobic pocket in the  $\beta$ -space is not opened in contrast to the situation in the crystal structure of FXIIIa° in complex with ZED1301 (b), in consequence the indole ring of the

*superimposed ligand ZED1301 would plunge into the surface of FXIIIa° using the coordinates of the complex with Mi0621.*

Sub-pockets of the active site of FXIIIa° have not yet been assigned. This relates to the fact that there is no crystal structure of FXIIIa° available in complex with a bound substrate. In the structure of FXIIIa° with Mi0621 a deviating interaction pattern is observed than with ZED1301 except for the residues at the sites P<sub>1</sub>'-P<sub>3</sub>' (*Fig. 4.6*). This suggests that natural substrates interact with FXIIIa° at these P<sub>1</sub>'-P<sub>3</sub>' positions similarly to ZED1301 and Mi0621. To support this assumption, crystallization with complete substrates would be necessary.

The P<sub>1</sub>' side chain, located next to the entrance of the hydrophobic tunnel, involves the carboxamide group of Asn 371 as interaction partner (*Fig. 4.7*). However, the S<sub>1</sub>' site can hardly be assigned as a visible pocket. The shallow S<sub>2</sub>' pocket is located adjacent to the hydrophobic tunnel and is formed by the residues Trp 370 and Ser 368. Finally, the S<sub>3</sub>' pocket consists of the side chain of Phe 339 and the methylene group and C<sub>α</sub> atom of Asn 371. Consequently, the S<sub>3</sub>' pocket possesses a more hydrophobic character. In contrast to the S<sub>2</sub>' pocket, the P<sub>3</sub>' side chain of a putative substrate occupying the S<sub>3</sub>' pocket points in direction of the surface of the protein.





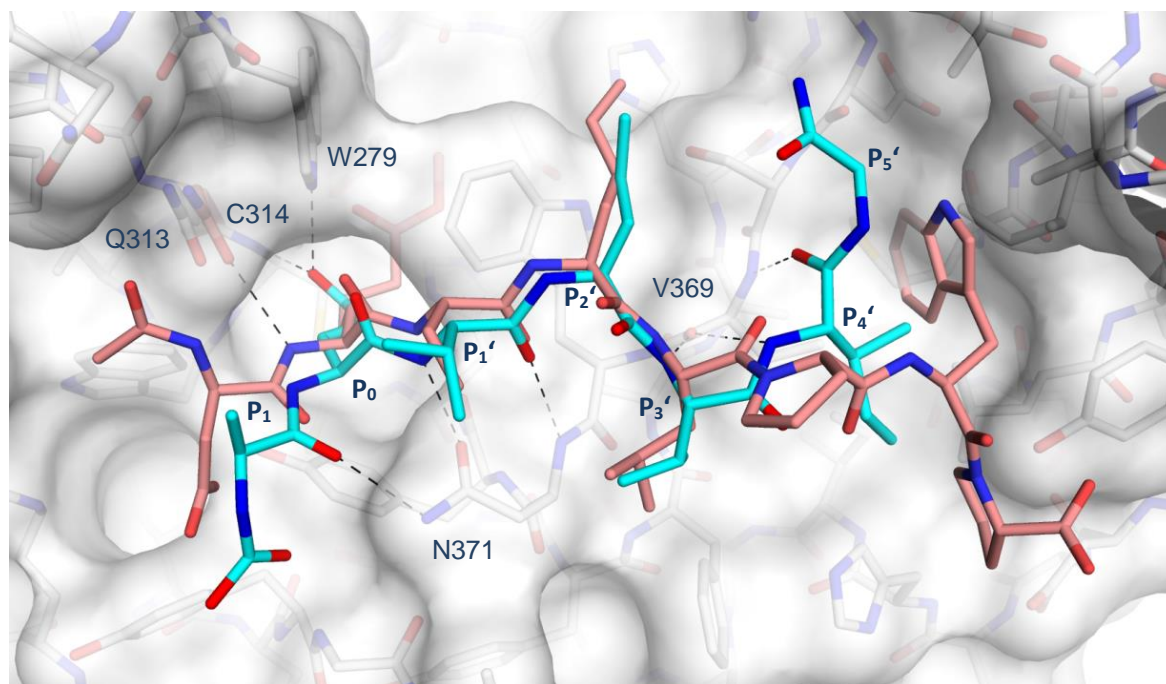
**Fig. 4.7:** Assignment of the  $S_2'$ - $S_3'$  sub-pockets of the active site of factor XIIIa<sup>o</sup> (PDB ID: 5MHL). The  $S_3'$  pocket is occupied by a valine side chain of Mi0621 (cyan) and the  $S_2'$  pocket is addressed by a lysine side chain (not completely defined by electron density – see Fig. 4.5a).

In total, Mi0621 interacts via eight hydrogens bonds with FXIIIa<sup>o</sup> (Fig. 4.8). The hydrogen bonding pattern in  $\beta$ -space adjacent to the catalytic center is identically established in both complexes. Both inhibitors form two H-bonds to the oxygen of the carboxamide side chain and the backbone NH of Asn 371 experiences an interaction with the backbone NH of the warhead residue ( $P_0$ ) and the backbone carbonyl oxygen of the  $P_1'$  amino acid, respectively. Furthermore, both inhibitors establish a hydrogen bond to the backbone oxygen of Val 369 using their peptide NH groups of the  $P_3'$  amino acid.

Compared to ZED1301, Mi0621 forms two additional hydrogen bonds in the region bridging the catalytic center and the hydrophobic region in the  $\beta$ -space. The  $P_4'$  backbone amino group of Mi0621 experiences an H-bond to the carbonyl oxygen of Val 369. The corresponding  $P_4'$  backbone carbonyl oxygen atom of Mi0621 forms a hydrogen bond to the backbone NH of Val 369. This H-bonded arrangement orients the side chain of the  $P_4'$  isoleucine residue exactly towards the hydrophobic region of the  $\beta$ -space.

Additionally, an H-bond is observed between the backbone carbonyl oxygen of the  $P_1$  amino acid and the backbone NH of Asn 371. Compared to ZED1301, there is no interaction of the

peptide bond to the backbone carbonyl oxygen of Gln 313. This might result from the used chloromethyl ketone warhead, which elongates geometry by one methylene group. The C-terminal glycine residue (P<sub>5</sub>' position) of Mi0621 forms no hydrogen bond with the protein. This is in line with the observation that omitting the latter amino acid does not affect affinity of the inhibitor.

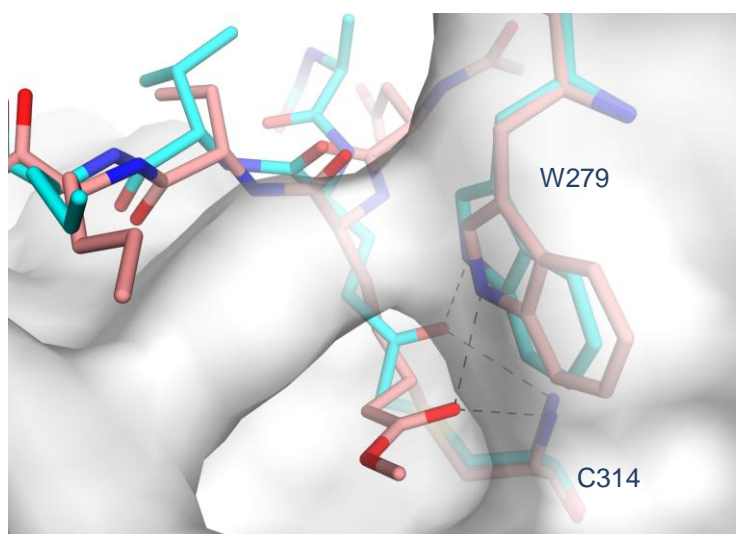


**Fig. 4.8:** Binding mode of Mi0621 (cyan, PDB ID: 5MHL) compared to ZED1301 (light red, PDB ID: 4KTY). In total, Mi0621 forms eight hydrogen bonds to FXIIIa<sup>o</sup>. Importantly, Mi0621 establishes two further hydrogen bonds with its P<sub>4</sub>' amino acid orienting the P<sub>4</sub>' side chain towards the hydrophobic region of the  $\beta$ -space.

This is the first crystal structure of a transglutaminase inhibited by a chloromethyl ketone inhibitor and confirms the assumed binding mode described in the introduction (Fig. 4.4b). Both inhibitors interact via their carbonyl oxygens of the warhead with the indole NH of Trp 279 and the backbone NH of Cys 314 (Fig. 4.9).

Moreover, Fig. 4.8 shows that Mi0621 penetrates not as deep as ZED1301 into the hydrophobic tunnel (distance between the C <sub>$\alpha$</sub>  carbon atoms of both warheads: 1.0 Å). This is mainly attributed to deviating linker length of the two warheads (Fig. 4.4 and Fig. 4.9). The additional

carbon atom introduced into the chloromethyl warhead dislocates Mi0621 out of this area. Considering the fact that the glutamine of the natural substrate is also connected via the C $\delta$  atom forming the thioester intermediate, the extended chloromethyl ketone deviates slightly from the binding geometry of a putative substrate. However, Mi0621 bears a carbonyl group at the same position ( $\delta$ ) as in the natural substrate (glutamine side chain). In case of ZED1301 the carbonyl oxygen of the  $\alpha,\beta$ -unsaturated carboxylic ester is located in the back part of the hydrophobic tunnel. This slightly relocates the indole ring of Trp 279 compared to the Mi0621 complex.

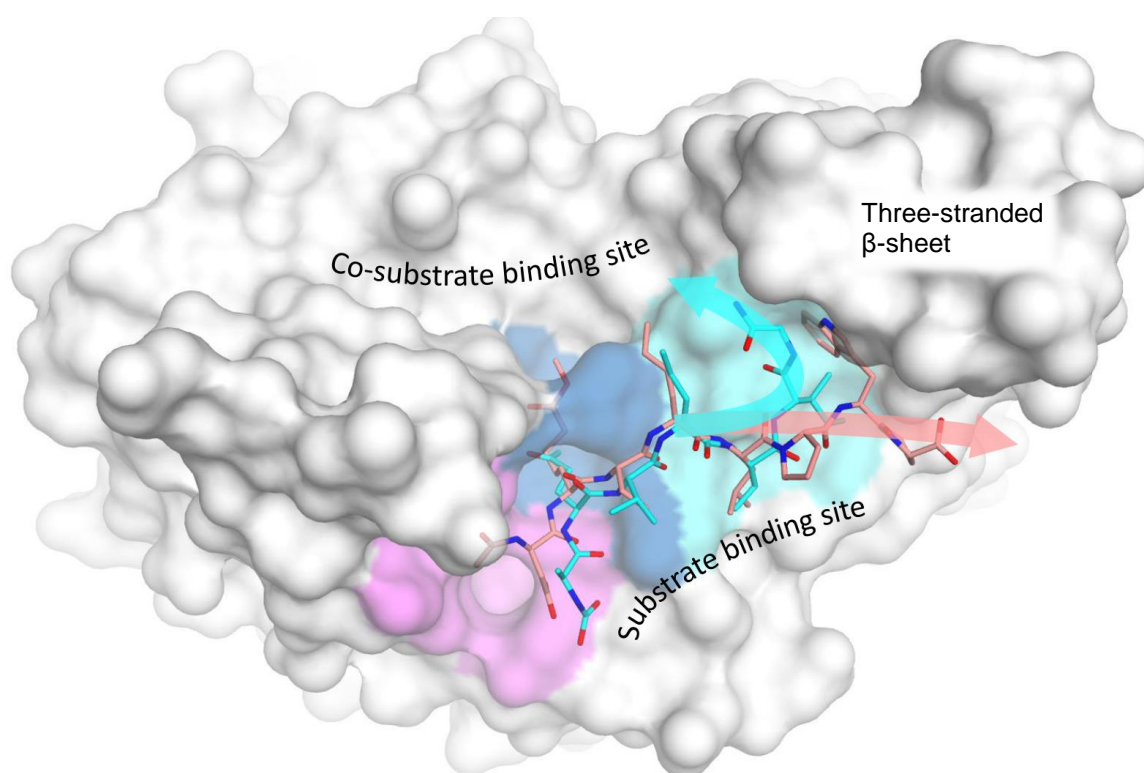


**Fig. 4.9:** Comparison of the warheads of Mi0621 (cyan, PDB ID: 5MHL) and ZED1301 (light red, PDB ID: 4KTY).

Remarkably, adjacent the S $_3$ ' pocket, the inhibitor Mi0621 orients towards the co-substrate binding site whereas ZED1301 is directed in opposite direction, likely determined by the proline residue at P $_4$ ' (Fig. 4.10). Both orientations could suggest putative orientations adopted by natural substrates. Though ZED1301 was derived based on phage screening<sup>[68]</sup>, the sequence QxxxP also occurs as part of natural substrates of FXIII (Chapter 4.2.3).<sup>[53]</sup> However, if a longer substrate would bind analogously to Mi0621, it would clash with the co-substrate. Therefore, it appears more likely that longer peptides orient similarly to ZED1301 in front of the three-stranded  $\beta$ -sheet. Fig. 4.10 also illustrates that the protein architecture of the active site places substrate and co-substrate into close vicinity. Consequently, a sterically demanding protein occupying the substrate binding site could facilitate by chemical interaction or sterically prevent

the occupancy of the co-substrate binding site. This allows the assumption that the specificity of FXIII (and likely also other transglutaminases) concerning the co-substrate could also be affected by the chemical structure of the proteinogenic substrate and co-substrate remote from the active site.

The orientation of the backbone of the inhibitor also differs in the  $\alpha$ -space. Consequently, further sub-pockets cannot be assigned unambiguously. In this respect, crystallization with additional substrate-like ligands extended at the N and C-terminus could provide valuable information.



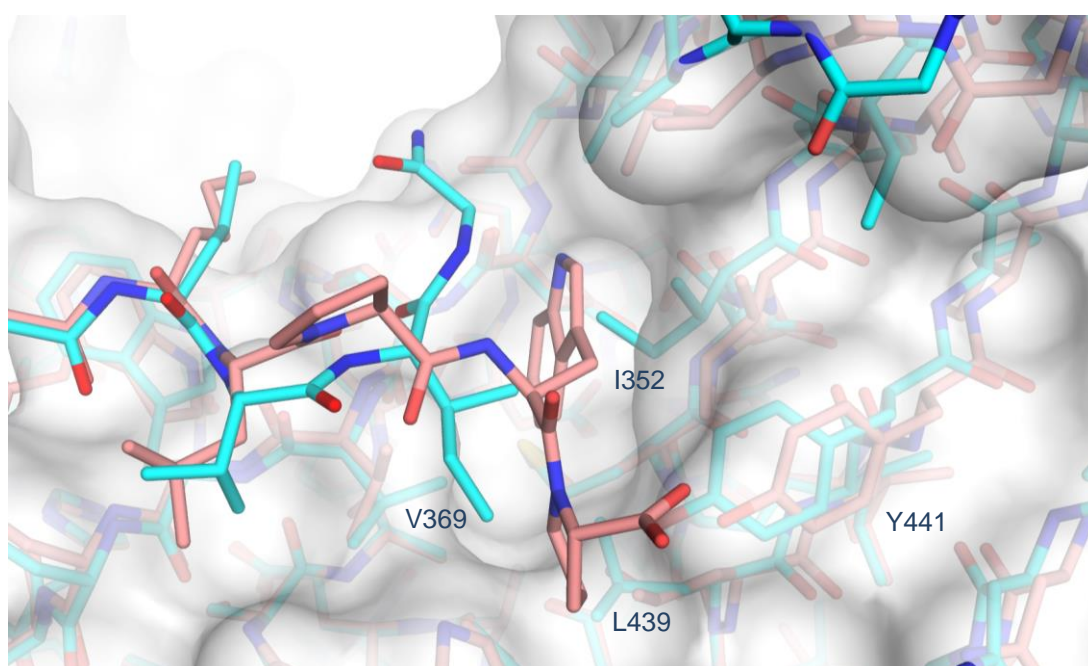
**Fig. 4.10:** Top view of the active site of FXIIIa<sup>o</sup>. The inhibitor Mi0621 (cyan, PDB ID: 5MHL) orients towards the co-substrate binding site whereas ZED1301 (light red, PDB ID: 4KTY) is placed in front of the three-stranded  $\beta$ -sheet.



### 4.2.2 The Three-Stranded $\beta$ -Sheet Adopts Different Conformations

The collapse of the hydrophobic pocket in the  $\beta$ -space is mainly attributed to a conformational transition of two amino acids in the region of the hydrophobic pocket (*Fig. 4.11*). The methyl group of Ile 352 relocates its position and mainly results in the closing of the pocket. Furthermore, a rotation of the isopropyl group of Val 369 also attributes to the adaptive behavior of this hydrophobic area. This means, the terminal tryptophan residue of ZED1301 stabilizes the opening of the hydrophobic pocket in the  $\beta$ -space.

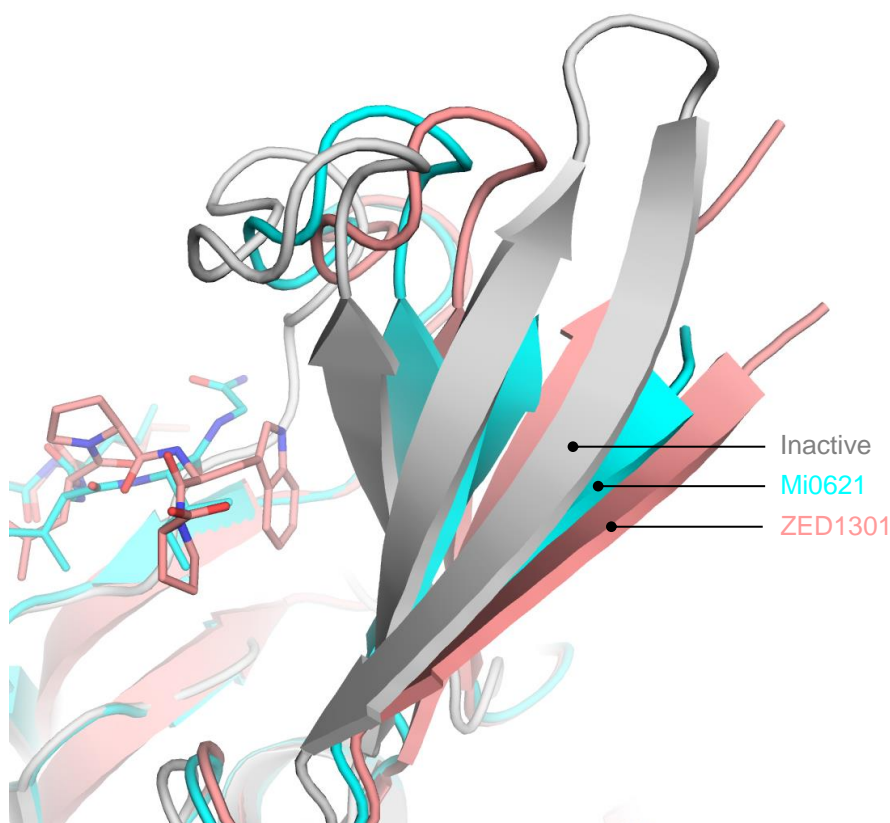
In addition to the conformational change of individual amino acids, a significant shift of the adjacent amino acids to the indole ring of ZED1301 (on average 1.0 Å between the  $C_\alpha$  atoms of Ile 352, Leu 439 and Tyr 441) forming the external part of the active site also occurs.



**Fig. 4.11:** Spatial superposition of the crystal structures of FXIII in complex with Mi0621 (cyan, PDB ID: 5MHL) and ZED1301 (light red, PDB ID: 4KTY) shows that a conformational change of Ile 352 and Val 369 along with a slight shift of Ile 352, Tyr 441 and Leu 439 are responsible for the collapse of the hydrophobic pocket, found in the ZED1301 complex, in the structure with Mi0621.

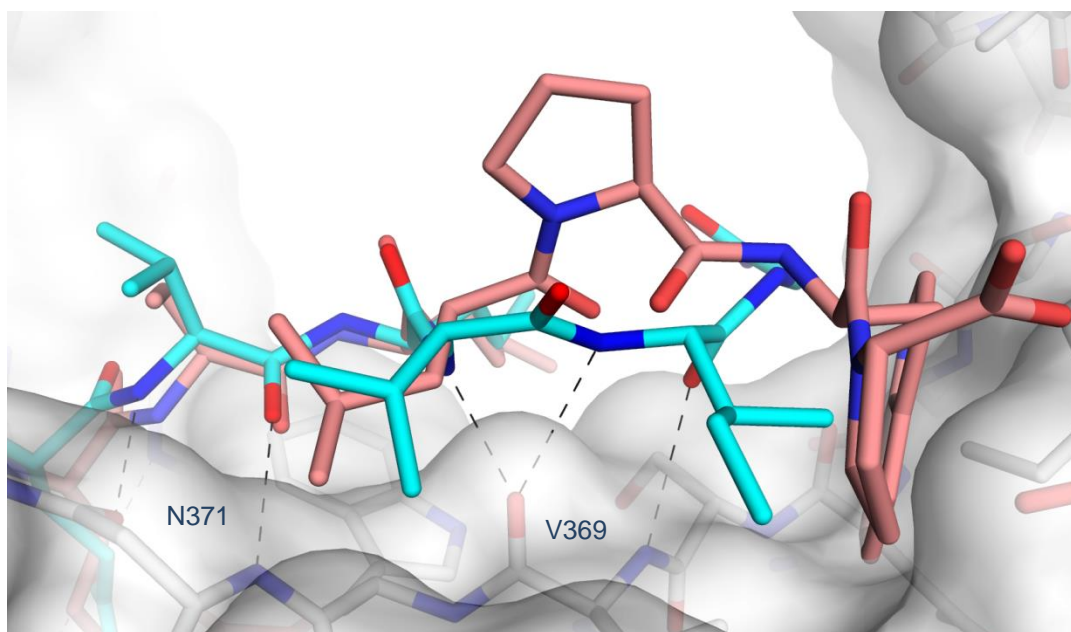
Interestingly, the latter shifted amino acids are part of a three-stranded  $\beta$ -sheet. A superposition of FXIII in the uncomplexed inactive state, in complex with Mi0621 and in complex with

ZED1301 shows that this three-stranded  $\beta$ -sheet adopts in the Mi0621 complex an orientation between the inactive state and the state induced by the inhibitors bearing a proline at P<sub>4</sub>' position (Fig. 4.12).



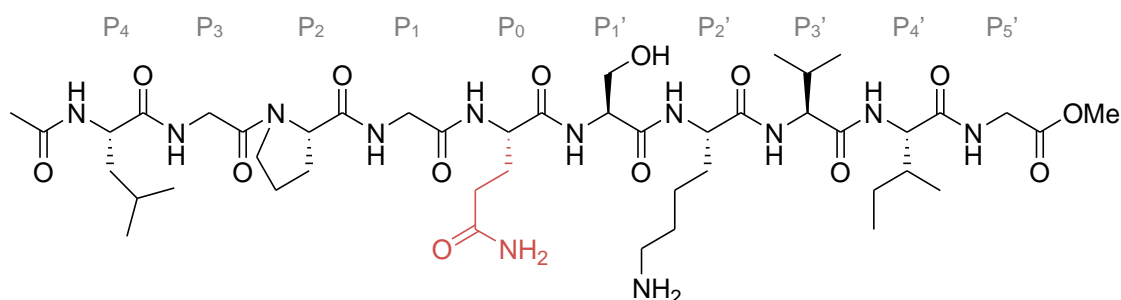
**Fig. 4.12:** Superposition of the crystal structures of FXIII in the inactive state (grey, PDB ID: 1F13), in complex with Mi0621 (cyan, PDB ID: 5MHL) and ZED1301 (light red, PDB ID: 4KTY). The three-stranded  $\beta$ -sheet of FXIIIa° in complex with Mi0621 adopts an orientation between the inactive state and FXIII in complex with ZED1301.

This observation can be explained by inspecting the binding site of FXIIIa° with bound Mi0621. The structural superposition of the crystal structure of FXIIIa° in complex with ZED1301 shows that the P<sub>5</sub>' tryptophan and the C-terminal proline at P<sub>6</sub>' position would clash with the three-stranded  $\beta$ -sheet (Fig. 4.13). Comparing the backbone trace of both inhibitors, ZED1301 adopts a bent conformation whereas Mi0621 binds closer to the protein surface.



**Fig. 4.13:** View onto the binding mode of Mi0621 (cyan, PDB ID: 5MHL) compared to ZED1301 (light red, PDB ID: 4KTY). Mi0621 forms two further hydrogen bonds (H-bonds to Val 369) and binds closer to the protein surface compared to ZED1301.

The crystal structure of FXIII in complex with Mi0621 can also be used to predict the binding mode of the factor XIII substrate K9 used in the Berichrom®-Assay<sup>[105]</sup> (Fig. 4.14). Mi0621 and K9 are both derived from  $\beta$ -casein. Both ligands only differ in the  $\alpha$ -space and at the P<sub>1</sub>' position. In the  $\alpha$ -space K9 possesses a Gly instead of a Phe at P<sub>1</sub> position and has three further amino acids with Leu, Gly and Pro at P<sub>4</sub>, P<sub>3</sub> and P<sub>2</sub> position, respectively. At P<sub>1</sub>' K9 bears Ser instead of Val. Consequently, K9 should interact in similar fashion as Mi0621 in the  $\beta$ -space.



**Fig. 4.14:** The substrate "K9" used in the Berichrom®-Assay should bind in the  $\beta$ -space in the same manner like the inhibitor Mi0621. The reactive glutamine is colored red.

### 4.2.3 Consensus sequence for FXIII substrates

In the past decades, huge efforts were made to determine a consensus sequence for FXIII substrates.<sup>[53, 68, 100]</sup> In 2006 Sugimura et al. performed a phage-display screen with a random peptide library suggesting the consensus sequence Qxx $\phi$ xWP (x and  $\phi$  represent a non-conserved x and hydrophobic amino acid  $\phi$ , respectively). However, this sequence is only matched by about 60 % of the peptides of the phage-display screen (red letters in *Tab. 4.2*). The other peptides do not agree with this sequence.<sup>[68]</sup> The ZED inhibitors with the general sequence QxxxPW were derived based on this screen (highlighted in light blue in *Tab. 4.2*).

As already discussed in *Chapter 4.2.2*, the inhibitor ZED1301 induces a more extensive rotation of the three-stranded  $\beta$ -sheet. This is likely because of a proline at P<sub>4</sub>' position inducing an orientation of the side chain of the amino acid at P<sub>5</sub>' position in direction of the three-stranded  $\beta$ -sheet. This calls upon the question, whether this sequence with a proline at P<sub>4</sub>' position does also occur in natural substrates of FXIII?

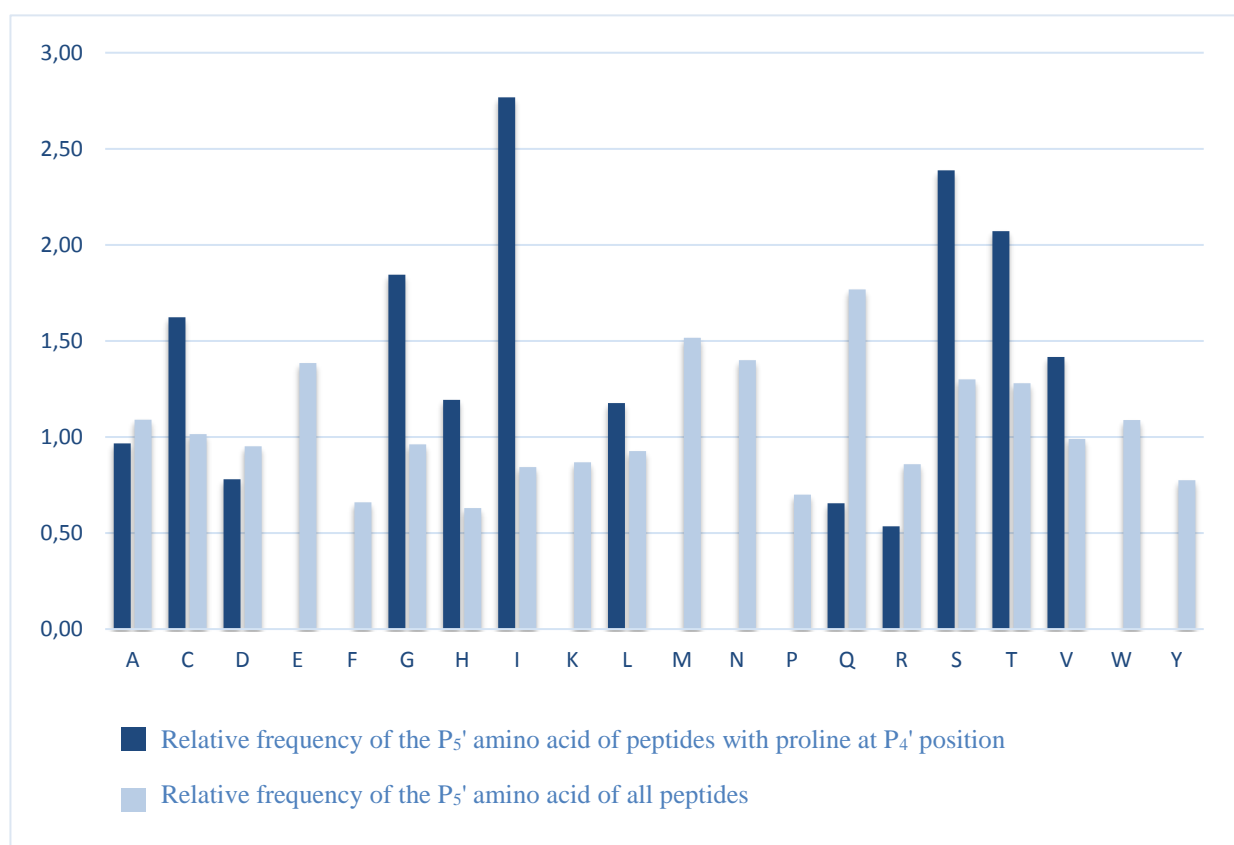
To answer this question, the data provided by Nikolajsen et al. in 2014 were used.<sup>[53]</sup> In this work the blood plasma was incubated with FXIII in presence of the amino co-substrate 5-(biotinamido)pentylamine (PBA). After digestion with trypsin the PBA-containing peptides were isolated via avidin binding and detected by a mass spectrometric approach (LC-MS/MS). An analysis of the provided 690 sequences of peptides with a reactive glutamine shows that 36 of the latter peptides bear proline at P<sub>4</sub>' position (*Tab. 4.3*). Consequently, the trace of the



backbone of ZED1301 induced by proline at P<sub>4</sub>' position is most likely also adopted by natural substrates.

The hydrophobic pocket of the  $\beta$ -space of FXIIIa<sup>o</sup> is occupied by the side chain of tryptophan at P<sub>5</sub>' position of ZED1301 (Fig. 4.6). What are the preferred amino acids of natural substrates at P<sub>5</sub>' of FXIII?

Comparing the 36 peptides of Tab. 4.3 bearing proline at P<sub>4</sub>' position with all the other substrate peptides shows that isoleucine is the most preferred amino acid at P<sub>5</sub>' (Fig. 4.15). Also other hydrophobic amino acids like cysteine, leucine and valine occur more frequently. Interestingly, there is also a preference for glycine, histidine, serine and threonine. However, none of the Pro-P<sub>4</sub>' peptides bear tryptophan at P<sub>5</sub>' position as the inhibitor ZED1301.



**Fig. 4.15:** Frequency of amino acids at P<sub>5</sub>' position of substrates with the consensus sequence QxxxPx (dark blue) compared to all substrates (light blue) based on a LC-MS/MS analysis published by Nikolajsen et al in 2014.<sup>[53]</sup> The relative frequency is shown meaning the absolute frequency number of an amino acid at fifth position to the reactive glutamine divided by the absolute frequency number of an

*amino acid at the fifth position with respect to all glutamine residues (also including the non-reactive ones).*

In total, Nikolajsen et al. analyzed 690 substrate peptides in order to find a consensus sequence of FXIII substrates. Even if the analyzed data show a slight preference for acidic residues at P<sub>1</sub> and P<sub>2</sub> position and no occurrence for proline at P<sub>1</sub>' position, the researchers concluded that there is no clear consensus sequence for FXIII substrates surrounding the active glutamine (P<sub>0</sub>). Doiphode et al. described a decrease in affinity upon substitution of glutamate by arginine at P<sub>1</sub>', glutamine by serine at P<sub>2</sub>' and introduction of glutamine at P<sub>3</sub>' position.<sup>[100]</sup> Furthermore, an increased frequency could be revealed for basic residues at surrounding the P<sub>8</sub> and P<sub>8</sub>' position.<sup>[53, 100]</sup>

Nevertheless, FXIII and other transglutaminases, are known to be specific concerning the cross-linking of their substrates.<sup>[1, 68]</sup> Obviously, there are areas for substrate recognition remote from the active site, as already suggested by Nikolajsen as well as Doiphode and co-workers.<sup>[53, 100]</sup> Regarding the active site of FXIIIa° (Fig. 4.10), substrate and co-substrate approach each other in such a close distance to interact with each other. Conceivably, the specificity of the transamidation can also be affected by the interaction or steric repulsion between the two proteins remotely from the active site and also remotely from FXIII itself. Moreover, the activity of plasmatic FXIII is accelerated by fibrin.<sup>[45-47]</sup> Thus, at least in the blood plasma, an impairment of the specificity by the B-subunit and/or fibrin cannot be excluded.

### 4.3 Summary & Conclusion

Taking into account all the information provided by the crystal structures of inhibitors of the two compound classes it is not surprising that all attempts for finding a consensus sequence of FXIII substrates binding to the region of the active site must have failed.

The  $\alpha$ -space is adaptive due to the flexible arginine. The P<sub>1</sub>' position can hardly contribute to the specificity because the side chain of the substrate at this position is oriented towards the solvent. The shallow pockets S<sub>2</sub>' and S<sub>3</sub>' provide only cavities as minor binding sites. In contrast, the S<sub>3</sub>' pocket should take highest impact on substrate specificity because here the side chain of the corresponding amino acid points towards the surface of FXIIIa° leading to minor tolerance concerning the accommodation of sterically demanding side chains. Interestingly, proline at P<sub>4</sub>' seems to affect the preference at P<sub>5</sub>' especially for isoleucine.

Moreover, the hydrophobic pocket in  $\beta$ -space is highly adaptive due to the ability of the three-stranded  $\beta$ -sheet to adopt different orientations as well as a conformational flexibility of Ile 352 and Val 369. Importantly, the three-stranded  $\beta$ -sheet is suggested as a key structural element in the activation of FXIIIa°. This assumption is corroborated by the interaction of highly conserved residues between the three-stranded  $\beta$ -sheet and calcium binding sites 1 and 2.<sup>[60]</sup> Furthermore, due to the high similarity between all members of the human transglutaminase family, the importance of the three-stranded  $\beta$ -sheet can supposedly be transferred to other transglutaminases as well.<sup>[63]</sup>

Finally, the crystal structures of the two compound classes shows that the trace of the backbone can occur in different orientations: in direction towards the co-substrate binding site or in front of the three-stranded  $\beta$ -sheet.

It is important to mention that we crystallized FXIIIa° in complex with covalently bound inhibitors. Both inhibitors are of peptidic nature and thus correspond to some degree to putative substrates in respect of their amino acid sequence. Considering protein-like substrates they are reduced in size at their N- and the C-terminus.<sup>[102, 106]</sup> In addition, the inhibitors bear a reactive chemical group for covalent linkage replacing the glutamine of the substrates. We therefore believe that our inhibitor structures provide some information with respect to possible binding

modes of natural substrates. However, protein-like substrates may differ in the interaction profile compared to our inhibitors.

## 4.4 Experimental Part

### 4.4.1 Crystallization, Data Collection and Structure Determination

The protein was crystallized utilizing the sitting drop vapor diffusion method at a concentration of 4.0 mg/mL and ambient temperature. 4  $\mu$ L protein solutions were mixed with 5  $\mu$ L precipitation solution (170 mM ammonium sulphate, 85 mM sodium cacodylate pH 6.5, 25.5% (w/v) PEG 8,000, 15% (v/v) glycerol). A suitable single crystal was obtained after removal of small crystallites from its surface. The extracted crystal was flash-frozen in liquid nitrogen and a dataset of 2000 images with 0.1 degree rotation at a wavelength of 0.9778 Å was collected at the ID29 beamline, European Synchrotron Radiation Facility in Grenoble, France.<sup>[107]</sup> Data were processed to a resolution of 2.40 Å in space group P1 using the program XDS.<sup>[98]</sup> The initial phases were calculated by Fourier synthesis using the structure (PDB ID: 4KTY) as starting model in the first refinement step with Phenix. The model was manually improved with Coot<sup>[108]</sup> and refined with Phenix.<sup>[88]</sup> At the beginning, a simulated annealing step was performed. The structure was refined with coordinates, individual B-factors, occupancies and TLS groups. The ligand restraints were created with eLBOW<sup>[109]</sup> implemented in Phenix. According to Procheck<sup>[90]</sup>, 1103 residues (90 %) are found in the most favored, 119 residues (9.8 %) in allowed, and 3 residues (0.2 %) occur in disallowed regions (Asp 270 of molecule 1 and 2 and Asp 139 of molecule 1).

## 4.5 Appendix

### 4.5.1 Crystallographic Table

**Tab. 4.1:** Data collection and refinement statistics.

<b>Inhibitor</b>	Mi0621
<b>PDB Entry</b>	5MHL
<b>Data collection and processing</b>	
Wavelength [Å]	0.97779
Beamline	ESRF – ID29
Space group	P1
<b>Unit cell parameters</b>	
a, b, c [Å]	57.1, 81.0, 103.3
$\alpha$ , $\beta$ , $\gamma$ [°]	88.1, 76.9, 81.5
<b>Diffraction data</b>	
Resolution range [Å]*	48.83 – 2.40 (2.54 – 2.40)
Unique reflections	65090 (9943)
R(I)sym [%]	8.7 (35.2)
Completeness [%]	92.6 (87.7)
Redundancy	2.0 (2.0)
I/ $\sigma$ (I)	7.36 (2.0)
<b>Refinement</b>	
Resolution range [Å]	48.83 – 2.40
Reflections used in refinement (total)	65031
Reflections used in refinement (work/free)	61777/ 3254
Final R values for all reflections (work/free) [%]	17.2/ 23.1

Protein residues (A/B)	692/ 692
Atoms Inhibitor (A/B)	51/ 59
Water molecules	368
Other ligand atoms	54
<b>RMSDs from ideality</b>	
Bond lengths[Å]	0.009
Bond angles [°]	1.4
<b>Ramachandran plot</b>	
Residues in most favored regions [%]	90.0
Residues in additional allowed regions [%]	9.6
Residues in generously allowed regions [%]	0.2
Residues in disallowed regions [%]	0.2
<b>Mean B-factor [Å<sup>2</sup>]</b>	
Protein total (A/B)	39.6 (41.5/ 37.8)
Protein main chain (A/B)	39.0 (41.0/ 36.9)
Protein side chain (A/B)	40.4 (42.0/ 38.8)
Inhibitor (A/B)	40.8 (41.7/ 40.1)
Calcium ions	39.5
Other ligand atoms	47.1
Water molecules	34.4

\*Highest resolution shell is shown in parenthesis

**Tab. 4.2:** Result of a phage-display screen performed by Sugimura et al.<sup>[68]</sup>

Peptide ID	P <sub>4</sub>	P <sub>3</sub>	P <sub>2</sub>	P <sub>1</sub>	P <sub>0</sub>	P <sub>1</sub> '	P <sub>2</sub> '	P <sub>3</sub> '	P <sub>4</sub> '	P <sub>5</sub> '	P <sub>6</sub> '
<i>Consensus sequence</i>					Q	x	x	ϕ	x	W	P
F23			L	D	Q	I	V	I	P	W	P
F11				D	Q	M	M	L	P	W	P
F18				D	Q	M	M	L	P	W	P
F19				D	Q	W	M	M	A	W	P
F9				E	Q	Y	Q	L	A	W	P
F14				E	Q	Y	Q	L	A	W	P
F2				E	Q	Y	Q	L	A	W	P
F16			S	E	Q	H	L	L	K	W	P
F7				E	Q	Q	Q	L	S	W	P
F12				S	Q	H	P	L	P	W	P
F26				S	Q	H	P	L	P	W	P
F6				S	Q	I	P	M	A	W	P
F31				T	Q	Y	T	M	T	W	P
F20				W	Q	Q	Q	M	K	W	P
F13				W	Q	I	P	V	D	W	P
F29				W	Q	M	Q	L	P	W	P
F35				G	Q	M	I	L	P	W	P
F17				E	Q	F	P	I	A	F	P
F21				H	Q	M	V	L	T	Y	P
F32				D	Q	Y	V	L	T	F	P
F4				W	Q	M	P	V	N	L	P



F15				W	Q	L	K	Y	P	W	I
F5				H	Q	I	P	I	Q	I	H
F28				W	Q	H	K	I	D	L	R
F27				W	Q	H	K	I	D	L	R
F22				D	Q	S	K	V	W	L	L
F30		W	H	Y	Q	V	P	Q	W	L	Y
F25	Y	H	P	L	Q	L	T	P	A	P	I

**Tab. 4.3:** Peptides of FXIII substrates with proline at P<sub>4</sub>' position based on a LC-MS/MS analysis published by Nikolajsen et al in 2014.<sup>[53]</sup>

Substrate	P <sub>4</sub>	P <sub>3</sub>	P <sub>2</sub>	P <sub>1</sub>	P <sub>0</sub>	P <sub>1</sub> '	P <sub>2</sub> '	P <sub>3</sub> '	P <sub>4</sub> '	P <sub>5</sub> '	P <sub>6</sub> '
Consensus sequence for Pro-P <sub>4</sub> ' inhibitors					Q	x	x	x	P	x	x
Alpha-2-antiplasmin			K	E	Q	Q	D	S	P	G	N
Apolipoprotein A-I				R	Q	G	L	L	P	V	L
Extracellular matrix protein 1	P	P	L	P	Q	E	A	V	P	L	Q
Inter-alpha-trypsin inhibitor heavy chain H2	K	F	Y	N	Q	V	S	T	P	L	L
Heparin cofactor 2	G	E	T	A	Q	S	A	D	P	Q	W
Inter-alpha-trypsin inhibitor heavy chain H1	T	V	S	Q	Q	Q	S	C	P	T	C
Plasma protease C1 inhibitor	E	D	M	E	Q	A	L	S	P	S	V
Inter-alpha-trypsin inhibitor heavy chain H2	K	F	Y	N	Q	V	S	T	P	L	L

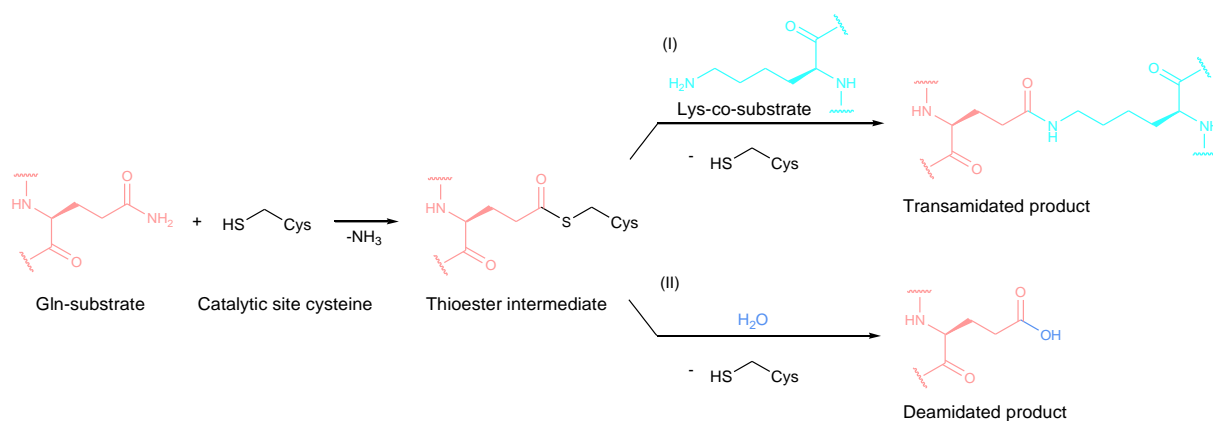
Inter-alpha-trypsin inhibitor heavy chain H1	T	V	S	Q	<b>Q</b>	Q	S	C	<b>P</b>	T	C
Serum albumin	K	K	V	P	<b>Q</b>	V	S	T	<b>P</b>	T	L
Complement C4-A		K	L	G	<b>Q</b>	Y	A	S	<b>P</b>	T	A
Hemopexin		K	L	L	<b>Q</b>	D	E	F	<b>P</b>	G	I
von Willebrand factor				K	<b>Q</b>	H	S	D	<b>P</b>	C	A
Carboxypeptidase N subunit 2				K	<b>Q</b>	L	V	C	<b>P</b>	V	T
Complement C2	K	S	S	G	<b>Q</b>	W	Q	T	<b>P</b>	G	A
Vitamin K-dependent protein C	A	T	L	S	<b>Q</b>	T	I	V	<b>P</b>	I	C
Fibrinogen gamma chain	Q	L	E	A	<b>Q</b>	C	Q	E	<b>P</b>	C	K
IgGFC-binding protein	C	K	P	G	<b>Q</b>	V	C	Q	<b>P</b>	S	G
Ig mu chain C region	A	A	T	S	<b>Q</b>	V	L	L	<b>P</b>	S	K
Complement factor B		K	Y	G	<b>Q</b>	T	I	R	<b>P</b>	I	C
Complement C3	R	A	E	L	<b>Q</b>	C	P	Q	<b>P</b>	A	A
Coagulation factor XIII A chain	T	V	E	L	<b>Q</b>	G	V	V	<b>P</b>	R	G
Carboxypeptidase N subunit 2	H	L	G	F	<b>Q</b>	V	T	W	<b>P</b>	D	E
Complement C1r subcomponent	R	F	C	G	<b>Q</b>	L	G	S	<b>P</b>	L	G
Protein AMBP	V	P	G	E	<b>Q</b>	E	P	E	<b>P</b>	I	L
Plasminogen	C	D	V	P	<b>Q</b>	C	A	A	<b>P</b>	S	F
Plasma protease C1 inhibitor	E	D	M	E	<b>Q</b>	A	L	S	<b>P</b>	S	V
IgGFC-binding protein		R	L	E	<b>Q</b>	Y	E	G	<b>P</b>	G	F

Ig alpha-1 chain C region	T	T	S	S	<b>Q</b>	L	T	L	<b>P</b>	A	T
Inter-alpha-trypsin inhibitor heavy chain H2		R	N	V	<b>Q</b>	F	N	Y	<b>P</b>	H	T
Apolipoprotein A-I				R	<b>Q</b>	G	L	L	<b>P</b>	V	L
Inter-alpha-trypsin inhibitor heavy chain H4				R	<b>Q</b>	L	G	L	<b>P</b>	G	P
Complement component C7	A	F	E	T	<b>Q</b>	S	C	E	<b>P</b>	T	R
Properdin	H	P	V	P	<b>Q</b>	H	G	G	<b>P</b>	F	C
Alpha-1B-glycoprotein	G	E	S	S	<b>Q</b>	V	L	H	<b>P</b>	G	N
Prolow-density lipoprotein receptor-related protein 1	P	F	D	L	<b>Q</b>	V	Y	H	<b>P</b>	S	R

## 5 Does TG2 Adopt Two Different Active States?

### 5.1 Introduction

Transglutaminases (TGs) represent a major protein class in humans involved in a variety of essential biochemical processes such as inflammation, programmed cell death, signal transduction and blood coagulation.<sup>[1, 63]</sup> These calcium dependent enzymes predominantly catalyze the cross-linking of two proteins by the formation of isopeptide bonds, a process known as transamidation. Additionally, transglutaminases catalyze the deamidation of the reactive glutamine into a glutamate residue. As shown in *Fig. 5.1*, transamidation occurs if the thioester intermediate reacts with the lysine co-substrate (I) whereas water penetrating the catalytic site can hydrolyze the thioester intermediate to a glutamate residue (II).



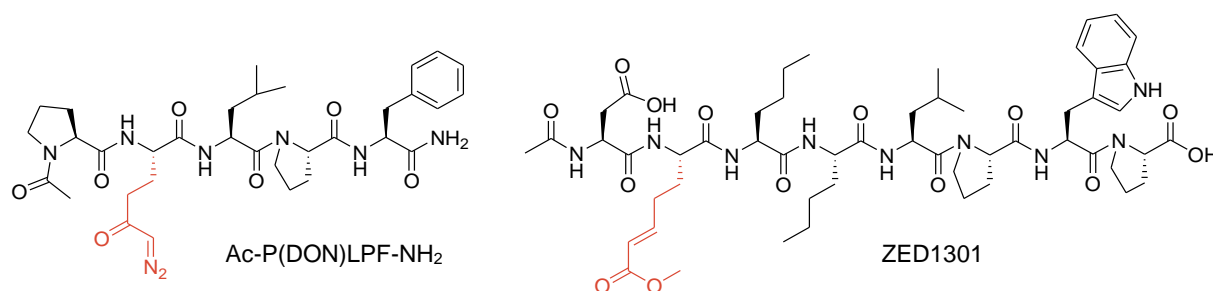
**Fig. 5.1:** Transglutaminases catalyze predominantly the transamidation of a Gln-substrate (red) with a lysine co-substrate (cyan) resulting in a isopeptide bond (I). However, transglutaminases are also capable to catalyze the deamidation of the Gln-substrate (II). Herein, the thioester intermediate is hydrolyzed by water (blue) entering the catalytic site.

Especially TG2 is regarded as a promising pharmacological target for many human diseases like celiac disease, Alzheimer's disease and cancer.<sup>[21, 24, 26, 32]</sup> Due to the multifunctional features of these enzymes, especially of TG2 (transamidation, deamidation, protein disulfide isomerase, kinase activity, ability to act as a G-protein), it is important to understand the mechanism of transglutaminases on molecular level. This allows the development of specific inhibitors influencing these particular functions to investigate the role of the individual functions of TG2 in the organism and, ultimately, to develop drugs with fewer side effects.<sup>[110]</sup> Transglutaminases has been crystallized in different states over the past decades. The first crystal structure of a transglutaminase (blood coagulation factor XIII) was published in 1994<sup>[17]</sup> and for the first time, scientists gained insights into the molecular architecture of this enzyme class. However, in this crystal structure the active site is sealed by the  $\beta$ -barrel 1 domain. Chemical modification and hydrogen-deuterium exchange experiments suggested a conformational change during the activation process.<sup>[70, 71]</sup>

Ahvazi et al. crystallized transglutaminase 3 in presence of calcium ions in 2002<sup>[18, 75]</sup>, where the protein was cleaved enzymatically between the catalytic domain and the  $\beta$ -barrel 1 domain. Here, TG3 adopts an active conformation with three fully occupied calcium binding sites. Nevertheless, the  $\beta$ -barrel 1 domain remains at the same position as in the inactive state most likely due to the absent of a bound inhibitor.

Five years later, the crystal structure of transglutaminase 2 was published by the lab of Chaitan Khosla featuring a covalent inhibitor (*Fig. 5.2*) attached to the cysteine residue in the active site<sup>[19]</sup>, in the following named TG2a\*.

Finally, in 2013 we were able to crystallize for the first time a transglutaminase (FXIIIa<sup>o</sup>) in complex with an inhibitor and with the three fully established calcium binding sites.<sup>[60]</sup> Interestingly, in contrast to the crystal structure of TG2a\* the enzyme adopts a globular geometry.



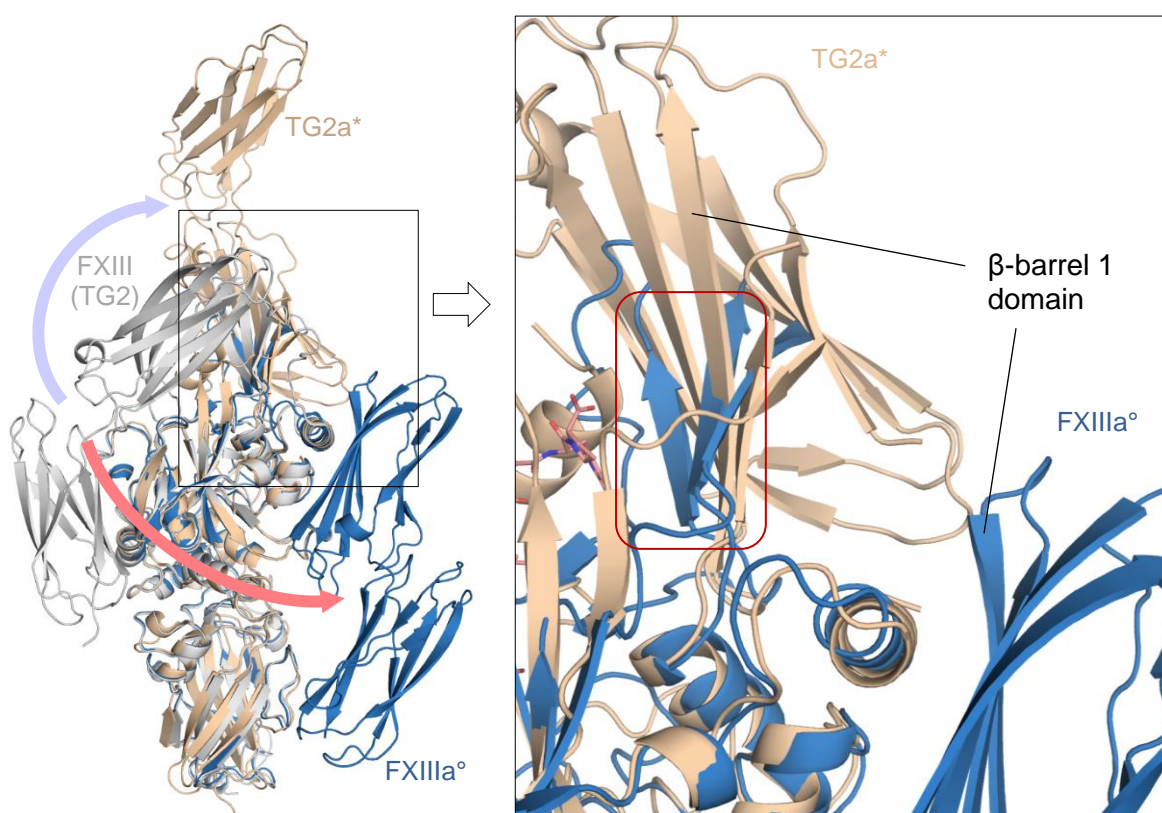
**Fig. 5.2:** Chemical structure of the TG2a\*-blocker Ac-P(DON)LPF-NH<sub>2</sub> (left) with the 6-diazo-5-oxonorleucine (DON) warhead (red)<sup>[19]</sup> and the FXIIIa°-blocker ZED1301 (right) with the  $\alpha,\beta$ -unsaturated methyl ester warhead (red).

In this chapter, the crystal structure of FXIIIa° is systematically compared with the crystal structure of TG2a\*, TG3a° and a homology model of TG2a constructed on the basis of the crystal structure of FXIIIa° to understand why the domains in the protein-ligand complexes of TG2 and FXIII are arranged in a different way. The proteins will be compared concerning their overall structures, as well as their calcium and substrate binding sites. Overall, human transglutaminases are moderately conserved with respect to their primary structure (sequence identity amounts about 30 %) but they are highly conserved with respect to their secondary and tertiary structures.<sup>[17-19, 60, 63]</sup> Consequently, we therefore hypothesize that structural features discovered in one of the crystal structure of an iso-enzyme can be transferred to the other members of this enzyme class, at least with respect to the highly conserved catalytic site and the calcium binding sites.

## 5.2 Comparison of FXIIIa° and TG2a\*

### 5.2.1 Overall Structure

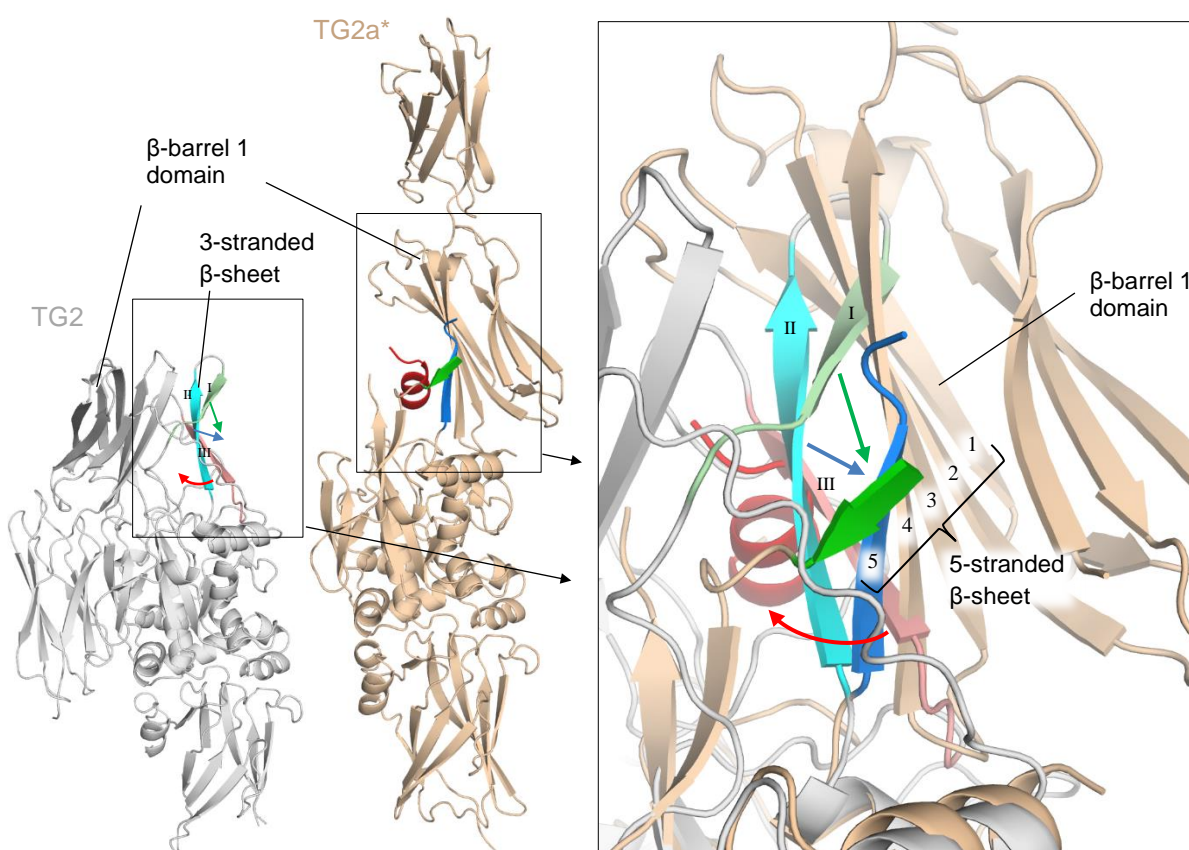
A superposition of FXIIIa° in the most likely active conformation (three calcium sites fully populated; inhibitor: ZED1301) and TG2a\* in the assumed active conformation (none of the calcium sites populated; inhibitor: Ac-P(DON)LPF-NH<sub>2</sub>) shows significant differences in the overall domain orientation (*Fig. 5.3a*). As illustrated in *Chapter 2.3.1*, the two  $\beta$ -barrel domains of FXIII swing aside (red arrow), whereas the four domains of transglutaminase 2 adopt a linear arrangement (purple arrow).<sup>[19, 60]</sup> Notably, the  $\beta$ -barrel 1 domain of TG2a\* is partially located at the position where the three-stranded  $\beta$ -sheet of FXIIIa° is located (red box in *Fig. 5.3b*). Consequently, in FXIIIa° the domains cannot be arranged in a linear geometry.



**Fig. 5.3:** Structural superposition of FXIII, FXIIIa° and TG2a\*. TG2a\* (beige, PDB ID: 2Q3Z) adopts in contrast to FXIIIa° (blue, PDB ID: 4KTY) a linear array of its domains. The inactive state of FXIII (PDB ID: 1F13) is colored in gray (only shown one monomer of the dimer). The movement of the two

$\beta$ -barrel domains of FXIII and TG2 are indicated by a red and a purple arrow, respectively. A similar linear arrangement of FXIIIa<sup>o</sup> cannot be adopted as the three-stranded  $\beta$ -sheet in the upper part of the catalytic domain prevents this geometry (red box).

Obviously, the three-stranded  $\beta$ -sheet experiences a conformational change during the transition from TG2 to TG2a\*. This becomes evident by superimposing the structures in both states. Strand I and II (green, blue) are translocated during the transition process. In contrast, strand III (red) reorganizes and forms an  $\alpha$ -helix (red arrow in Fig. 5.4).<sup>[111]</sup> In the active state, the position of the reorganized strand III is taken by strand 4 of the  $\beta$ -barrel 1 domain, forming a five-stranded  $\beta$ -sheet, whereas four strands (1-4) originate from the  $\beta$ -barrel 1 domain and strand 5 has its origin in the three-stranded  $\beta$ -sheet (strand II).



**Fig. 5.4:** Structural rearrangement of the three-stranded  $\beta$ -sheet of TG2 during transition from the inactive (grey, PDB ID: 1KV3)<sup>[112]</sup> to the active state (beige, PDB ID: 2Q3Z).<sup>[19]</sup> Strand I (green) and II (blue) of the inactive state are only shifted in space, whereas strand III and an adjacent loop region (red) adopts an  $\alpha$ -helical structure in the active state. Exactly at the previous position of strand III of



*the inactive state, strand 4 of the  $\beta$ -barrel 1 domain in the active state is located and forms a five-stranded  $\beta$ -sheet, whereas strand 1-4 originates from the  $\beta$ -barrel 1 domain and strand 5 was originally strand II of the three-stranded  $\beta$ -sheet.*

## 5.2.2 Calcium Binding Sites

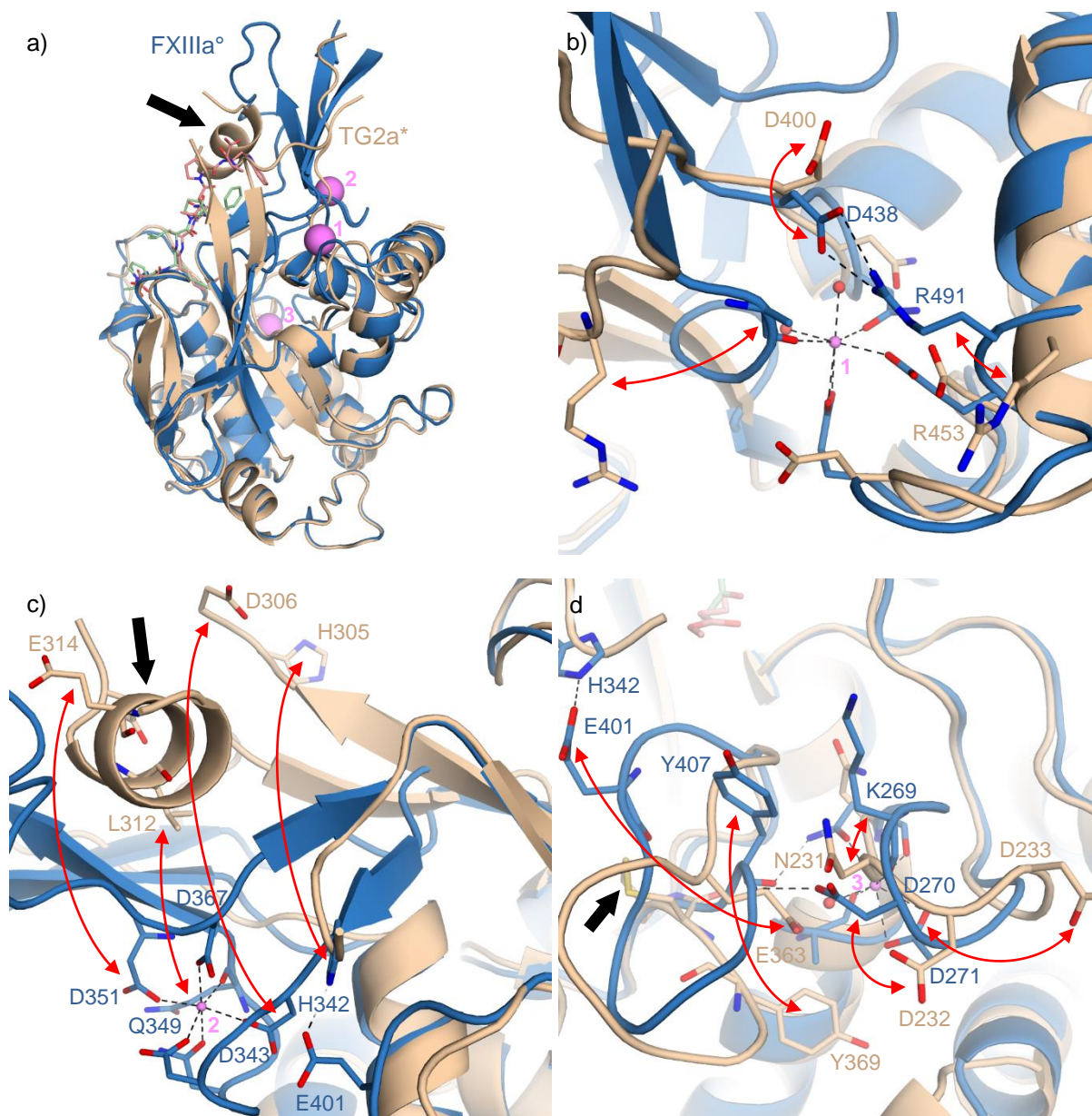
Even though, TG2a\* adopts an active conformation with respect to the arrangement of the  $\beta$ -barrel domains, the calcium binding sites are not established.<sup>[19]</sup>

As shown in *Fig. 5.5b*, in the crystal structure of TG2a\* the side chains of Arg 491 and Asp 438 (FXIII nomenclature) do not interact owing to the missing calcium ion at site 1 (*Chapter 2.3.2*). Notably, both residues are highly conserved among human transglutaminases (*Chapter 2.6.1*).

The amino acids which are according to the crystal structure determined for FXIIIa° in the assumed active state involved in calcium binding site 2, are found with large distance from the location of calcium binding site 2 (*Fig. 5.5c*). Importantly, two amino acids (Glu 314 and Leu 312) of calcium binding site 2 are even part of the  $\alpha$ -helix. Consequently, calcium coordination at site 2 is not compatible with the presence of the  $\alpha$ -helix and would even prevent the reorganization of strand III to the  $\alpha$ -helix. Remarkably, in TG2a\* the  $\alpha$ -helix contributes to the active site and if the binding pose of ZED1301 is hypothetically transferred to TG2a\* the tryptophan moiety of ZED1301 in FXIIIa° would clash with the  $\alpha$ -helix. Most likely, the lack of the established calcium binding site 2 prevents the formation of the active site in the crystal structure of TG2a\* as observed in the crystal structure of FXIIIa°. Moreover, the coordination of calcium ions at site 2 triggers the formation of the catalytic dyad (*Chapter 2.3.2*)<sup>[60]</sup>, which, confirmed by mutational studies, is essential for the enzyme function to catalyze the transamidation reaction.<sup>[76]</sup>

Finally, calcium binding site 3, also involved in the formation of the co-substrate binding site in the crystal structure of FXIIIa°, is not established in TG2a\*. In particular, the TG2 residues of Tyr 407 and Lys 269 (FXIII nomenclature) do not adopt the corresponding positions, as observed in the assumed active state of FXIII (*Fig. 5.5d*). Furthermore, TG2a\* possesses a disulfide bond (black arrow in *Fig. 5.5d*) that is formed under oxidative conditions and has been discussed to inactivate TG2.<sup>[15, 16]</sup> Notably, this disulfide bond is part of the loop, which

reorganizes during the activation process of FXIIIa° and is responsible for shaping of the co-substrate binding site and the formation of the catalytic dyad.



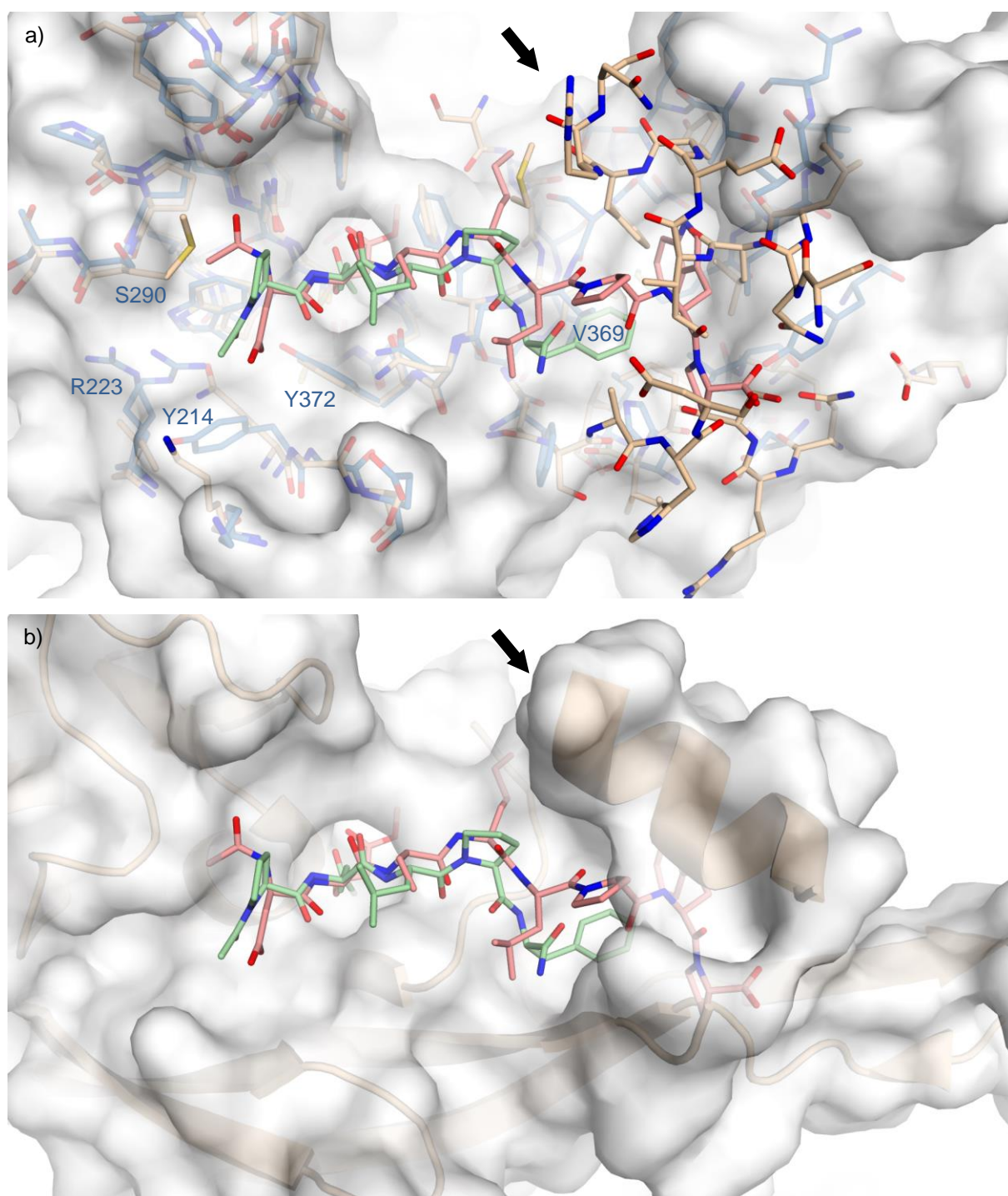
**Fig. 5.5:** In contrast to the most likely active state of FXIIIa° (blue, PDB ID: 4KTY), in the superimposed structure of TG2a\* (beige, PDB ID: 2Q3Z) the three calcium (magenta spheres) binding sites are not established (red arrows indicate significant differences between both structures). a) Overall representation with the inhibitor ZED1301 of FXIIIa° (light red) and the inhibitor Ac-P(DON)LPF-NH<sub>2</sub> of TG2a\* (green). b) In the crystal structure of FXIIIa°, the residues Arg 491 and Asp 438 forms a strong salt bridge. Due to the lack of calcium at binding site 1 the corresponding residues of TG2a\*

*Arg 453 and Asp 400 are found remote from each other in very different locations. c) Two amino acids of calcium binding site 2 (Glu 314 and Leu 312) are involved in the formation of the  $\alpha$ -helix (black arrow) which had reorganized from strand III of the three-stranded  $\beta$ -sheet. Due to lacking calcium coordination at site 2 in the crystal structure of TG2a\*, the catalytic dyad (His 305 and Glu 363) is not formed. d) The missing calcium ion at site 3 results in a different positioning of the co-substrate containing amino acids like Asn 231 (Lys 269 of FXIII) and Tyr 369 (Tyr 407 of FXIII). The formation of a disulfide bond (black arrow) in TG2 restricts the conformational flexibility of the corresponding loop. Notably, this loop region belongs to the co-substrate binding site.*

### 5.2.3 Active Site

How do the active sites differ between FXIIIa<sup>o</sup> and TG2a\*? As expected, due to the fact that the activation process does not affect the  $\alpha$ -space, this region is structurally similar concerning the backbone arrangement structure. However, four amino acids disagree in the chemical composition of their side chains (*Fig. 5.6a*). Notably, instead of the flexible arginine residue (Arg 223) TG2 exhibits an isoleucine at this position.

The  $\beta$ -space of the crystal structure of TG2a\* in complex with the inhibitor Ac-P(DON)LPF-NH<sub>2</sub> (*Fig. 5.2*) is reduced in its size compared to FXIIIa<sup>o</sup> in complex with ZED1301. Remarkably, this is attributed to the reorganized  $\alpha$ -helix that prevents access to the rear part of the  $\beta$ -space as established in the FXIIIa<sup>o</sup> structure (*Fig. 5.6b*). The superposition of both crystal structures (TG2a\* and FXIIIa<sup>o</sup>) shows that the inhibitor ZED1301 would clash with its C-terminal part (Pro-Trp-Pro) into the  $\alpha$ -helix (black arrow in *Fig. 5.6b*) present in TG2a\*.

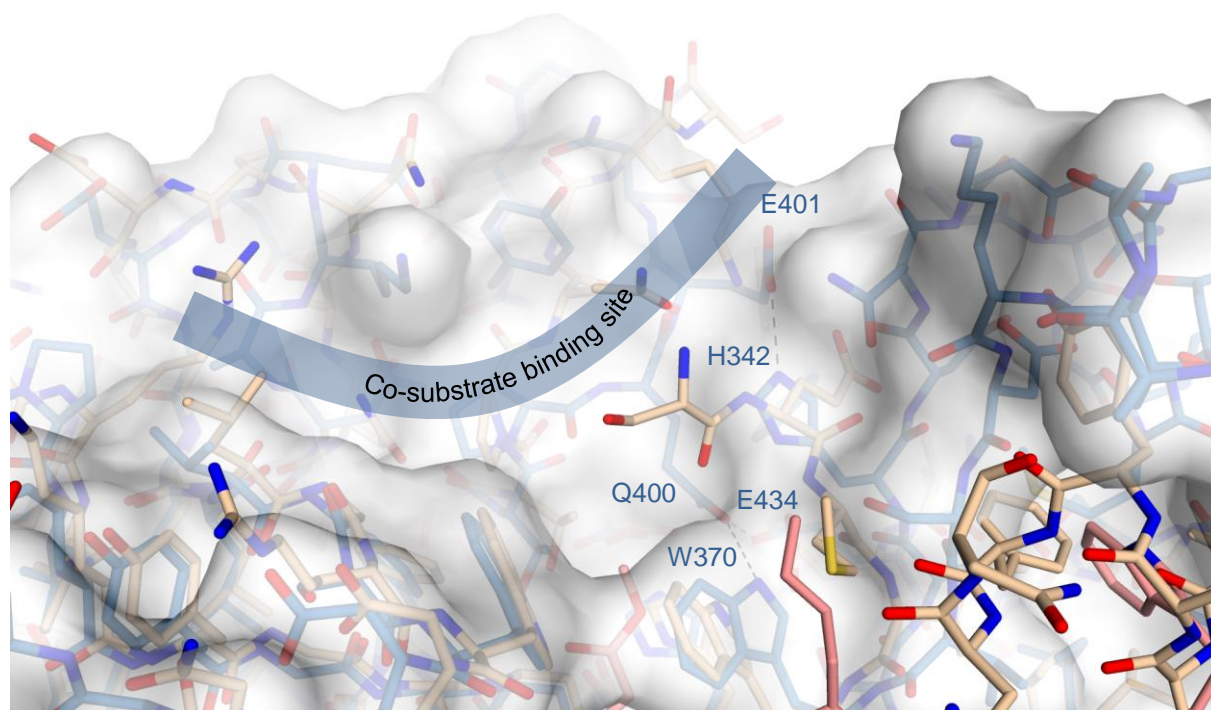


**Fig. 5.6:** Substrate binding sites of the superimposed crystal structures of FXIIIa° (blue, PDB ID: 4KTY) and TG2a\* (beige, PDB ID: 2Q3Z). a) Surface is shown for FXIIIa°. In the  $\alpha$ -space four amino acids differ in their side chain composition compared to FXIIIa° (Ser 290  $\rightarrow$  Met, Arg 223  $\rightarrow$  Ile, Tyr 214  $\rightarrow$  Gln, Tyr 372  $\rightarrow$  Phe). The shorter inhibitor Ac-P(DON)LPF-NH<sub>2</sub> (green) of TG2a\* compared to the inhibitor ZED1301 of FXIIIa° (light red) would clash via its phenyl ring with Val 369 of FXIIIa° if



transferred using the adopted conformation in the TG2a\* complex. However, the clash can be avoided by minor conformational rearrangements of Val 369 or the phenyl ring of the inhibitor. b) Instead, the extended FXIII inhibitor ZED1301 would clash using the observed binding pose in FXIIIa° with the  $\alpha$ -helix in TG2a\* indicated by a black arrow (shown: surface and cartoon representation of TG2a\*). With other words, the  $\alpha$ -helix diminishes the size of the substrate binding site, particularly of the  $\beta$ -space of TG2a\* and does not allow the binding of the bulkier FXIII-inhibitors of the  $\alpha(wh)xxxPW$ -type to TG2 in the linear state (TG2a\*).

Based on the knowledge of the crystal structure of FXIIIa°, the co-substrate binding site as well as the catalytic dyad (His 342, Glu 401; FXIII nomenclature) cannot be established in TG2a\* because of the missing calcium coordination. A superposition of both crystal structures (FXIIIa° and TG2a\*) provides strong evidence for this assumption (Fig. 5.7). Also the H-bond described in Chapter 3.2.4 between the indole ring of Trp 370 and the carboxamide group of Gln 400 (FXIII nomenclature) is not formed in the crystal structure of TG2a\*.



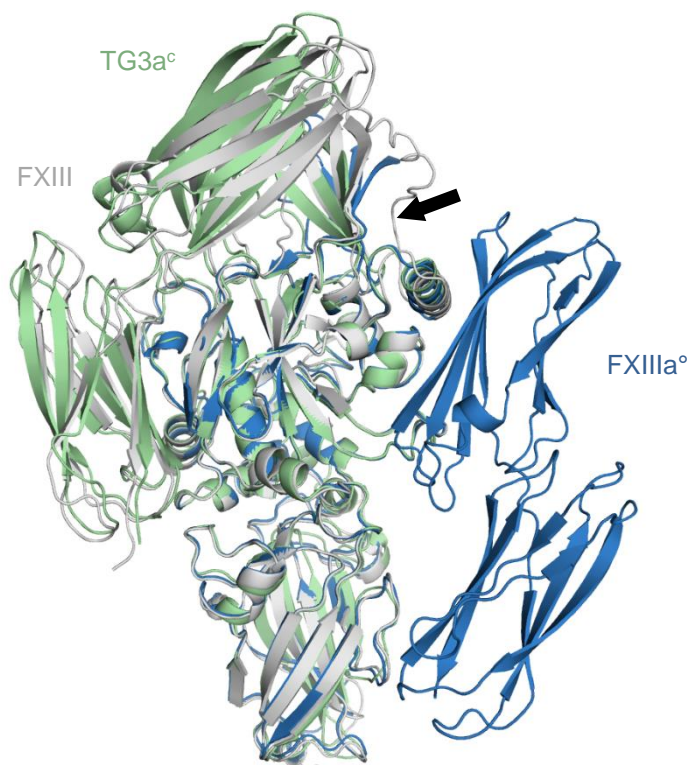
**Fig. 5.7:** A superposition of the crystal structures of FXIIIa° (blue, PDB ID: 4KTY) and TG2a\* (beige, PDB ID: 2Q3Z) shows that a co-substrate binding site is not established in TG2a\* (no structural agreement with FXIIIa° can be found). Also the catalytic dyad that is necessary for the deamidation reaction (His 342 and Glu 401, FXIII nomenclature) and the two amino acids Gln 400 and Glu 434

*(FXIII nomenclature) of which Gln 400 accepts a hydrogen bond of the indole moiety of Trp 370 are not formed.*

## 5.3 Comparison of FXIIIa° and TG3a<sup>c</sup>

### 5.3.1 Overall Structure

Transglutaminase 3 was crystallized in the presence of calcium ions, whereas the protein was enzymatically truncated between the catalytic domain and the  $\beta$ -barrel 1 domain (in this chapter named as TG3a<sup>c</sup>; c: cleaved).<sup>[18, 75]</sup> Interestingly, as shown in *Fig. 5.8*, the enzyme adopts the same domain arrangement as FXIII in the inactive state. This is likely due to the fact that TG3 was crystallized in the absence of an active site ligand, hence, the domain rearrangement and substrate binding seems to be most likely, a concerted process.

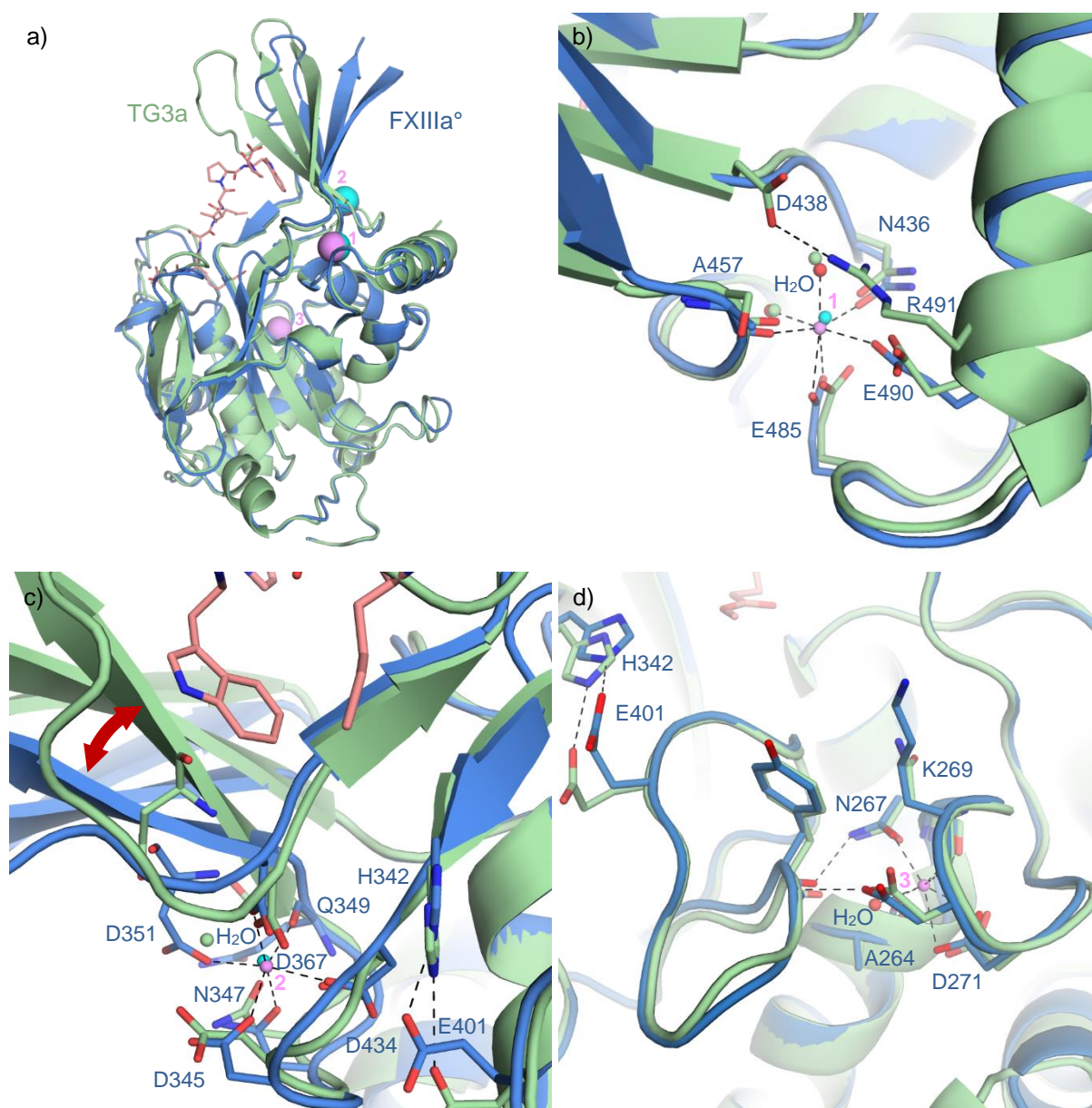


**Fig. 5.8:** Superposition of the inactive state of FXIII (grey, PDB ID: 1GGU), the active state of FXIII (blue, PDB ID: 4KTY) and the crystal structure of TG3a<sup>c</sup> (green, PDB ID: 1L9N) enzymatically truncated between the catalytic domain and the  $\beta$ -barrel 1 domain (black arrow). Herein, TG3a<sup>c</sup> adopts the same domain arrangement as FXIII in the inactive state.

### 5.3.2 Calcium Binding Sites

In the crystal structure of TG3a<sup>c</sup>, the three calcium binding sites are populated and the catalytic dyad is established (*Fig. 5.9*).<sup>[17, 75]</sup> With respect to the coordination spheres of the calcium ions the active form of FXIII (FXIIIa<sup>o</sup>) and the structure of TG3a<sup>c</sup> show remarkable similarity. Nevertheless, there is an important difference at site 2. In the case of FXIIIa<sup>o</sup>, the side chain of Asp 351 interacts with the calcium ion at this site, whereas in the TG3a<sup>c</sup> structure a water molecule occupies this coordination site. As a consequence, in TG3a<sup>c</sup> the adjacent three-stranded  $\beta$ -sheet is only shifted and not rotated. In FXIIIa<sup>o</sup>, complete rotation is essential to fully establish the substrate binding pocket, especially the hydrophobic pocket of the  $\beta$ -space. Therefore, it can be assumed that the TG3a<sup>c</sup> structure does not represent the fully active state. Remarkably, an Asp residue at this pivotal position is highly conserved across all transglutaminases except TG2, which contains a Glu residue at this position.



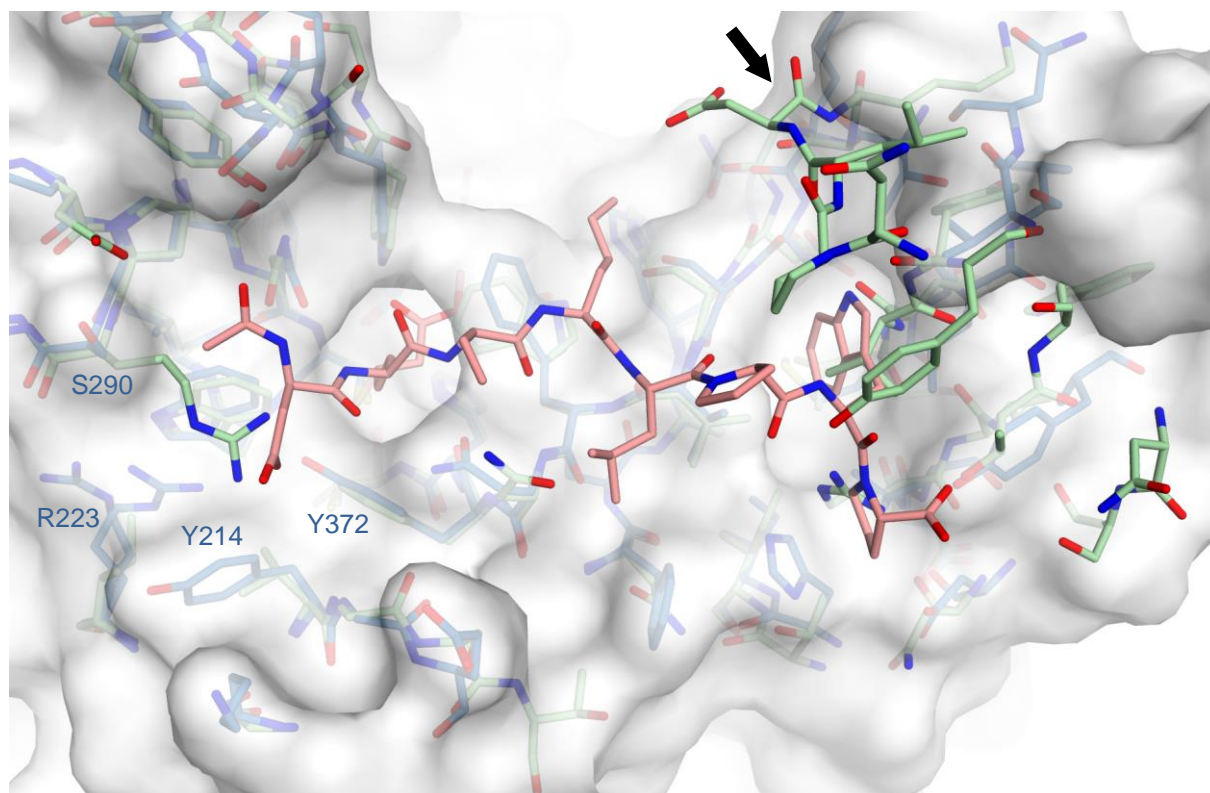


**Fig. 5.9:** Comparison of the three  $\text{Ca}^{2+}$  binding sites in FXIIIa° (blue, PDB ID: 4KTY,  $\text{Ca}^{2+}$  magenta) and TG3a° (green, PDB ID: 1L9N,  $\text{Ca}^{2+}$  cyan). a) The superimposed catalytic domains of FXIIIa° and TG3a° only differ in the orientation of the three-stranded  $\beta$ -sheet. b) At site 1 virtually the same coordination geometry as well as the salt bridge between the arginine and glutamate are found in both structures. c) Compared to the active state of FXIII (FXIIIa°), the coordination polyhedron around the calcium ion is not yet fully established as comparable interactions to Gln 349 are lacking and the coordination site of Asp 351 is not occupied by a functional group of the protein but by a water molecule in the TG3a° structure. As a consequence, the three-stranded  $\beta$ -sheet is not shifted (red arrow) and rotated and the substrate binding site which accommodates the ligand's indole moiety (ZED1301) in

active FXIIIa<sup>o</sup> is not formed. d) The third calcium ion binding site shows similar geometry as at the site in active FXIIIa<sup>o</sup>.

### 5.3.3 Active Site

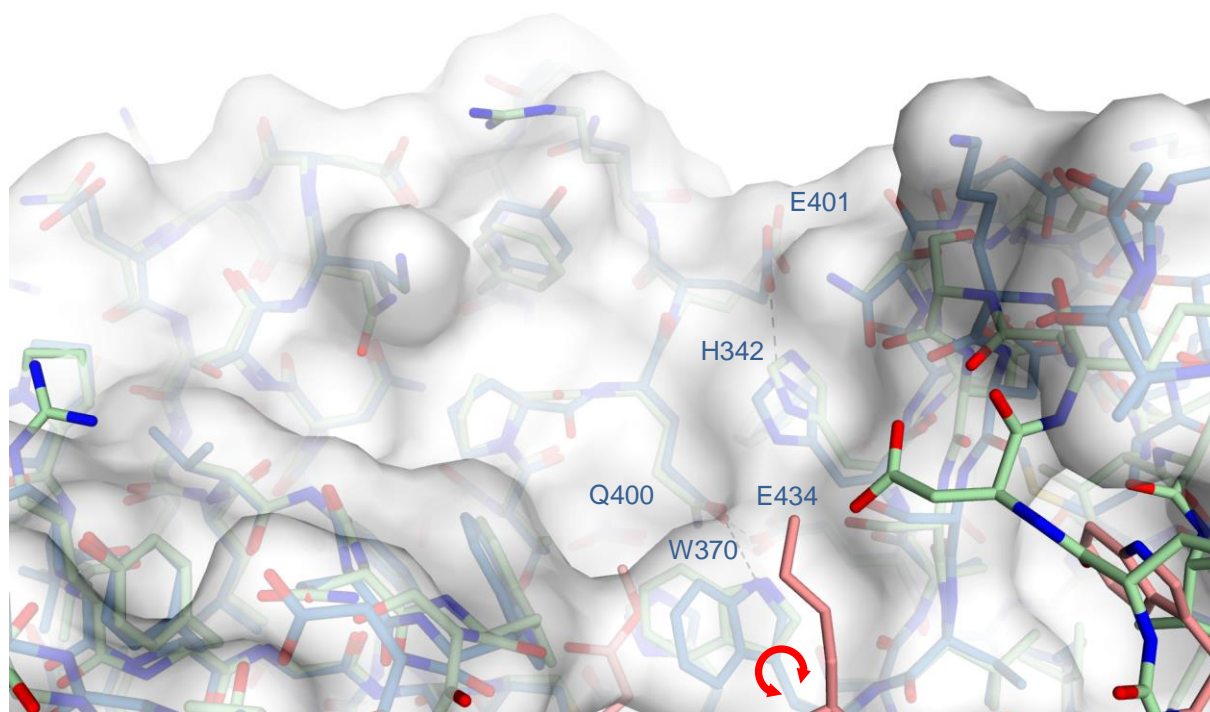
Comparable to TG2a\*, the arrangement of the amino acids of TG3a<sup>c</sup> in the  $\alpha$ -space corresponds to that found in FXIIIa<sup>o</sup> (Fig. 5.10). As in case of TG2a\* (Fig. 5.6a), only exchanges of amino acid side chains of the same four residues are observed. However, the variations of these amino acids can be used for the design of selective inhibitors of the corresponding transglutaminases. The  $\beta$ -space of TG3a<sup>c</sup> is also geometrically reduced compared to FXIIIa<sup>o</sup> due to the three-stranded  $\beta$ -sheet (black arrow) that remains, despite of the incomplete calcium coordination at site 2, in an orientation found for the inactive state of FXIIIa<sup>o</sup>.



**Fig. 5.10:** Substrate binding sites of the superimposed crystal structures of FXIIIa<sup>o</sup> (blue, PDB ID: 4KTY) and TG3a<sup>c</sup> (green, PDB ID: 1L9N). The surface is shown of FXIIIa<sup>o</sup>. In the  $\alpha$ -space four amino acids differ in their side chains compared to FXIIIa<sup>o</sup> (Ser 290 → Arg, Arg 223 → Ile, Tyr 214 → Val,

*Tyr 372 → Phe). The size of the  $\beta$ -space of TG3a<sup>c</sup> is slightly reduced because of the non-rotated three-stranded  $\beta$ -sheet (black arrow).*

In contrast to TG2a\*, in TG3a<sup>c</sup> the co-substrate binding site is fully established as in the crystal structure of FXIIIa<sup>o</sup>. The catalytic dyad and the H-bond to the indole ring of Trp 370 (FXIII nomenclature) are also formed. Even though, the arrangement of the  $\beta$ -barrel domains of TG3a<sup>c</sup> corresponds to the inactive state of a transglutaminase, TG3a<sup>c</sup> adopts a geometry obvious approaching closer the active state when compared to the TG2a\* structure in terms of transamidation activity. FXIIIa<sup>o</sup> and TG3a<sup>c</sup> differ only in the orientation of the three-stranded  $\beta$ -sheet and the orientation of the indole ring of Trp 370. The three-stranded  $\beta$ -sheet adopts an arrangement as observed in the inactive state of FXIII possibly due to the lack of a bound inhibitor. Interestingly, the indole ring of Trp 370 also adopts an orientation found for the inactive state, and not for the active one (red arrow in *Fig. 5.11*). Though, as already discussed in *Chapter 4.2.5*, the H-bond is formed between the indole NH and the carboxamide oxygen of Gln 400 (FXIII nomenclature).



**Fig. 5.11:** The co-substrate binding site is completely established as suggested by the superimposed crystal structures of TG3a<sup>c</sup> (green, PDB ID: 1L9N) and FXIIIa<sup>o</sup> (blue, PDB ID: 4KTY). This transition

*also comprises the catalytic dyad (His 342 and Glu 401, FXIII nomenclature) and the two amino acids Gln 400 and Glu 434 (FXIII nomenclature) of which Gln 400 accepts a hydrogen bond of the indole NH of Trp 370. In contrast to FXIIIa<sup>o</sup>, the indole ring of TG3a<sup>c</sup> adopts a conformation resembling the inactive state (red arrow), this observation might also result from the fact that the TG3a<sup>c</sup> structure does not host any active site ligand.*

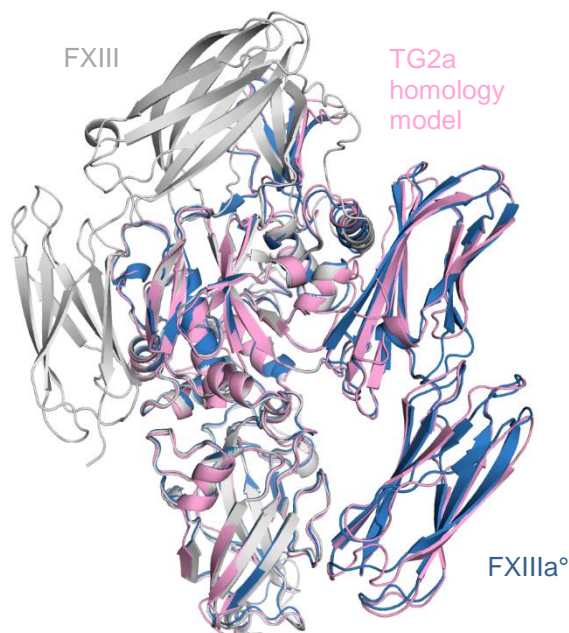


## 5.4 Comparison of FXIIIa° and a Homology Model of TG2a Based on the Structure of FXIIIa°

### 5.4.1 Overall Structure

A homology model of TG2a based on the crystal structure of FXIIIa° was generated (*Fig. 5.12*) to investigate whether active TG2 can in principle adopt the same conformation as FXIIIa° resulting from the fully established calcium coordination sites as well as the substrate and co-substrate binding sites (*Chapter 5.4.2 and 5.4.3*).

The generated homology model of TG2a\* has 0.8 % outliers and 80.2 % residues in most favored regions in the Ramachandran plot, whereas the template structure has 0.3 % outliers and 89.6 % residues in most favored regions (see *Chapter 5.7*). Thus, the quality of the model appears sufficient in terms of the backbone geometry. As constrained, in the homology model TG2a adopts the same arrangement of its domains as in the active state of FXIII (*Fig. 5.12*).

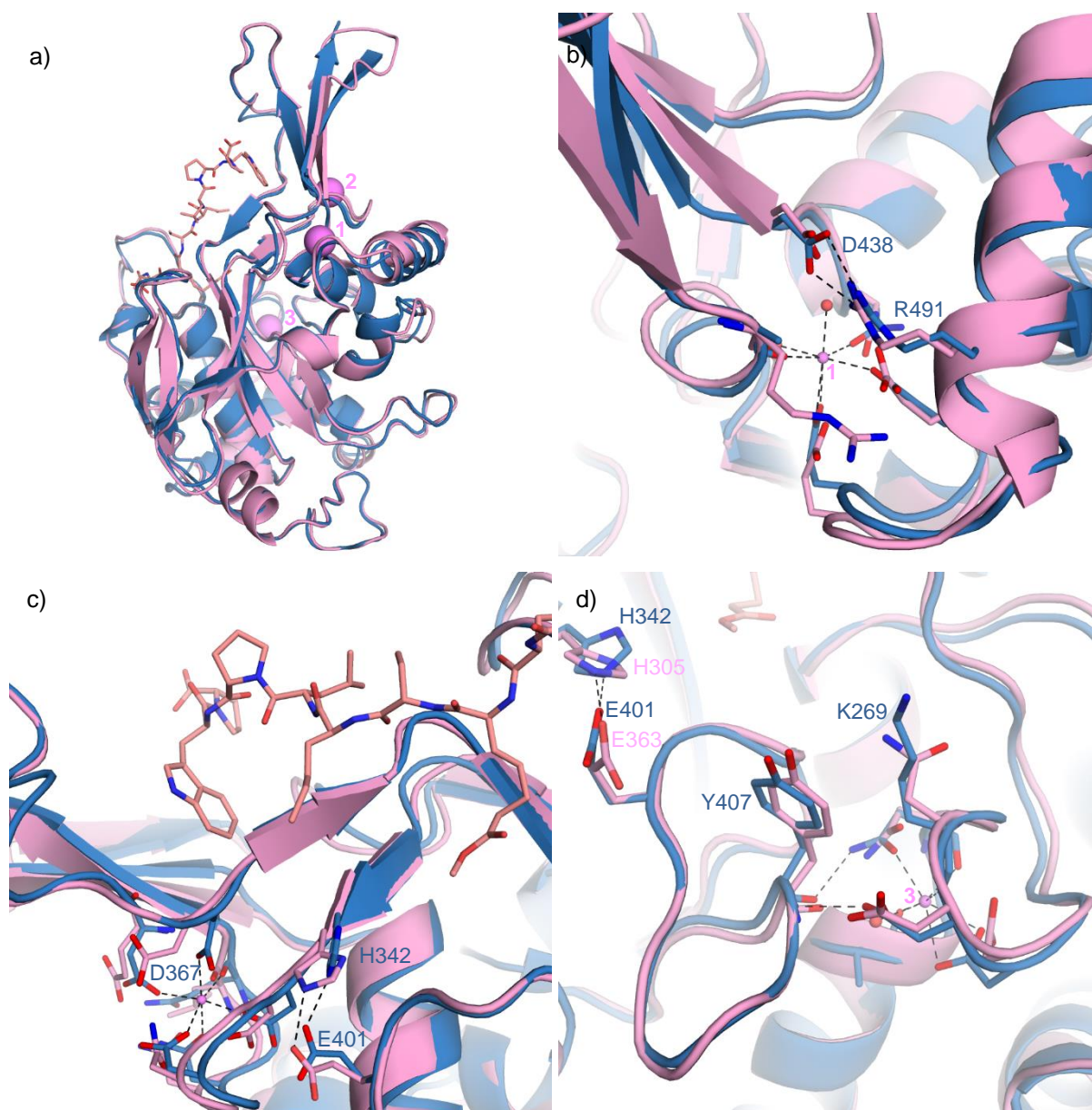


**Fig. 5.12:** Superposition of the inactive state of FXIII (grey, PDB ID: 1F13), the active state of FXIII (blue, PDB ID: 4KTY) and the homology model of TG2a (magenta) generated based on the crystal structure of FXIIIa°.

### 5.4.2 Calcium Binding Sites

As expected, the three calcium binding sites of the homology model can be equally formed with respect to the orientation of corresponding secondary structural elements like  $\beta$ -sheets,  $\alpha$ -helices and loop regions (*Fig. 5.13*). Much more importantly, those amino acids serving as coordination partners for the calcium ions, bear negatively charged side chains and are localized at the same positions as in the active state of FXIIIa°. Finally, the catalytic dyad is also established in the homology model of TG2a (*Fig. 5.13c,d*).

As previously stated in *Chapter 2.3.2*, the calcium coordinating amino acid Asp 351, that belongs to the three-stranded  $\beta$ -sheet, is a glutamate in TG2. Due to the longer side chain, the orientation of the three-stranded  $\beta$ -sheet is not affected to the same extent as in FXIIIa°. Additionally, Asp 367 of FXIII is in case of TG2 a glutamate. Thus, the corresponding loop might also not to be shifted downwards to the same amount as suggested by the homology model. Importantly, this loop affects the shape of the co-substrate binding site.



**Fig. 5.13:** Comparison between the three  $\text{Ca}^{2+}$  binding sites in FXIIIa° (blue) and in a homology model of TG2a based on FXIIIa° (magenta). a) The superimposed catalytic domains of FXIIIa° and the homology model are identical in the arrangement of secondary structure elements. b) At site 1 virtually the same coordination geometry as well as the salt bridge between the highly conserved arginine and glutamate are found in both structures. c) Calcium binding site 2 is established similarly to FXIIIa°, whereas Asp 351 and Asp 367 (FXIII nomenclature) are glutamates in TG2 resulting in a possibly less extensive rotation of the three-stranded  $\beta$ -sheet and minor loop reorientation. d) Also calcium binding site 3 can be perfectly established as in FXIIIa°.

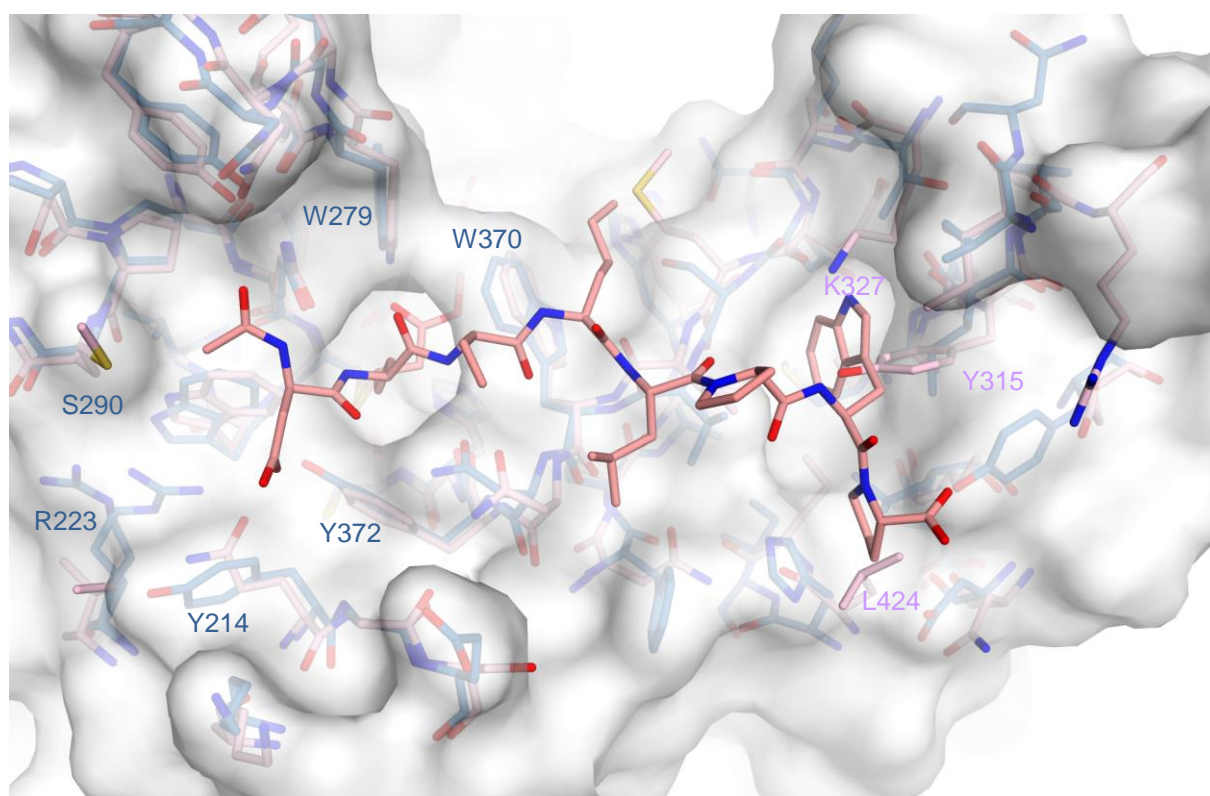
### 5.4.3 Active Site

In *Fig. 5.14* a superposition of the substrate binding sites of the homology model of TG2a and FXIIIa° is shown. The highly conserved amino acids Trp 279 and Trp 370 (FXIII nomenclature) which form the hydrophobic tunnel, are located in virtually identical positions. The same applies to the catalytic triad (not shown). Furthermore, the  $\alpha$ -space possesses the same amino acid arrangement as was the case in the crystal structure of TG2a\* in complex with the inhibitor Ac-P(DON)LPF-NH<sub>2</sub> (*Fig. 5.6a*).

Importantly, in contrast to the crystal structure of TG2a\* (*Fig. 5.6*), the inhibitor ZED1301 of the general  $\alpha(\text{wh})\text{xxxPW}$ -type fits into the active site of the homology model (*Fig. 5.14*). The inhibitor only clashes at the rear end of the  $\beta$ -space with the amino acids Tyr 315, Lys 327 and Leu 424. However, all these side chains appear flexible enough to move out of the space thus avoiding further clashes with an inhibitor of the general  $\alpha(\text{wh})\text{xxxPW}$ -type. This also confirms the quality of the homology model.

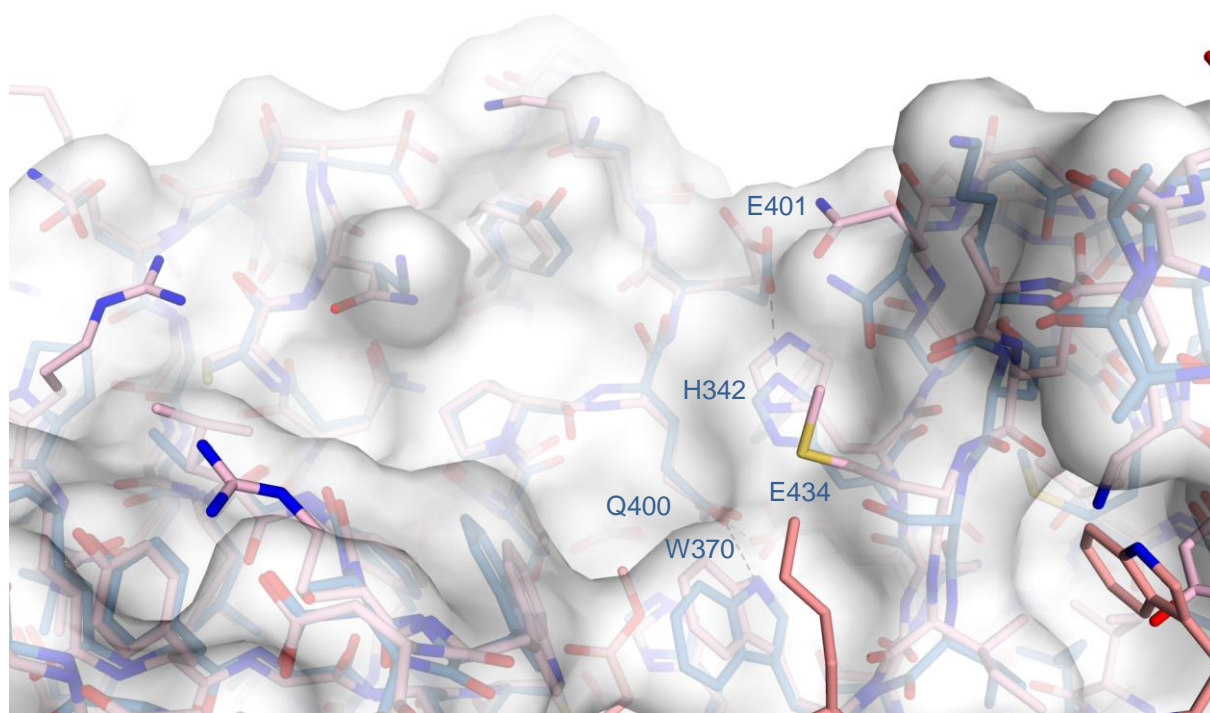
It should be mentioned that the affinity of ZED1301 towards TG2 (IC<sub>50</sub>: 2.9  $\mu\text{M}$ ) is 20-fold lower compared to FXIII (IC<sub>50</sub>: 110 nM). However, ZED1630 (*Chapter 3.2.1*) belongs also to the  $\alpha(\text{wh})\text{xxxPW}$ -type and binds to both enzymes with similar affinity (IC<sub>50,FXIII</sub>: 139 nM, IC<sub>50,TG2</sub>: 262 nM). The same observation was made for an inhibitor (not shown) that only differs in the  $\alpha$ -space and at P<sub>2</sub>' position (IC<sub>50,FXIII</sub>: 56 nM, IC<sub>50,TG2</sub>: 102 nM). Thus, the preferred binding of ZED1301 to FXIII should be predominantly affected by the  $\alpha$ -space. Considering the homology model (*Fig. 5.14*), TG2 bears a phenylalanine instead of a tyrosine residue (Tyr 372; FXIII nomenclature). Consequently, the negatively charged carboxylate residue of ZED1301 would be located in a more apolar environment. Additionally, the H-bond between the carboxylate oxygen atom of ZED1301 and the hydroxyl group of Tyr 372 cannot be established if ZED1301 would bind to TG2a, according to the homology model.





**Fig. 5.14:** Substrate binding site of the superimposed crystal structure of FXIIIa° (blue) and the homology model of TG2a based on FXIIIa° (magenta). The surface is shown of FXIIIa°. Conserved functional amino acids (like Trp 279 and Trp 370) are located at the corresponding positions to FXIIIa°, corroborating the relevance of the homology model. The inhibitor ZED1301 (light red) fits into the active site of the homology model. Clashes with the amino acids Tyr 315, Lys 327 and Leu 424 can potentially be avoided by side chain reorientation of the latter residues.

Finally, the co-substrate binding site of the homology model is also identically shaped to FXIIIa°. All conserved residues are at the right positions in contrast to the crystal structure of TG2a\* (Fig. 5.7). The catalytic dyad (His 342, Glu 401; FXIII nomenclature) is established as well as the corresponding residues forming the H-bond to the indole of Trp 370 (Gln 400, Glu 434; FXIII nomenclature).



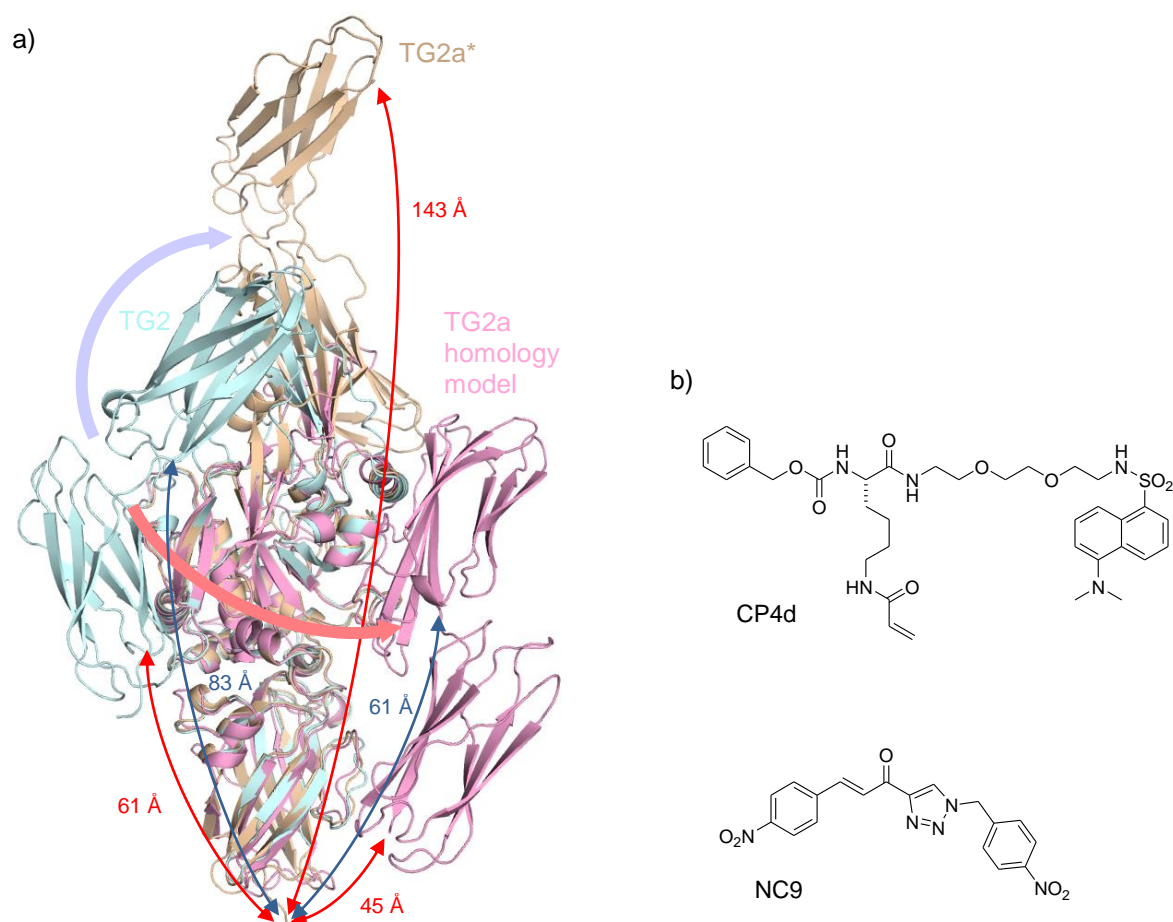
**Fig. 5.15:** The co-substrate binding site of the superimposed homology model of TG2a (magenta) onto the crystal structure of FXIIIa° (blue) suggested that it is fully established, comprises the catalytic dyad (His 342 and Glu 401, FXIII nomenclature) and the two amino acids Gln 400 and Glu 434 (FXIII nomenclature) of which Gln 400 accepts a hydrogen of the indole moiety of Trp 370.

It can be concluded that all functional and major structural elements of the homology model of TG2a are located at correct positions. The catalytic site comprising the catalytic triad, catalytic dyad and hydrophobic tunnel match well with the crystal structure of FXIIIa°. The calcium binding sites of the homology model correlate with those in FXIIIa° concerning the arrangement of negatively charged side chains around the calcium ions. Additionally, in contrast to the crystal structure of TG2a\*, the active site of the homology model provides sufficient space to accommodate the bulky  $\alpha$ (wh)xxxPW inhibitors.

As elucidated in *Chapter 5.2.2*, in contrast to FXIIIa\* and TG3a<sup>c</sup>, in TG2a\* the catalytic dyad is not established and mutational studies reveal that the highly conserved catalytic dyad is essential for transamidation activity.<sup>[76]</sup> Consequently, the linear state found in the crystal structure of TG2a\* most likely does not represent the active state of TG2 with respect to the transamidation activity. The fact that TG2 can act as a transglutaminase suggests that TG2 can adopt the same conformation as FXIIIa°. This assumption is reinforced by the homology model

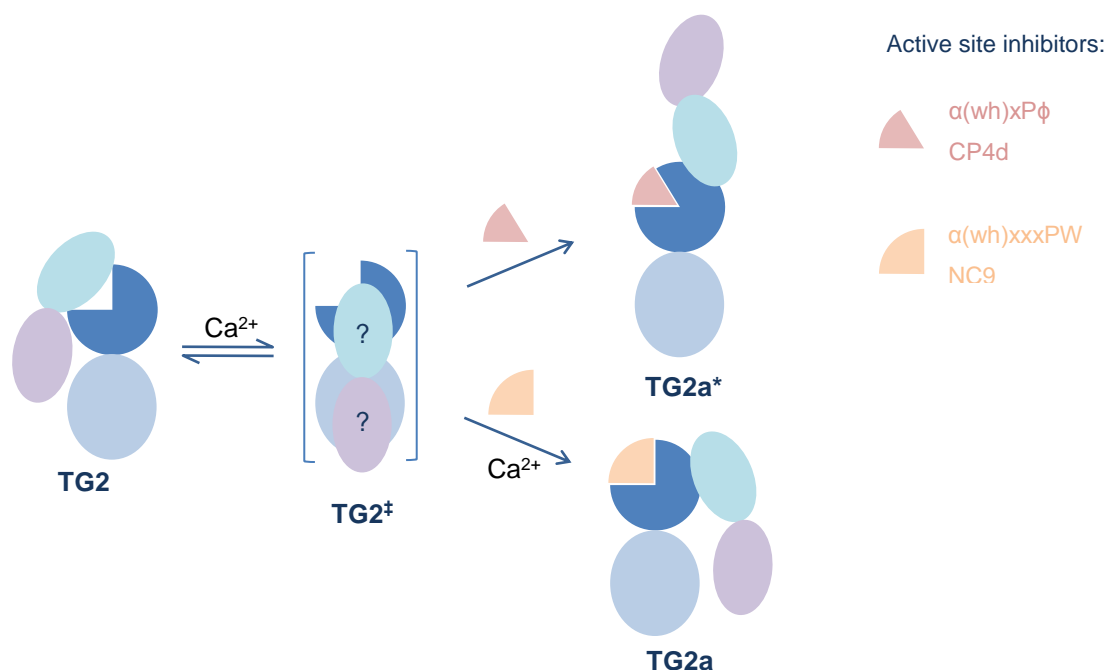
of TG2a generated based on the crystal structure of FXIIIa<sup>o</sup>. Herein, the catalytic dyad is located at the same position as of FXIIIa<sup>o</sup> and TG3a<sup>c</sup>.

The hypothesis that active TG2 can also exist in the same conformation as FXIIIa<sup>o</sup> is supported by FRET measurements with the irreversible inhibitor NC9 and the competitive, reversible inhibitor CP4d (*Fig. 5.16b*) performed by Caron and co-workers.<sup>[113]</sup> For this experiment, FRET donor and acceptor were located at N- and C-terminus of TG2. NC9 behaves as expected, meaning the FRET signal decreases owing to a larger distance between the N- and C-terminus in TG2a\* with its linear domain arrangement. Remarkably, CP4d leads to an increase of the FRET signal, indicating that TG2 adopts a conformation where the distance is shorter between N and C-terminus. This observation is in agreement with the homology model of TG2a based on FXIIIa<sup>o</sup> (see distances in *Fig. 5.16a*). In the globular active state, the distance between the  $\beta$ -sandwich domain and the  $\beta$ -barrel 2 domain is 45 Å, in contrast to the inactive state where the distance is 61 Å (red arrows). A distortion of the flexible  $\beta$ -barrel 2 domain reinforced by the crystal packing can be excluded by calculating the distances between the  $\beta$ -sandwich domain and the  $\beta$ -barrel 1 domain (blue arrows). Here the ratio of the distances is similar to the  $\beta$ -barrel 2 domain (61:45 vs 83:61).



**Fig. 5.16:** a) Superposition of the inactive state of TG2 (cyan), the crystal structure of TG2a\* (beige) and the homology model of TG2a (magenta). The distances between the N-terminus (Met 1 of TG2a\*) and the respective  $\beta$ -barrel 2 domain are indicated by red arrows. Since in the template (crystal structure of FXIIIa<sup>o</sup>) the  $\beta$ -barrel 2 domain does not interact with the protein and the arrangement can be a result of the crystal packing, the distances between the N-terminus and the respective  $\beta$ -barrel 1 domain were measured as well (blue arrows). b) Inhibitors used in the FRET experiment. CP4d inhibits the linear active state TG2a\*<sup>[113]</sup>, whereas NC9 inhibits most likely the globular active state TG2a.

Notably, some inhibitors synthesized by Zedira for FXIIIa<sup>o</sup> with the general  $\alpha(\text{wh})\text{xxxPW}$ -type ( $\alpha$ :  $\alpha$ -space addressing moiety, wh: warhead, x: any amino acid) have despite of their size similar affinity in the two to three-digit nanomolar range to FXIII and TG2 (see Chapter 5.4), also indicating that TG2 can adopt a very similar conformation in its active state as observed for FXIIIa<sup>o</sup> (Fig. 5.17).



**Fig. 5.17:** As suggested by the comparison of the different crystal structures, the homology model of TG2a based on the crystal structure of FXIIIa<sup>o</sup> and the FRET experiment<sup>[19, 60, 113]</sup>, TG2 appears to adopt depending on the inhibitor two different active states. Inhibitors of the  $\alpha(\text{wh})\text{xP}\phi$ -type and CP4d stabilize TG2 in the linear active state (TG2a\*) whereas inhibitors of the  $\alpha(\text{wh})\text{xxxPW}$ -type and NC9 force the enzyme into globular active state (TG2a). Since calcium ions already induces in the absence of an inhibitor or substrate a conformational change of TG2<sup>[19, 113]</sup>, TG2a adopts possibly a calcium-dependent pre-activated transition state (TG2<sup>‡</sup>). Herein, the location of the  $\beta$ -barrel domains cannot be defined. Departing from this transition state, TG2 can transform into its linear or calcium-dependent in its globular active state.

Why does TG2 adopt two different states distinguishing from the inactive state? The conformation adopted by FXIIIa<sup>o</sup> appears to represent the active state of a transglutaminase in terms of catalyzing the transamidation reaction due to the existence of the catalytic dyad.

Even though, in the linear state (TG2a\*) TG2 cannot act as a transglutaminase (catalyzing of transamidation) due to the lacking catalytic dyad, the catalytic site is accessible for the glutamine side chain of a putative substrate. Consequently, TG2a\* could still be able in the given conformation to hydrolyze the glutamine to a glutamate side chain. Thus, with some care the hypothesis can be put forward that TG2a\* could correspond to the deamidation state.

The following question arises: Which influencing factors decide whether TG2 adopts the linear or the globular active state? Judging by the crystal structures, the homology model and the FRET experiments<sup>[113]</sup>, it seems that the local calcium concentration, the inhibitor and, most likely, also the macromolecular substrate represents the crucial determinants. Chemical complementary inhibitors to the active site of TG2a\* keep the enzyme in the linear state, whereas chemical complementary inhibitors to the active site of TG2a force the enzyme into the globular active state. This would mean, possibly the shape and composition of the substrate determines whether the corresponding glutamine placed into the active site tunnel is trans- or deamidated. Exactly this assumption is in agreement with various experiments showing that TG2 specifically deamidates particular glutamine residues whereas other glutamine residues of the same protein are specifically transamidated.<sup>[7, 114, 115]</sup> Originally, it was assumed that only the presence or absence of the lysine-containing co-substrate and the local calcium concentration as well as the pH determines whether a substrate is deamidated or transamidated.<sup>[7, 97]</sup>

Regarding the crystal structure of TG2a\*, calcium coordination at site 2 possibly prevents formation of the  $\alpha$ -helix that in turn seems to trigger the linear arrangement of the domains. Overall with increased local calcium concentration, the linear active state (TG2a\*) is supposedly less preferred as the globular active state (TG2a). Conclusively, both the shape of the substrate and the local calcium concentration appears to determine which state (linear or globular active state) is preferably adopted by the enzyme.

In *Chapter 3.2.4*, it has been speculated about the role of the function of the highly-conserved H-bond between the indole ring of Trp 370 and Gln 400 (FXIII nomenclature). If TG2a\* would still exhibit deamidation activity without forming this H-bond, then this provides evidence that the H-bond functions as a required stabilizing element to keep the indole ring in the tunnel position and thus preventing water penetration into the catalytic site of transglutaminases. If TG2a\* represents the deamidation state of a transglutaminase, water penetration is even necessary to hydrolyze the thioester intermediate to the corresponding glutamate residue.

## 5.5 Summary & Conclusion

Detailed structural analysis of the crystal structure of FXIIIa<sup>o</sup>, TG2a\*, TG3a<sup>c</sup> and a homology model of TG2a based on FXIIIa<sup>o</sup> suggests that TG2 can adopt two calcium-induced active states, a linear (TG2a\*) and a globular active state (TG2a).

The linear domain arrangement of TG2a\* seems to be attributed to a rearrangement of strand III of the three-stranded  $\beta$ -sheet to an  $\alpha$ -helix. Importantly, amino acids of the  $\alpha$ -helix contributing to calcium binding site 2. Thus, in the linear state, the enzyme should not be able to establish transamidation activity, because the catalytic dyad cannot be formed (triggered by calcium binding site 2). Therefore, it is hypothesized that another state of the enzyme must exist, capable to catalyze the transamidation reaction (transamidation state). This state is most likely a globular active form similar to the one adopted by FXIIIa<sup>o</sup> in complex with ZED1301. This hypothesis is also corroborated by the homology model of TG2a generated based on the crystal structure of FXIIIa<sup>o</sup>, the FRET experiments performed by Caron and co-workers<sup>[113]</sup> and the fact that the FXIIIa<sup>o</sup> inhibitors of the  $\alpha$ (wh)xxxPW-type also inhibit TG2.

Even if TG2 cannot possess transamidation activity in the linear state (TG2a\*), the catalytic site is accessible for putative Q-substrates. Thus, it is assumed that the linear active form (TG2a\*) is competent to catalyze deamidation (deamidation state of TG2), whereas the globular active state (TG2a) represents the enzyme as catalyst for transamidation (transamidation state of TG2).

The FRET experiments<sup>[113]</sup> and the ability of FXIIIa<sup>o</sup> inhibitors of the  $\alpha$ (wh)xxxPW-type also to bind TG2 suggests that the complementarity of a substrate to the active site of TG2a or TG2a\* decides whether a substrate will be transamidated or deamidated. Remarkably, this is in line with experimental investigations revealing that TG2 is capable to specifically transamidate and deamidate glutamine side chains of the same protein.<sup>[7, 114, 115]</sup>

The fact that several amino acids are reorganized in an  $\alpha$ -helix and belong to calcium binding site 2 has to also be taken into account with respect to the development of drugs inhibiting TG2. The local calcium concentration can be important for the transition of the  $\beta$ -strand to the  $\alpha$ -

helix, an important structural feature to generate TG2a\*. Consequently, an increasing local calcium concentration could shift the equilibrium between TG2a\* and TG2a towards the latter. Possibly, even the affinity of different TG2 inhibitors could depend on the local calcium concentration and correlate to the state stabilized by these inhibitors. Finally, it can be hypothesized that the shape of TG2 blockers can possibly be designed in a way to inhibit more strongly the linear or the globular active state of TG2 which might give the option to develop ligands either inhibiting TG2 in diseases where the deamidation activity plays a crucial role. In contrast, a globular active state inhibitor could be more suitable where the transamidation activity is involved in pathogenesis.



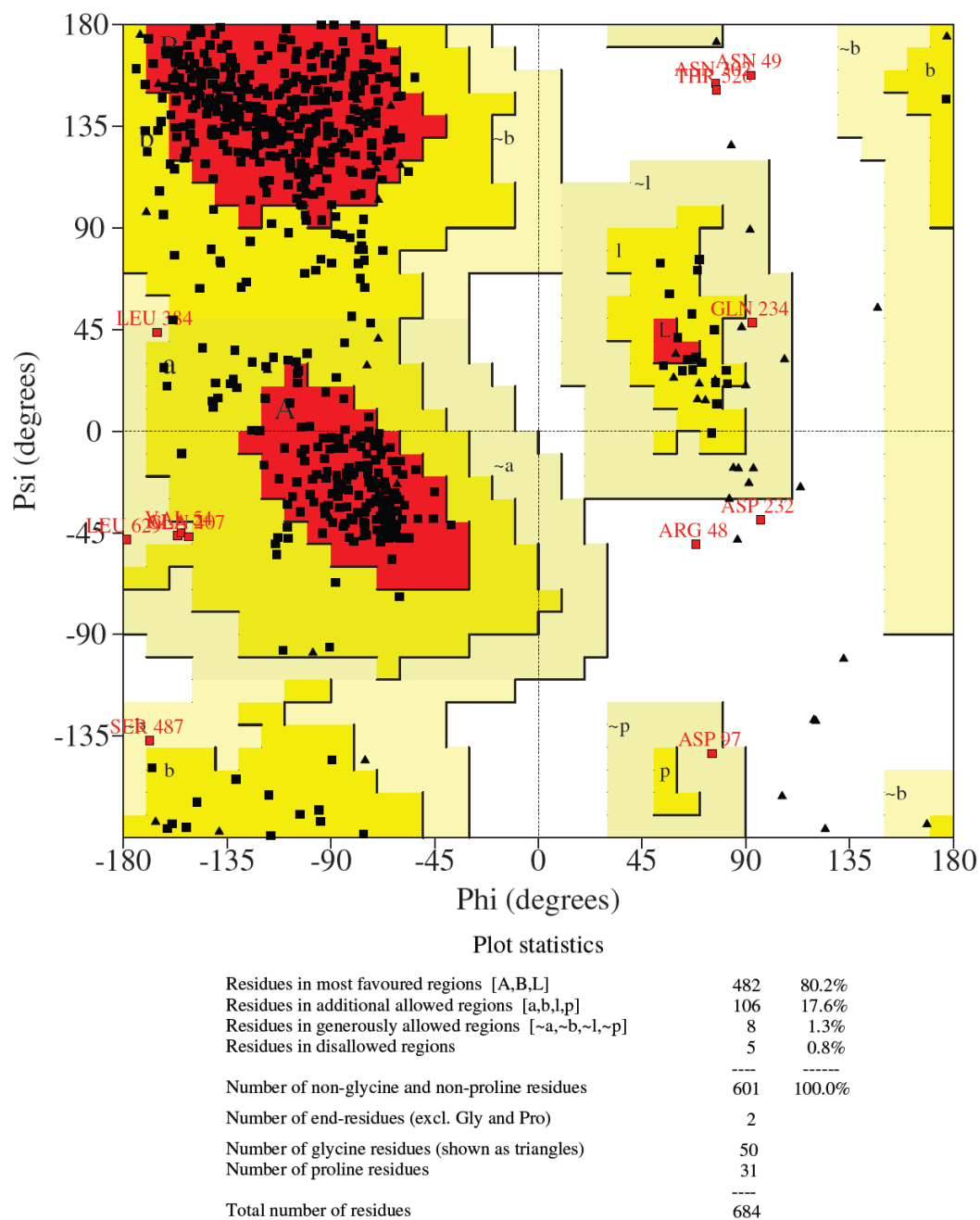
## 5.6 Experimental Part

### 5.6.1 Homology Model Building of TG2a

The homology model of TG2a was generated based on the crystal structure of FXIIIa<sup>o</sup> in complex with ZED1301 (PDB ID: 4KTY) by using the Molecular Operating Environment (MOE).<sup>[116-118]</sup> After sequence alignment of TG2 and FXIII (using the UniProt entries P21980 for TG2 and P00488 for FXIII), five main chain models were generated. For each main chain model one side chain sample was generated. Subsequently, each model was submitted to an electrostatics-enabled minimization run, which terminated when the gradient fell below the specified RMS gradient of 1. The temperature was set to 300 K and the Merck Molecular Force Field 94x (MMFF94x) was chosen. After protonation, the models were minimized as described above with an RMS gradient of 0.5. The five models were visually inspected and validated by comparing conserved regions of human transglutaminases comprising calcium binding sites, catalytic triad, catalytic dyad and hydrophobic tunnel. A Ramachandran plot of the homology model and the template structure was generated with Procheck.<sup>[90]</sup>

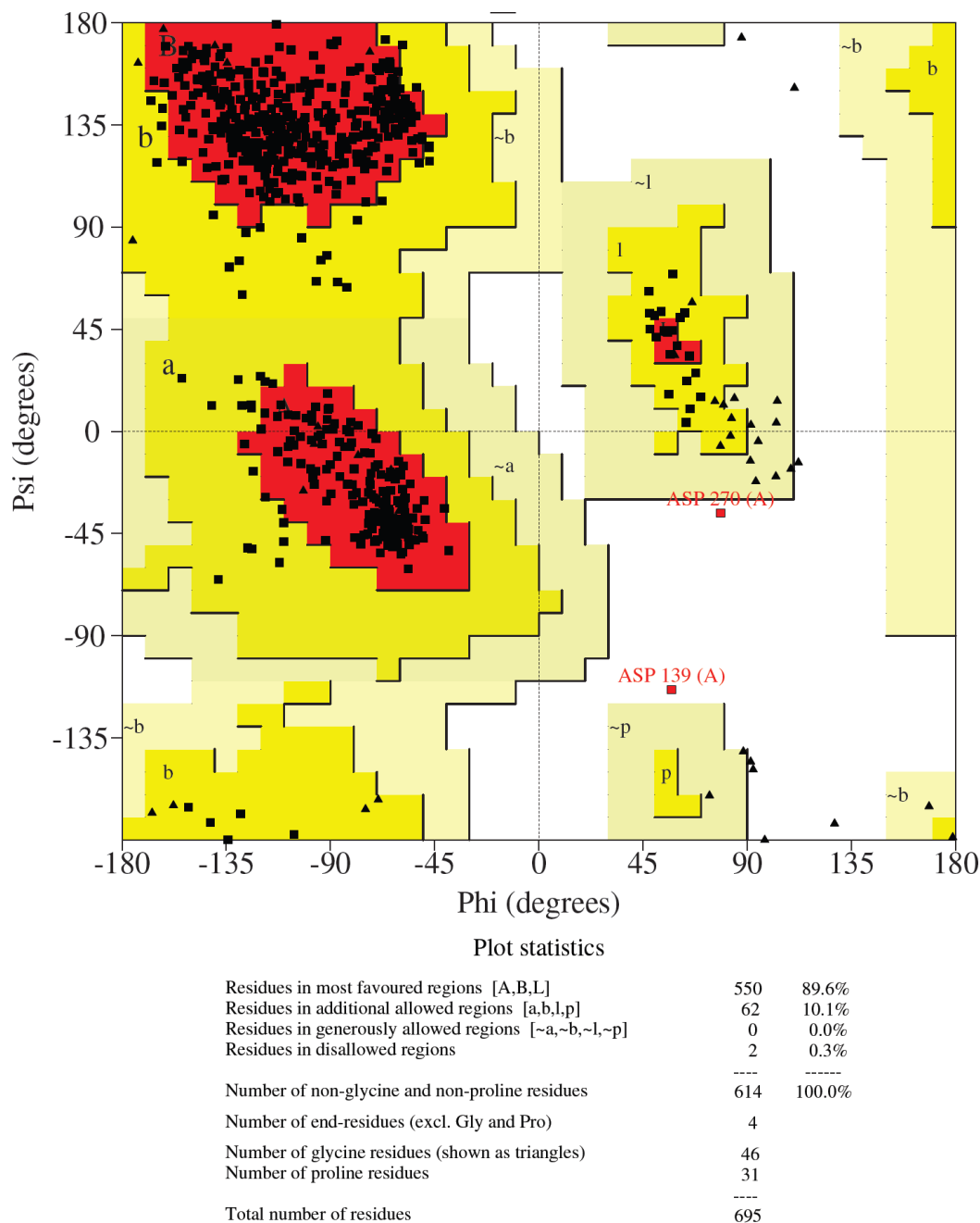
The distances between the  $\beta$ -sandwich domain and the  $\beta$ -barrel 2 domain of TG2 (PDB ID: 1KV3), TG2a\* (PDB ID: 2Q3Z) and the homology model of TG2a were measured with Pymol.<sup>[119]</sup> The crystal structures were structural superimposed on the basis of the C $_{\alpha}$  atoms of the  $\beta$ -sandwich and catalytic domain. Since in the crystal structure of TG2, TG2a\* and FXIIIa<sup>o</sup> the C- and the N-terminus are not completely defined in the electron density, the distance was measured between the C $_{\alpha}$  atom of Met 1 of TG2a\* and the C $_{\alpha}$  atom of Ile 683 of the respective structures. The distance between the  $\beta$ -sandwich domain and the  $\beta$ -barrel 1 domain was measured between the C $_{\alpha}$  atom of Met 1 of TG2a\* and the C $_{\alpha}$  atom of Tyr 583 of the respective structures.

## 5.7 Appendix



**Fig. 5.18:** Ramachandran-Plot (generated with Procheck<sup>[90]</sup>) of the homology model of TG2 based on the crystal structure of FXIIIa<sup>o</sup> (PDB ID: 4KTY). In contrast to the template structure (see next page),

the homology model has three further residues in disallowed regions und 80.2 % instead of 89.6 % in most favored regions.



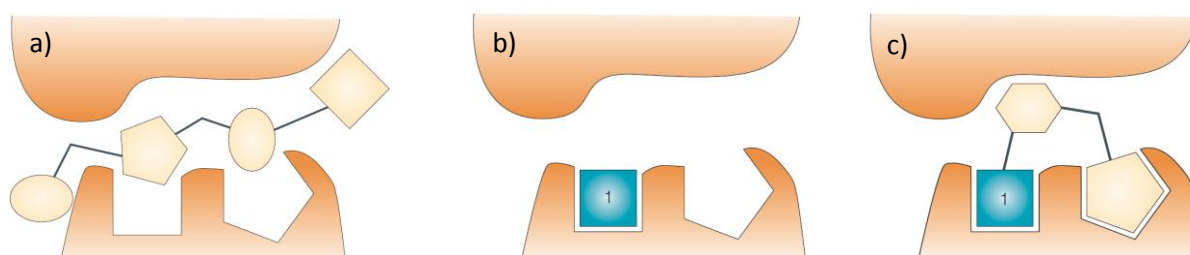
**Fig. 5.19:** Ramachandran-Plot (generated with Procheck<sup>[90]</sup>) of the crystal structure of FXIIIa° (PDB ID: 4KTY).

## 6 Crystallographic Fragment Screening Using the Example of the Surrogate Protease Endothiapepsin

### 6.1 Introduction

The development of FXIIIa blockers demonstrate the disadvantage of the popular approach in drug design, which starts with the search for initial hits with sufficient binding affinity to allow a structural characterization and to subsequently improve ligand's binding efficiency by tailored chemical modification. High molecular weight hits yielded by high-throughput screening usually possess reasonable starting affinity. However, the lead compound already adopts a certain binding geometry (*Fig. 6.1a*) restricting the subsequent optimization process. Moreover, the improvement of binding affinity by replacement of functional groups and interaction partners can turn out to be enormously time consuming as for example the search of a non-peptidic shortcut between the catalytic site and the hydrophobic pocket in case of FXIII. Furthermore, upon binding of a high molecular weight hit, the enzyme may be forced into a tensed geometry reducing binding affinity and therefore rendering the subsequent optimization more difficult.

A powerful alternative strategy to the above described classical approach is to start with much smaller ligands showing a molecular weight of up to 250 Da (fragments) and binding into different regions of the active site of the protein (*Fig. 6.1b*).<sup>[120, 121]</sup> Subsequently, fragments can be used as starting points to grow a ligand into further subpockets (*Fig. 6.1c*) or to chemically connect two fragments resulting in high affinity inhibitors of still rather low molecular weight.



**Fig. 6.1:** Hits from high-throughput screening of high molecular weight molecules do not fill the active site of an enzyme in an optimal way (a). In contrast, fragments occupy commonly only one subpocket with optimal interaction geometry (b) that can serve as a starting point for addressing further subpockets (c). Image taken from Rees et al. (slightly modified).<sup>[121]</sup>

The main issue of this method named in the literature as fragment-based lead discovery (FBLD) is to find fragments binding to the enzyme. The reason for this difficulty arises from the fact that small molecules generally represent weak binders. Usually putative binders are detected by applying a broad range of biophysically diverse assays like native mass spectrometry (MS), thermal-shift assay (TSA), saturation-transfer difference NMR, surface plasmon resonance (SPR) or micro thermophoresis (MST).<sup>[122-125]</sup> These methods are associated with a non-negligible amount of time but do provide only limited insights on how the molecule interacts with the target protein. Hence, it appears obvious to start immediately with a screening method that provides structural information directly like protein crystallography. Of course, crystallography is still rather time-consuming as well but this fact can be overcome by automatization of the entire process from crystallization via data collection to structure refinement.

FBLD would also be a powerful approach for the development of transglutaminase inhibitors, especially, concerning non-covalent transglutaminase blockers but also for the optimization within the  $\alpha$ -space and  $\beta$ -space. Unfortunately, the active site of transglutaminases is not accessible for the fragments because it is covered by the  $\beta$ -barrel 1 domain and the affinity of fragments is most likely not sufficient to successfully displacing this domain. Thus, removing the two  $\beta$ -barrel domains by molecular biology would provide a promising strategy to make a molecular species available for crystallization of weak binders like fragments and the development of alternative transglutaminase inhibitors possibly of non-peptidic character.

However, the obstacle producing such a truncated variant and then of finding new crystallization conditions would have to be resolved.

To investigate the above outlined strategy and start directly with a crystallographic screening, a library comprising 361 fragments were soaked into endothiapepsin crystals and for each fragment exposure a dataset was collected.<sup>[126-129]</sup> Due to the time-consuming process including data collection, processing of the corresponding datasets, visual inspection of the electron density, structure refinement and deposition, the fragment project has been split and transferred to several members of our research group. The range studied by the author of the present thesis covered the fragments 211 to 256. Here, eight binders were discovered and the corresponding structures were deposited in the protein data bank (see *Chapter 6.5.2*).

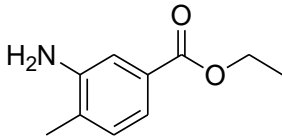
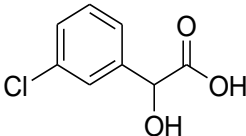
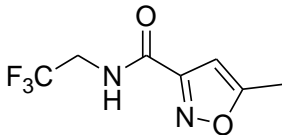
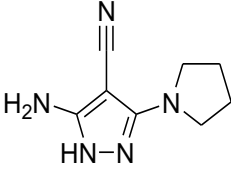
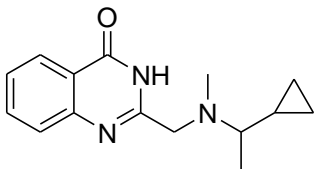
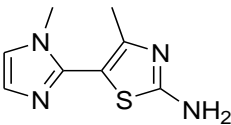
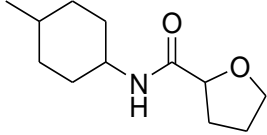
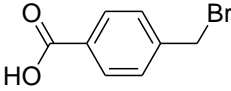
To classify the hits found in this project, fragments are related to the areas to which they bind namely the S<sub>1</sub>'-S<sub>3</sub>' pockets and in remote areas beyond the active site (remote binders).<sup>[126]</sup>

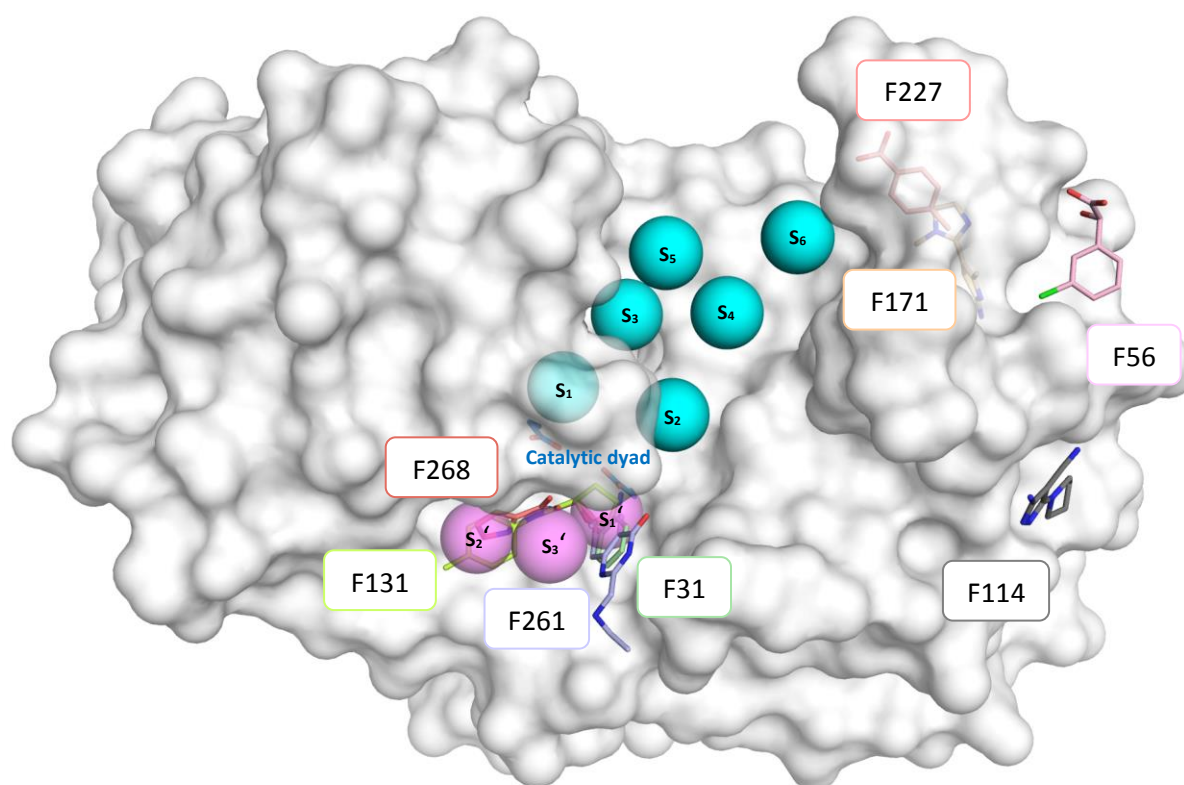
## 6.2 Results & Discussion

### 6.2.1 Overview

Four fragments (F31, F131, F261, F268) of the crystallographic screen bind to the  $S_1'$ ,  $S_2'$  and  $S_3'$  pocket and four fragments (F56, F114, F171, F227) bind remotely from the catalytic site of endothiapepsin (*Tab. 6.1, Fig. 6.2*). It should be mentioned that F114, F171 and F261 represent multiple binders. A further copy of F114 and F171 interact with the catalytic dyad, F114 directly and F171 via the catalytic water. Two additional copies of F261 occupies the  $S_3/S_5$  pockets.

**Tab. 6.1:** Fragments binding to the  $S_1'$ ,  $S_2'$  and  $S_3'$  pocket and remotely from the active site of endothiapepsin.

$S_1'$ , $S_2'$ and $S_3'$ pocket		Remote binder	
<b>F31</b>		<b>F56</b>	
<b>F131</b>		<b>F114</b>	
<b>F261</b>		<b>F171</b>	
<b>F268</b>		<b>F227</b>	



**Fig. 6.2:** Four fragments (F31, F131, F261, F268) bind to the  $S_1'$ - $S_3'$  pockets and four fragments (F56, F114, F171, F227) bind remotely from the active site of endothiapepsin. The subpockets of the active site are indicated by cyan ( $S_1$ - $S_6$ ) and magenta ( $S_1'$ - $S_3'$ ) spheres. The surface is shown of the apo structure (PDB ID: 4Y5L). Additional copies of multiple binders (F114, F171, F261) are not shown.

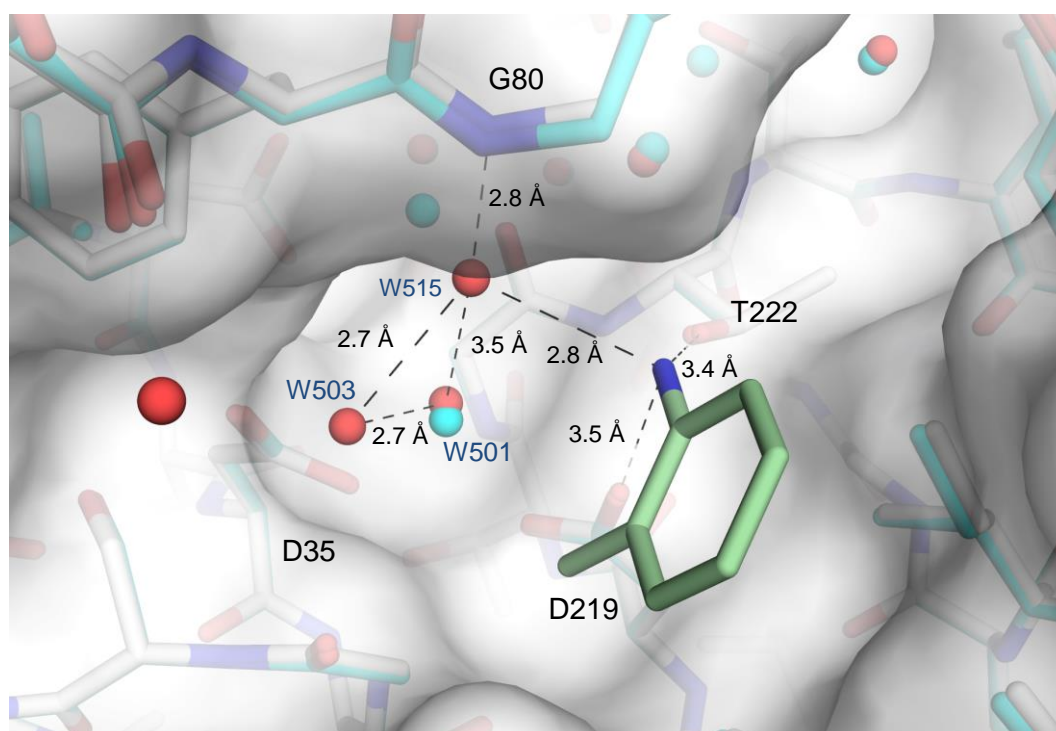
In the following, the exact interaction pattern of those fragments binding to the  $S_1'$ - $S_3'$  pockets and the remote binders will be described.



## 6.2.2 Fragments Addressing the S<sub>1</sub>'-S<sub>3</sub>' Subpockets

### F31

Fragment 31 binds to the S<sub>1</sub>' pocket, whereas the exocyclic nitrogen atom forms a weak hydrogen bond (3.4 Å) with the hydroxyl group of Thr 222 (*Fig. 6.3*). Much more interesting, the fragment induces a water network involving the catalytic water 501. In detail, the exocyclic amino nitrogen atom interacts with water 515 that in turn establishes a hydrogen bond with water 503. Finally, water 503 interacts with the catalytic water (W501). With other words, fragment 31 interacts water-mediated both with the flap region and the catalytic dyad. The exocyclic nitrogen atom also forms a direct H-bond with the carboxylate oxygen of Asp 219 of the catalytic dyad. The weak binding character (3.5 Å) can be presumably attributed to the methyl group. Considering that the ethyl ester of F31 is not defined in the electron density, an aniline molecule could represent an interesting fragment to address the catalytic dyad and the S<sub>1</sub>' pocket.

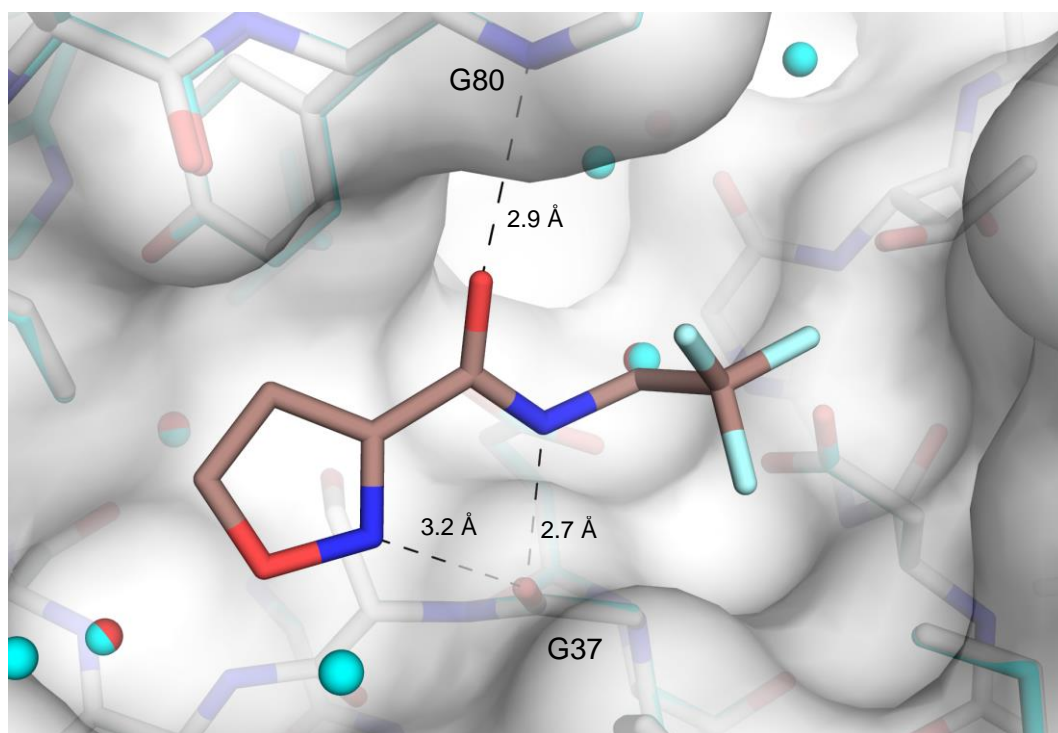


**Fig. 6.3:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 5DPZ) in complex with fragment 31 (green). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan. To simplify the representation, hydrogen bonds between the catalytic water (W501) and the catalytic dyad (Asp 35 and Asp 219) are not shown.

## F131

Fragment 131 occupies the  $S_1'$  pocket with its trifluoromethyl group and the  $S_3'$  pocket hosts its oxazole (Fig. 6.4). The methyl group of the oxazole is not resolved in the electron density map. The central carboxamide interacts with the flap by accepting a hydrogen bond of Gly 80 NH with the carboxamide oxygen, whereas the adjacent carboxamide NH forms a hydrogen bond to the opposite oxygen of Gly 37 that is additionally addressed by the isoxazole NH of F131. However, the  $pK_a$  of the isoxazole nitrogen falls into a range making protonation unlikely even though the pH of the crystallization buffer was 4.6. Thus, the nitrogen predominantly exists in the deprotonated state. Accordingly, replacement of the oxazole against e.g. a pyrrolidine heterocycle might be an obvious option to improve fragment binding. Furthermore,

the hydrogen bond has only a moderate strength with 3.2 Å. Replacement by a six-membered ring (piperidine) could show whether the different angle relationships in the larger ring allows for better interactions with the binding site.

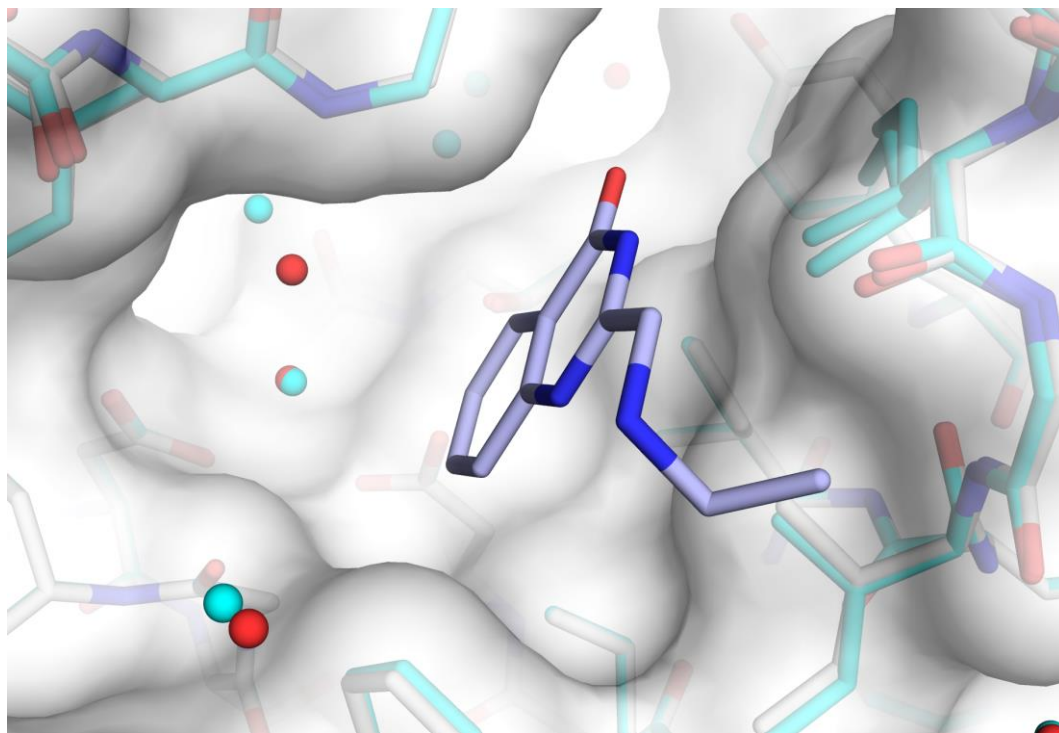


**Fig. 6.4:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4Y4E) in complex with fragment 131 (brown). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

## F261

Compound F261 binds with its quinazolinone in the area of the  $S_1'$  pocket (Fig. 6.5), whereas the cyclopropyl moiety points towards the solvent and thus it is not completely defined in electron density (see Chapter 6.5.1). Interestingly, the fragment does not form any polar interactions with the protein. The occupancy of the fragment was refined to 0.86 suggesting a weak affinity of the fragment towards the target enzyme endothiapepsin in this binding pose. However, ITC measurements reveal F261 as a potent binder with a  $K_d$  of 0.6 mM.<sup>[129]</sup> But F261 represents a multiple binder and consequently there is no direct correlation between one of the

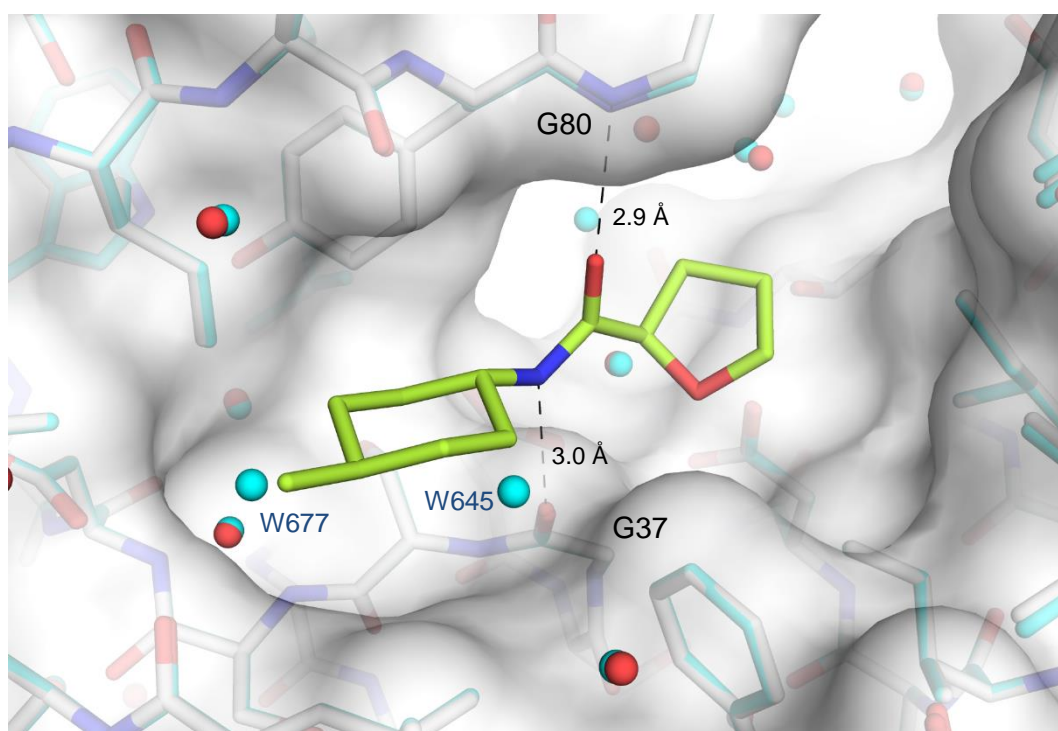
three binding poses and binding affinity because the binding affinity distributes on the three different binding poses.



**Fig. 6.5:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4Y5B) in complex with fragment 261 (light purple). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

## F268

This fragment interacts with its carboxamide similar to F131 (*Fig. 6.6*). The carbonyl oxygen accepts the hydrogen atom of the backbone NH of Gly 80 and the carboxamide nitrogen interacts via a hydrogen bond with the backbone oxygen of Gly 37. However, the carboxamide is reversely oriented compared to the latter meaning the tetrahydrofuran portion occupies the  $S_1'$  pocket and the methyl cyclohexyl part is located in the  $S_2'$  pocket whereas the methyl cyclohexyl ring displaces two water molecules (water 677 and 645) in the  $S_2'$  pocket.

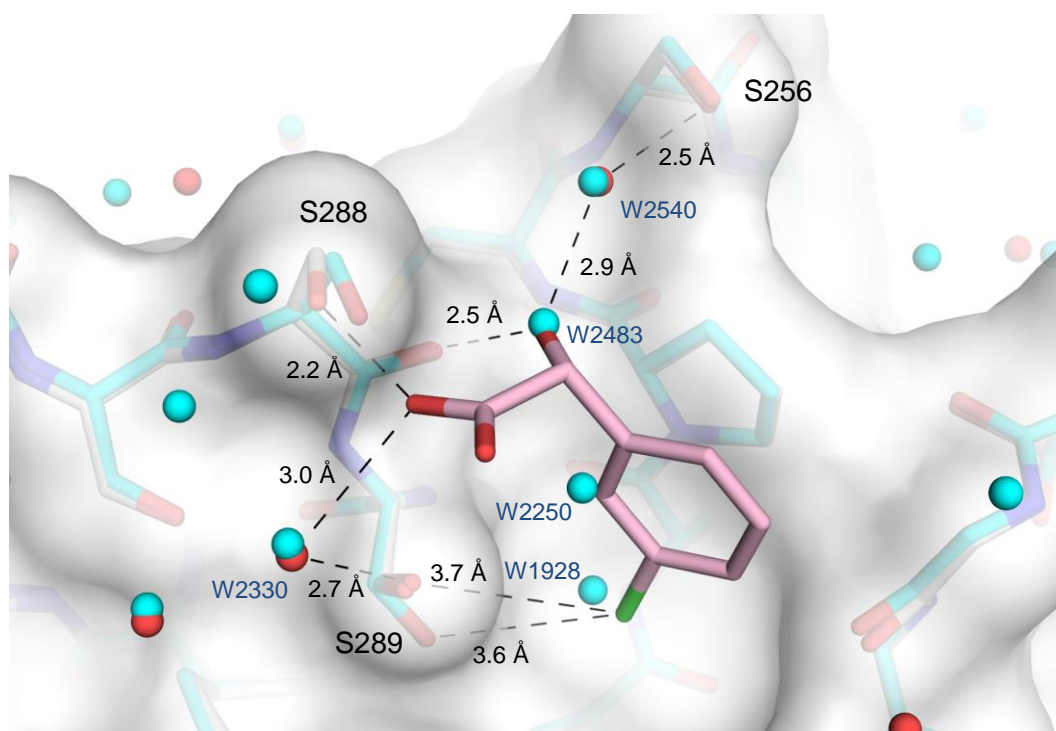


**Fig. 6.6:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 5Y5E) in complex with fragment 268 (green). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

### 6.2.3 Fragments Bind Remotely from the Active Site

#### F56

Fragment 56 belongs to the group of remote binders (Fig. 6.7). One carboxylate oxygen of fragment 56 acts as a hydrogen acceptor for the hydroxyl group of Ser 288 and interacts water-mediated (water 2330) with the hydroxyl group of Ser 289. The hydroxyl group of F56 donates a hydrogen bond to the carbonyl oxygen of Ser 288 and interacts via water 2540 with the hydroxyl group of Ser 256. Comparison of the crystal structure of endothiapepsin in complex with F56 with the apo structure shows that the fragment displaces three water molecules. The hydroxyl group of F56 is exactly located at the position where water 2483 is located in the apo structure.

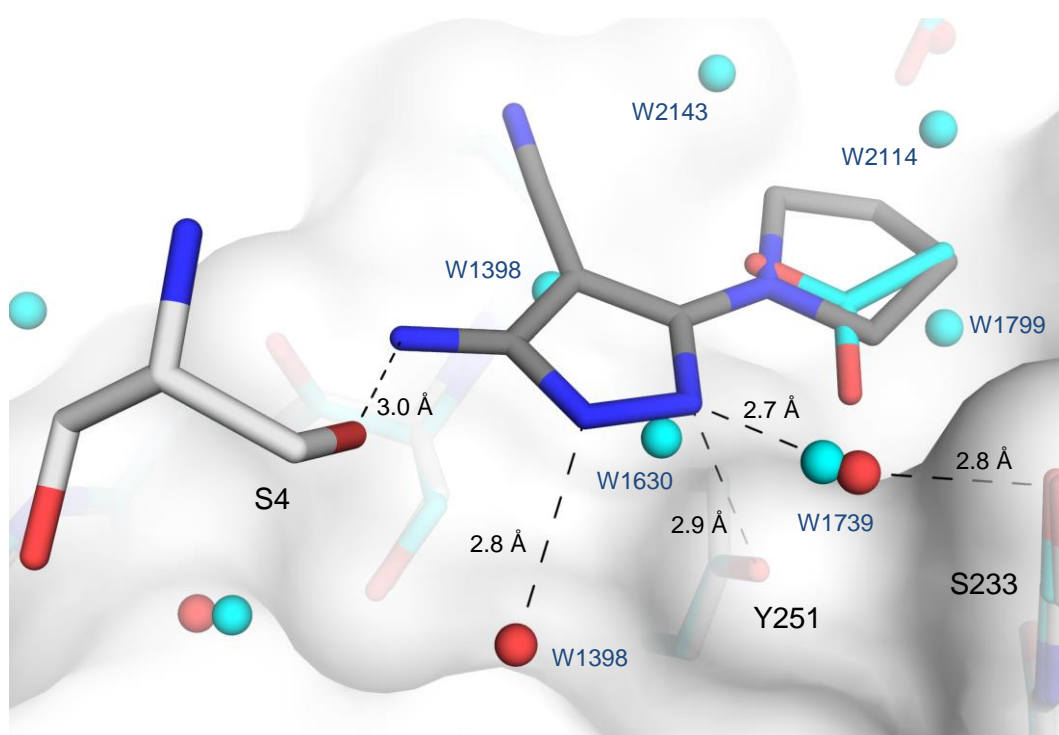


**Fig. 6.7:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4Y54) in complex with fragment 56 (pink). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

## F114

This fragment interacts directly with the protein via one hydrogen bond (Fig. 6.8). Here, the endocyclic pyrazole nitrogen forms a hydrogen bond with the hydroxyl group of Tyr 251. Additionally, the nitrogen interacts water-mediated (W1739) with the hydroxyl group of Ser 233. The second endocyclic nitrogen forms a hydrogen bond with water 1398. Finally, the exocyclic nitrogen of F114 interacts with the hydroxyl group of Ser 4 of a symmetry-related molecule via a hydrogen bond. The fragment displaces five water molecules, and the pyrrolidine portion is located at the position where an acetate ion is found in the apo structure.

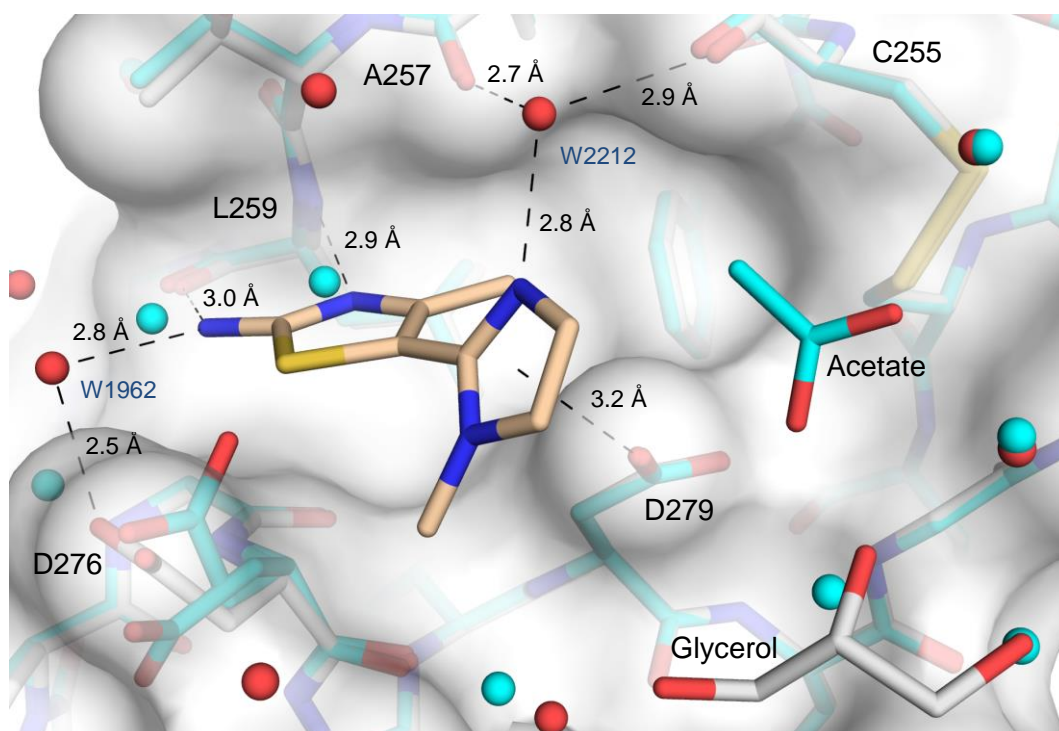




**Fig. 6.8:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4Y4T) in complex with fragment 114 (dark grey). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

## F171

Fragment 171 forms two direct hydrogen bonds to the protein (*Fig. 6.9*). One is formed by the endocyclic nitrogen of the thiazole moiety and the backbone nitrogen of Leu 259. The other one is formed between the exocyclic nitrogen of the thiazole moiety and the backbone carbonyl oxygen of Leu 259. The exocyclic nitrogen interacts furthermore water-mediated with Asp 276. The imidazole nitrogen interacts with the backbone oxygen of Ala 257 and Cys 255 via water 2212. Moreover, the  $\pi$ -electron system of the imidazolium ring interacts with the carboxyl oxygen of Asp 279. Notably, due to the interaction profile between F171 and the protein and for steric reasons, the thiazole and imidazole ring are twisted against each other by an angle of 63.0 degrees. In close vicinity to the fragment, a glycerol molecule is bound (lower left part of *Fig. 6.9*) which does not interact with F171.

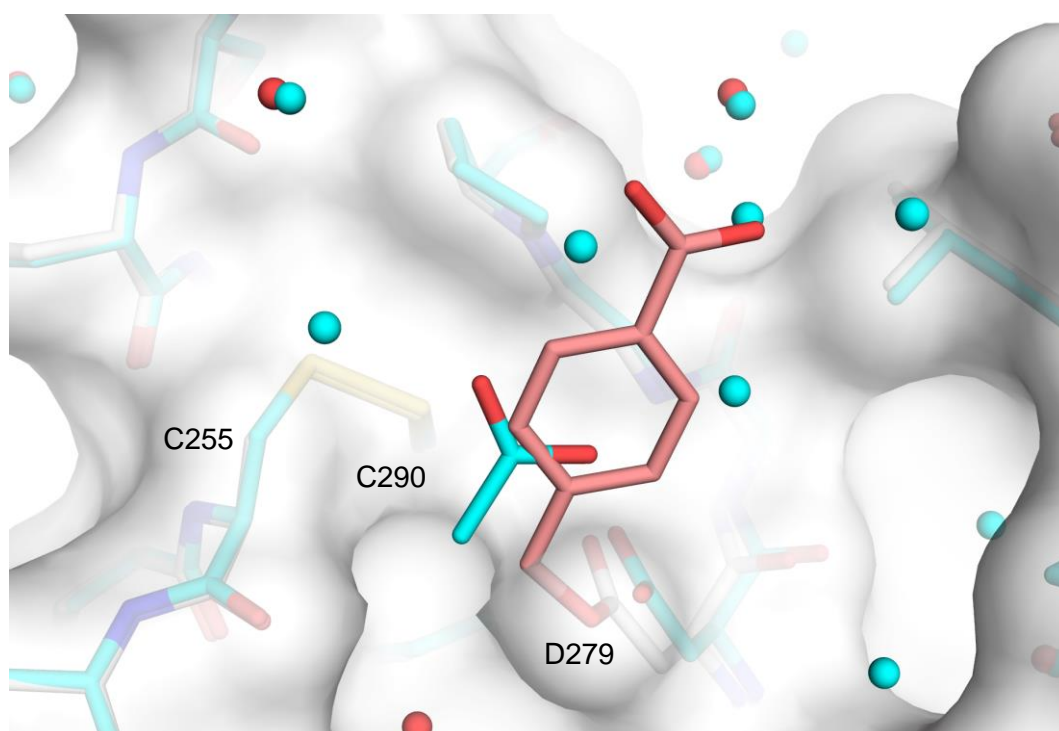


**Fig. 6.9:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4Y3X) in complex with fragment 171 (beige). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

## F227

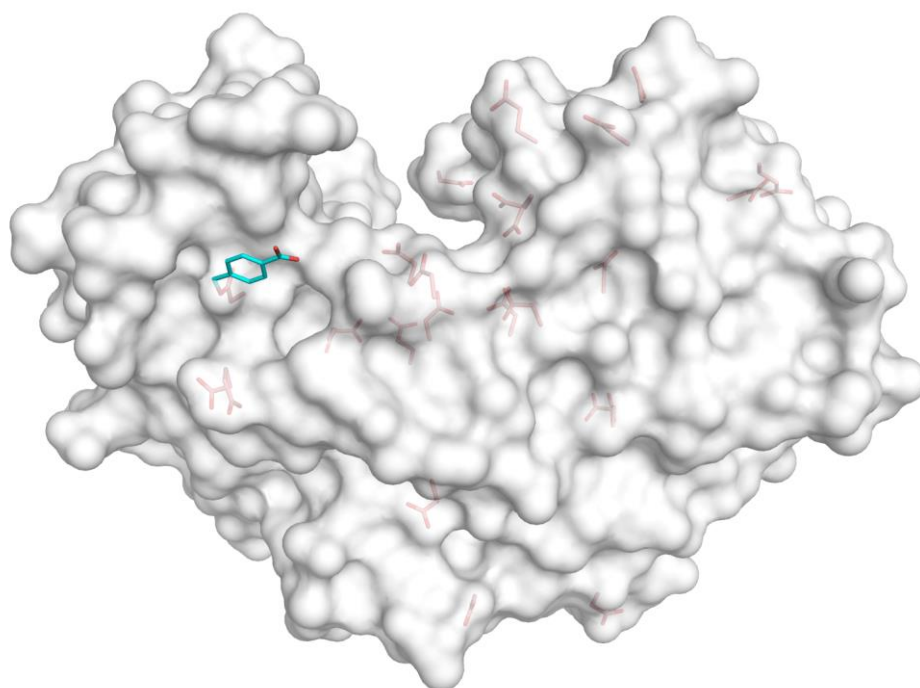
Probably, the most interesting fragment of the entire screen represents F227, which binds covalently to the carboxyl group of Asp 279 of endothiapepsin (Fig. 6.10). The fragment is a benzyl bromide in its unbound state, and the bromine has been substituted by one of the aspartate oxygen atoms of Asp 279 in the bound state.





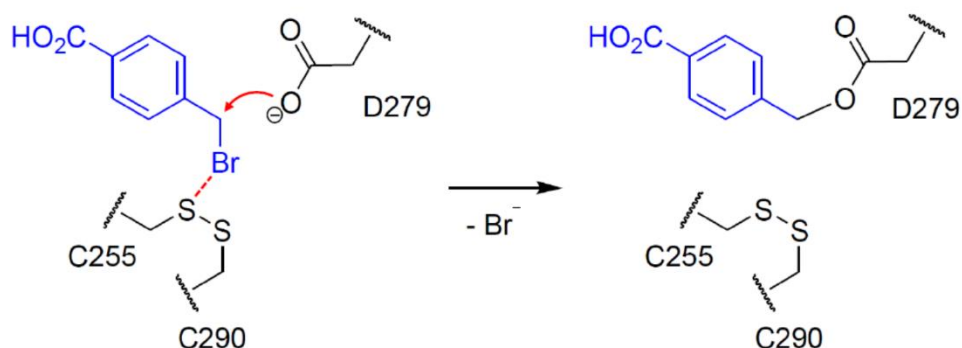
**Fig. 6.10:** Crystal structure of endothiapepsin (grey sticks, white surface, PDB ID: 4YD4) in complex with fragment 227 (light red). The apo structure is shown in cyan (PDB ID: 4Y5L). The water molecules of the complex structure are colored in red and of the apo structure in cyan.

Remarkably, as shown in Fig. 6.11, the fragment only binds to Asp 279, although in total there are 21 aspartate and glutamate side chains located at the surface of endothiapepsin and therefore in principle accessible for F227. Thus, the covalent attachment seems to be highly selective. Regarding the binding site in terms of polar interaction partners available to the fragment which could facilitate the binding to Asp 279, gives no indication for the specific binding of F227. Also the influence of symmetry-related molecules can be excluded after inspection of the crystal packing. Nevertheless, it is interesting to notice that an acetate ion is found close to the binding site of F227 in the apo structure which is replaced upon fragment accommodation.



**Fig. 6.11:** Endothiapepsin possesses 21 solvent-exposed aspartate (red, PDB ID: 4YD4) and glutamate residues but the fragment 227 (cyan) binds selectively to Asp 279.

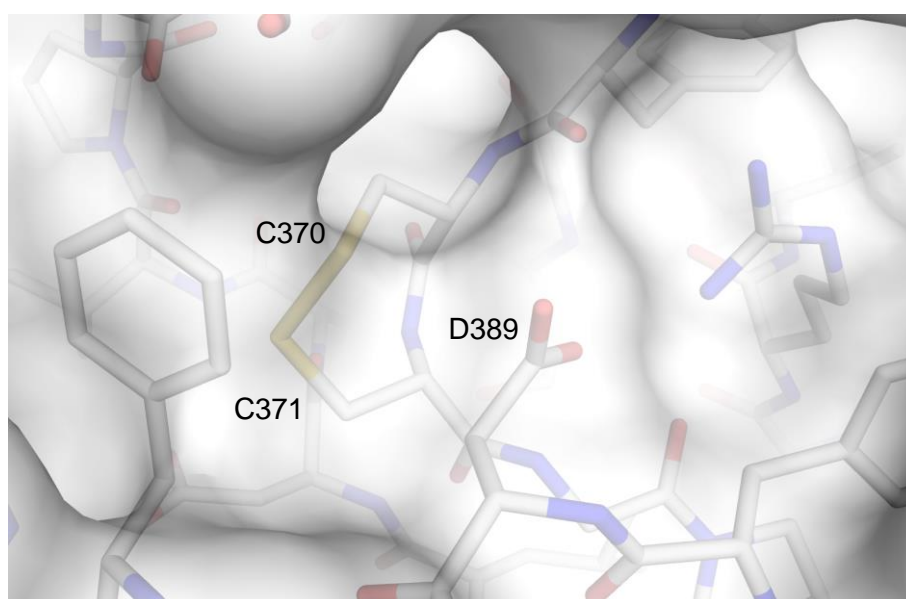
Interestingly, regarding the binding site of F227 (Fig. 6.10), a disulfide bond is present, formed between Cys 255 and Cys 290 in close vicinity (3.8 Å) to the bound fragment. Since disulfides can act as electrophiles, as for instance in thiol-disulfide exchange reactions<sup>[130]</sup>, possibly, an interaction of the bromine of F227 with the disulfide facilitates the nucleophilic attack of the carboxyl oxygen of Asp 279 (Fig. 6.12).



**Fig. 6.12:** Proposed mechanism of the covalent attachment of F227 to Asp 279.

Regarding the apo structure of endothiapepsin, the displaced acetate ion binds next to the disulfide bond in close proximity at the position where F227 is found. This is even more surprising by taken into account that there are no positively charged residues in this area. Why does the acetate ion exactly binds to this site of the enzyme? It can be speculated whether the disulfide bridge serves with its electrophilic character as an interaction partner for the negatively charged acetate ion.

The above assumed influence of the disulfide bond on carboxylate groups has to be investigated by evaluating other proteins. Transglutaminase 2 could serve as an appropriate candidate since also its disulfide bond is found in the neighborhood of an aspartate (4.0 Å) with similar geometry as in endothiapepsin (*Fig. 6.13*). Notably, this disulfide bond is assumed to be responsible for the disulfide isomerase activity of TG2.<sup>[15, 16]</sup> Thus, it can be speculated whether a corresponding modification of the benzyl bromide could even work as a low molecular inhibitor that affects the disulfide isomerase activity.



**Fig. 6.13:** Disulfide bond with an adjacent aspartate of Transglutaminase 2 (PDB ID: 2Q3Z).

### 6.3 Summary & Conclusion

Four fragments (F31, F131, F261, F268) out of the crystallographic screen comprising 361 molecules were observed to bind to the  $S_1'$ - $S_3'$  pockets. Four fragments (F56, F114, F171, F227) bind remotely from the active site of endothiapepsin. Both fragment 131 and 268 interacts with the flap by forming a hydrogen bond to the backbone nitrogen of Gly 80 and serves as potent interaction partner for fragments binding to adjacent subpockets of the active site. The remote binders show impressively that certain fragments bind to the surface of a protein without the presence of deep pockets. Thus, the FBLD approach could possibly be applied to generate ideas how to address the  $\beta$ -space by appropriate moieties in the development of transglutaminase inhibitors.

Fragment 227, that binds covalently and highly selective to Asp 279, turned out as exceedingly interesting. Possibly an adjacent disulfide bond facilitates the formation of covalent attachment with the benzyl bromide by nucleophilic substitution reaction.

In academic and industrial research, it is highly desired to selectively modify proteins in mixtures with other proteins with labels that enable specific detection (fluorescence microscopy), protein purification (His-tag), protein labeling (SPR spectroscopy, investigation of conformational changes with ESR spectroscopy) and therapeutic conjugation (PEGylation for improvement of the plasma half-life of therapeutic proteins).<sup>[131-135]</sup>

Usually, proteins are modified at the C- or N-terminus.<sup>[132, 133]</sup> Unfortunately, the C- and N-terminus are not always accessible or such additional modifications affect the activity of the protein in an undesired way. An alternative approach is the selective mutation of particular amino acids at the surface of the protein by lysine or cysteine residues which can be subsequently modified by covalent linkage. In case of lysine mutation the remaining lysine residues have to be mutated to guarantee selective modifications. However, these mutations may affect the isoelectric point and enhance the immunogenicity which is problematic especially for therapeutic proteins. The introduction of cysteine mutations can result in undesired cross-linking of the modified protein by formation of disulfide bonds with other thiol groups.

The benzyl bromide discovered in the present example could give a unique possibility to modify a protein at special position. Notably, transglutaminases also represent a promising tool for specific modification of proteins (PEGylation of therapeutic proteins).<sup>[136]</sup>

## 6.4 Experimental Part

### 6.4.1 Processing and Refinement

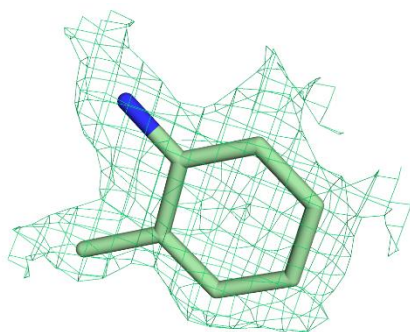
Datasets of each fragment in the range of F211-F256 were collected at the BESSY II electron storage ring<sup>[79]</sup> and reprocessed using XDS<sup>[98]</sup>. Datasets with insufficient statistical power (resolution > 2.0 Å, R-merge > 10 %) were recollected and processed again.

Afterwards the 45 datasets were subjected to Fourier synthesis using Phenix<sup>[88]</sup>, and the crystal structure of endothiapepsin with the PDB ID 3PCW (P2<sub>1</sub>) or 4Y3M (P2<sub>1</sub>2<sub>1</sub>2<sub>1</sub>) served as starting model depending on the space group. As first two refinement steps, rigid body refinement (at 4.0 Å) and simulated annealing were applied. Subsequently, the difference maps (> 3.0  $\sigma$ ) of the whole enzyme of all datasets were visually inspected. Those datasets where the difference map indicated a bound fragment were further refined anisotropic and water molecules were assigned manually. Finally, the structures with clearly defined difference electron density for the fragment were completely refined and deposited in the protein data bank. Model building was performed with Coot.<sup>[87, 108]</sup> Every dataset of bound fragments was refined anisotropic in respect to the B factors with and without water, with TLS groups and with hydrogen atoms. Restraints of the fragments were generated using the GRADE server.<sup>[99]</sup>

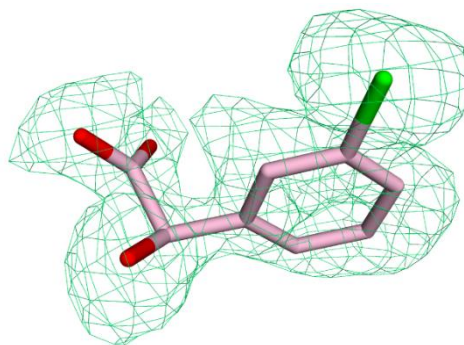
## 6.5 Appendix

### 6.5.1 Difference Electron Density Map (2.5 sigma)

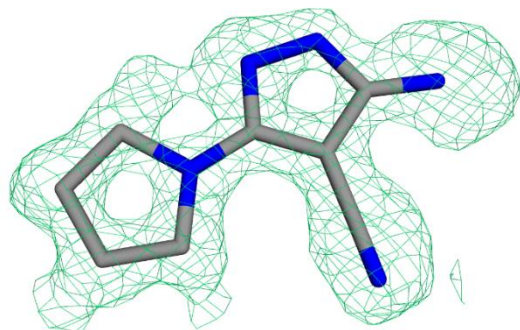
F31



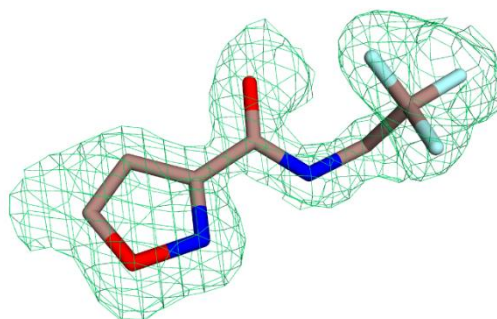
F56



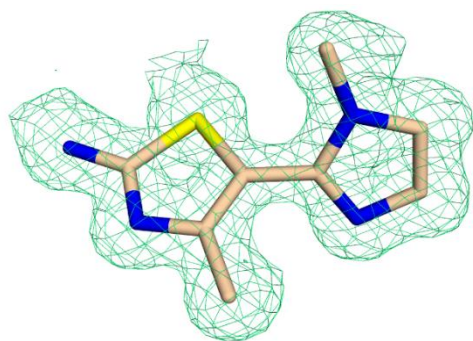
F114



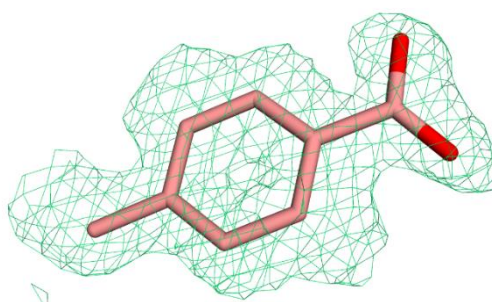
F131



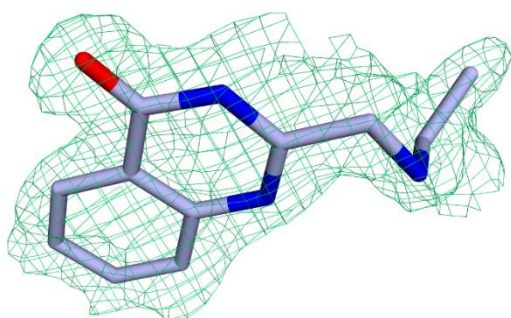
F171



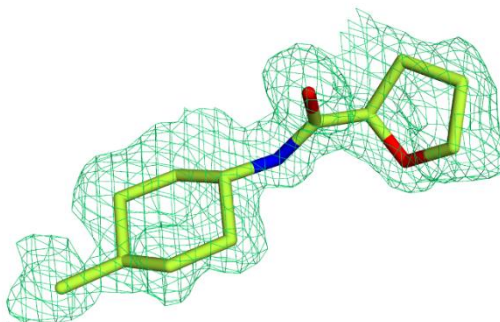
F227



F261



F268





## 6.5.2 Crystallographic Table

**Tab. 6.2:** Data collection and refinement statistics of fragment 211 and 216.

Fragment	211	216
PDB Entry	4YCK	4YCT
<b>Data collection and processing</b>		
Wavelength [Å]	0.91841	0.91841
Beamline	14.1	14.2
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>
<b>Unit cell parameters</b>		
a, b, c [Å]	45.3, 73.0, 52.8	45.2, 73.0, 52.6
$\alpha$ , $\beta$ , $\gamma$ [°]	90.0, 109.7, 90.0	90.0, 109.5, 90.0
<b>Diffraction data</b>		
Resolution range [Å]*	42.7 - 1.07 (1.13 - 1.07)	42.6 - 1.13 (1.20 - 1.13)
Unique reflections	137772 (21639)	114817 (17169)
R(I)sym [%]	4.8 (54.5)	4.8 (45.1)
Completeness [%]	96.7 (94.0)	95.4 (88.5)
Redundancy	3.8 (3.7)	4.1 (3.8)
I/ $\sigma$ (I)	14.2 (2.2)	17.4 (3.0)
<b>Refinement</b>		
Resolution range [Å]	36.5 - 1.07	34.7 - 1.13
Reflections used in refinement (total)	137767	114806
Reflections used in refinement (work/free)	130879/ 6888	109066/ 5740
Final R values for all reflections (work/free) [%]	11.9/ 13.6	13.8/ 15.7
Protein residues	330	328
Atoms Fragment	12	13

Water molecules	302	267
Other ligand atoms	32	27
<b>RMSDs from ideality</b>		
Bond lengths[Å]	0.006	0.006
Bond angles [°]	1.2	1.2
<b>Ramachandran plot</b>		
Residues in most favored regions [%]	93.9	93.5
Residues in additional allowed regions [%]	6.1	6.5
Residues in generously allowed regions [%]	0	0
Residues in disallowed regions [%]	0	0
<b>Mean B-factor [Å<sup>2</sup>]</b>		
Protein total	10.7	11.1
Protein main chain	9.84	10.3
Protein side chain	11.5	11.8
Fragment	12.2	16.3
Other ligand atoms	19.3	27.2
Water molecules	25.6	24.6

\*Highest resolution shell is shown in parenthesis

**Tab. 6.3:** Data collection and refinement statistics of fragment 218 and 224.

<b>Fragment</b>	<b>218</b>	<b>224</b>
<b>PDB Entry</b>	4YCY	4YD3
<b>Data collection and processing</b>		
Wavelength [Å]	0.91841	0.91841
Beamline	14.2	14.1
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>
<b>Unit cell parameters</b>		
a, b, c [Å]	45.6, 72.9, 52.7	45.3, 73.2, 52.7
$\alpha$ , $\beta$ , $\gamma$ [°]	90.0, 108.9, 90.0	90.0, 109.7, 90.0
<b>Diffraction data</b>		
Resolution range [Å]*	19.2 - 1.70 (1.80 - 1.70)	42.7 - 1.25 (1.32 - 1.25)
Unique reflections	35814 (5648)	86779 (13216)
R(I)sym [%]	5.8 (49.1)	3.9 (33.1)
Completeness [%]	99.4 (98.5)	96.5 (91.0)
Redundancy	4.2 (4.2)	3.4 (3.3)
I/ $\sigma$ (I)	19.5 (3.3)	18.9 (3.3)
<b>Refinement</b>		
Resolution range [Å]	19.3 - 1.70	42.7 - 1.37
Reflections used in refinement (total)	35813	86776
Reflections used in refinement (work/free)	34022/ 1791	82437/ 4339
Final R values for all reflections (work/free) [%]	16.6/ 19.8	11.1/ 13.2
Protein residues	323	330
Atoms Fragment	16	34
Water molecules	164	293

Other ligand atoms	19	59
<b>RMSDs from ideality</b>		
Bond lengths[Å]	0.006	0.006
Bond angles [°]	1.0	1.2
<b>Ramachandran plot</b>		
Residues in most favored regions [%]	93.0	94.2
Residues in additional allowed regions [%]	7.0	5.8
Residues in generously allowed regions [%]	0	0
Residues in disallowed regions [%]	0	0
<b>Mean B-factor [Å<sup>2</sup>]</b>		
Protein total	20.4	11.4
Protein main chain	19.7	10.4
Protein side chain	21.4	12.3
Fragment	38.5	15.6
Other ligand atoms	35.8	23.6
Water molecules	27.8	27.0

\*Highest resolution shell is shown in parenthesis

**Tab. 6.4:** Data collection and refinement statistics of fragment 227 and 236.

<b>Fragment</b>	<b>227</b>	<b>236</b>
<b>PDB Entry</b>	4YD4	4YD5
<b>Data collection and processing</b>		
Wavelength [Å]	0.89440	0.91841
Beamline	14.3	14.1
Space group	P2 <sub>1</sub>	P2 <sub>1</sub>
<b>Unit cell parameters</b>		
a, b, c [Å]	45.3, 73.2, 52.7	45.3, 73.0, 52.7
$\alpha$ , $\beta$ , $\gamma$ [°]	90.0, 109.7, 90.0	90.0, 109.3, 90.0
<b>Diffraction data</b>		
Resolution range [Å]*	42.7 - 1.27 (1.35 - 1.27)	42.8 - 1.21 (1.28 - 1.21)
Unique reflections	84492 (13403)	96832 (14290)
R(I)sym [%]	4.8 (48.3)	5.2 (22.4)
Completeness [%]	97.8 (96.5)	98.1 (89.8)
Redundancy	4.2 (4.2)	3.6 (2.8)
I/ $\sigma$ (I)	18.8 (3.2)	13.7 (3.6)
<b>Refinement</b>		
Resolution range [Å]	29.6 - 1.27	42.8 - 1.21
Reflections used in refinement (total)	84482	96831
Reflections used in refinement (work/free)	80260/ 4222	91989/ 4842
Final R values for all reflections (work/free) [%]	12.2/ 14.7	11.9/ 13.6
Protein residues	330	330
Atoms Fragment	10	13
Water molecules	237	226

Other ligand atoms	34	34
<b>RMSDs from ideality</b>		
Bond lengths[Å]	0.007	0.006
Bond angles [°]	1.2	1.2
<b>Ramachandran plot</b>		
Residues in most favored regions [%]	94.2	94.2
Residues in additional allowed regions [%]	5.8	5.8
Residues in generously allowed regions [%]	0	0
Residues in disallowed regions [%]	0	0
<b>Mean B-factor [Å<sup>2</sup>]</b>		
Protein total	11.9	11.6
Protein main chain	11.0	10.6
Protein side chain	12.8	12.5
Fragment	29.0	18.6
Other ligand atoms	36.7	23.6
Water molecules	27.3	25.5

\*Highest resolution shell is shown in parenthesis

## 7 Summary and Outlook (Transglutaminases)

To treat and prevent thromboembolic events like strokes, lung embolism and heart attacks, anticoagulants such as heparin, vitamin K-antagonists (i.e. Markumar and Warfarin, respectively) and the new oral direct thrombin and factor Xa blockers are administered.<sup>[61]</sup> These substances, commonly called blood thinners, reduce the tendency for blood coagulation and thus the thrombotic risk as well. However, such substances also significantly increase the risk of experiencing a life-threatening bleeding event. Moreover, the administration of vitamin K-antagonists requires time-consuming monitoring and results in an extra restriction of quality of life of patients.

An agent blocking the enzyme factor XIII (FXIII) of the blood coagulation cascade could prove to be a superior anticoagulant with a lower bleeding tendency. The advantage of inhibiting FXIII, in contrast to all the other enzymes of the coagulation cascade, is that it does not affect the plasma level of thrombin (FII). Thrombin represents the central enzyme of the blood coagulation catalyzing the activation of fibrinogen to fibrin (*Fig. 3.1*). Subsequently, fibrin polymerizes and forms a blood clot along with other blood components (i.e. erythrocytes and thrombocytes). Apart from the activation of fibrinogen, thrombin is also involved in the activation of the primary hemostasis (aggregation of red blood cells) and even exhibits an antithrombotic impact.<sup>[62]</sup>

FXIII belongs to the family of transglutaminases, mainly catalyzing the cross-linking of proteins by formation of an isopeptide bond between a glutamine and a lysine side chain.<sup>[1]</sup> FXIII represents the last enzyme of the blood coagulation cascade providing final stability to the blood clot by cross-linking of fibrin fibers (*Fig. 3.1*).<sup>[52]</sup> Consequently, inhibition of FXIII would still allow the formation of a weak blood clot, most likely resulting in a lower bleeding risk by using a FXIII blocker compared to other anticoagulants.

The primary goal of the thesis was to obtain a crystal structure of factor XIII in its active conformation (FXIIIa) and subsequently to use the gained information about the chemical composition of the active site for the development of FXIIIa blockers.

In collaboration with the biotech company Zedira GmbH FXIIIa<sup>o</sup> could be crystallized in complex with the covalently-attached inhibitor ZED1301 (*Fig. 2.3*).<sup>[60]</sup> The crystal structure shows that an enormous conformational change occurs during the transition from the inactive to the active state, whereby two  $\beta$ -barrel domains move out of space exposing the active site for inhibitors or substrates (*Fig. 2.8*). The inhibitor ZED1301 binds covalently by its warhead ( $\alpha,\beta$ -unsaturated carboxylic ester) to Cys 314 of the active site. The binding of the ligand induces the formation of a hydrophobic tunnel by rotation of the indole ring of Trp 370 (*Fig. 2.11*). The natural function of this tunnel is most likely the shielding of the intermediately formed thioester from hydrolytic cleavage. The substrate binding site can be subdivided into three areas: The catalytic site, an area located N-terminal named  $\alpha$ -space as well as an area located C-terminal named  $\beta$ -space (*Fig. 2.7*). In the  $\beta$ -space a hydrophobic pocket is formed, presumably induced by the tryptophan indole ring of the inhibitor.

In close collaboration with Zedira the lead compound was optimized in  $\alpha$ -space. A thiazole proved as a promising building block. Starting with the thiazole (*Fig. 3.5*), the hydrophobic pocket in the  $\alpha$ -space could be addressed by the introduction of an ethyl ester (*Fig. 3.9*). This leads to an improvement of the affinity by about one order of magnitude.

The inhibitors developed in the present thesis exhibit a short plasma half-life due to the peptidic chemical structure. Moreover, owing to their high molecular weight the inhibitors are not orally available. Thus, in the future, it should be intended to substitute the part of the inhibitor bridging between the catalytic center and the hydrophobic pocket in the  $\beta$ -space by non-peptidic moiety, in order to prevent degradation by proteases and to reduce molecular weight.

Apart from the usage for the structure-based development of FXIIIa-blockers, the crystal structure of FXIIIa<sup>o</sup> provides mechanistic insights at atomic level that in turn are of fundamental importance for the development of inhibitors for FXIIIa and other human transglutaminases. By means of the crystal structure of FXIIIa<sup>o</sup> it could be understood for the first time how calcium ions are involved in the activation process (*Fig. 2.10*). The coordination of calcium ions induces the rearrangement of loop-regions being parts of the co-substrate binding site



(*Fig. 2.12, Fig. 3.17*). Moreover, the coordination of the calcium ions triggers the formation of a catalytic dyad possessing the function to deprotonate the nitrogen atom of the side chain of the lysine co-substrate (*Fig. 2.2*). In contrast to the inactive conformation, the indole ring of Trp 370 of the hydrophobic tunnel forms a hydrogen bond with the carboxylate oxygen of Gln 400 (*Fig. 3.13*). The carboxamide NH of Gln 400 in turn interacts with the carboxylate oxygen of Glu 434. Also, the conformational reorientation of these amino acids is induced by complexation to the calcium ions (*Fig. 3.14*). Interestingly, both amino acids, the catalytic dyad and the calcium binding sites are highly conserved within the human transglutaminase family (*Chapter 2.6.1, Fig. 3.15*). Moreover, a rotational movement of a three-stranded  $\beta$ -sheet during the transition from the inactive to active state is observed which is assisted by calcium coordination of Asp 351 (*Fig. 2.10c*). The three-stranded  $\beta$ -sheet is part of the  $\beta$ -space. Thus, the extent of the rotation determines the accessible space for ligand binding to the  $\beta$ -space.

Furthermore, FXIIIa<sup>o</sup> could be crystallized with the inhibitor Mi0621 in a collaboration with the academic partner of the laboratory of Prof. Steinmetzer, and this inhibitor differs significantly from the inhibitors of the ZED1301-type (*Fig. 4.3*). Mi0621 contains neither the rigid amino acid proline nor a tryptophan necessary for the occupation of the hydrophobic pocket in the  $\beta$ -space. Instead, Mi0621 possesses an isoleucine residue with an apolar side chain. However, the latter residue is located at the same position as the proline in the other type of inhibitors. The crystal structure of FXIIIa<sup>o</sup> in complex with Mi0621 shows that the hydrophobic pocket in  $\beta$ -space is not shaped and is most likely only induced by the indole ring of the Zedira inhibitors (*Fig. 4.6*). However, the hydrophobic region of the  $\beta$ -space is addressed by the isoleucine side chain. Interestingly, in the crystal structure of FXIIIa<sup>o</sup> in complex with Mi0621 the three-stranded  $\beta$ -sheet adopts an orientation between the inactive state and the crystal structure of FXIIIa<sup>o</sup> in complex with ZED1301 (*Fig. 4.12*).

Including FXIII there are eight human catalytically active transglutaminases (FXIII, TG12-TG7).<sup>[1]</sup> Apart from FXIII, transglutaminase 2 is of particular pharmacological interest and represents a promising target protein for treatment of celiac disease, Alzheimer's disease and diabetic nephropathy (*Fig. 1.3*).<sup>[21]</sup>

In the crystal structure of transglutaminase 2 in complex with a covalent inhibitor (named TG2a\*) published in 2007, the enzyme adopts an active state, as in case of FXIIIa°, however not with a globular but linear conformation (*Fig. 5.3*).<sup>[19]</sup> Based on the new structural and mechanistic insights obtained by the crystal structure of FXIIIa° and further experimental data from the literature, it has been attempted to explain the occurrence of two different active conformations adopted by FXIII and TG2.

First, at closer inspection of the crystal structure of TG2a\*, strand III of the three-stranded  $\beta$ -sheet of the  $\beta$ -barrel 1 domain is displaced, forming a five-stranded  $\beta$ -sheet out of the four-stranded  $\beta$ -sheet of the  $\beta$ -barrel 1 domain and strand II of the original three-stranded  $\beta$ -sheet (*Fig. 5.4*). The displaced strand III now adopts an  $\alpha$ -helical structural element in the region of the  $\beta$ -space which thus is enormously reduced in size. Remarkably, amino acids forming the  $\alpha$ -helix belong to calcium binding site 2 (*Fig. 5.5c*). Consequently, occupancy of calcium binding site 2 does not allow a linear domain arrangement because the  $\beta$ -barrel 1 domain would clash with the three-stranded  $\beta$ -sheet (*Fig. 5.3*). Since only the occupancy of calcium binding site 2 enables the formation of the catalytic dyad (*Fig. 2.10c*), the linear state should not exhibit transamidase activity.

By use of a homology model based on the crystal structure of FXIIIa° (*Fig. 5.12 - Fig. 5.15*), it could be shown that TG2 can adopt a conformation equivalent to FXIII in the active state (FXIIIa°). This is in accordance with the observation that some FXIII inhibitors of the  $\alpha(\text{wh})\text{xxxPW}$ -type have the same affinity against FXIII and TG2. However, due to their size these inhibitors should not be able to bind to the active site of the linear state of TG2 (TG2a\*) since they would clash with the  $\alpha$ -helix in  $\beta$ -space (*Fig. 5.6*). Inhibitors of the  $\alpha(\text{wh})\text{xxxPW}$ -type bind most likely only to the globular active conformation of TG2 as suggested in the present thesis.

Consequently, the inhibitor seems to determine if TG2 adopts the linear or globular active state (*Fig. 5.17*). This is also indicated by FRET measurements performed by the Keillor lab in

Canada showing that, depending on the inhibitor, the enzyme adapts to different conformational states (*Fig. 5.16*).<sup>[113]</sup> One state corresponds to the linear conformation, the other state to a more compact conformation. The measurement of the distances between the N- and the C-terminus of the inactive and the active globular state shows that the distances between the N- and the C-terminus decrease during the transformation to the active conformation. Thus, the compact state found by the FRET experiment corresponds most likely the globular active state.

Transferring these findings to the endogenous biochemical function of transglutaminase 2, the complementarity of substrates binding to the linear or the globular active state determines whether the linear active state is formed and the glutamine residue is hydrolyzed in the catalytic center (deamidation) or the globular active state is formed and the glutamine residue is cross-linked with the lysine residue of a co-substrate by formation of an iso-peptide bond. The hypothesis of the substrate-induced chemoselectivity could give an explanation that simultaneously particular glutamine residues of a protein become deamidated whereas other glutamine residues of the same protein become transamidated.<sup>[7, 114, 115]</sup>

Whether a substrate is deamidated or transamidated can also be affected by the chemical environment (i.e. pH value).<sup>[97]</sup> Analysis of the crystal structures also shows that the calcium concentration might impair the ratio of the linear (TG2a\*) to the globular (TG2a) state, since the occupancy of calcium binding sites presumably does not allow the formation of the linear state. Consequently, with increasing calcium concentration the adoption of the globular state should be preferred. However, at this point it cannot be determined which influencing factor (substrate complementary or calcium concentration) plays the prominent role concerning the chemoselectivity.

Nevertheless, an investigation of the transamidation of FXIII substrates of different sizes by Siebenlist and co-workers shows that the higher the molecular weight the lower the calcium concentration which is required for substrate turnover.<sup>[137]</sup> It appears that the calcium concentration is less mandatory in case of binding of macromolecular substrates to FXIII. This indicates that calcium shifts the equilibrium from the inactive state to a pre-active species where the catalytic center becomes exposed subsequently to a conformational change of the two  $\beta$ -barrel domains. Thus, at high calcium concentrations even low molecular weight substrates with minor affinity might bind to FXIII. At this point it should be mentioned that FXIII exists

in the inactive state as a dimer and calcium might also shift the equilibrium between the dimeric and the monomeric state.

As already described, the substrate induces depending on its structural and chemical constitution the formation of the linear or globular active state. Reversely, also the present conformation of the active TG2 (TG2a\* or TG2a) might determine which inhibitors are potently bound. If the hypothesis that the enzyme adopts the globular active state depending on certain influencing factors turns true, a potential inhibitor of the linear active state could exhibit a strikingly lower affinity in the organism under certain conditions as the *in-vitro* experiment suggested. Therefore, it might be useful to develop both, TG2-blockers active on the linear and the globular state. Consequently, drugs might result for diseases where the deamidase or the transamidase activity is disease relevant.

Ultimately, further experimental investigations are necessary to corroborate the hypothesis that TG2 can adopt a linear deamidation or a globular transamidation state depending on the complementarity of the substrates and the local calcium concentration. In the following, a short outlook is provided in terms of further investigations to confirm the hypothesis made in the present thesis.

Clear evidence for the hypothesis that TG2 can also adopt an active globular state would be provided by a crystal structure of TG2 in complex with an inhibitor of the ZED1301 type. However, for this purpose new crystallization condition will be needed. Also, the FRET experiment with TG2 in complex with TG2-potent inhibitors of the ZED1301 type could provide a valuable insight. In this experiment, the FRET signal should have a similar intensity as in case of inhibitor binding that induce the distance to decrease between the  $\beta$ -sandwich and the  $\beta$ -barrel 2 domain. Also, native gel electrophoresis could give some insights concerning the induced conformation of TG2 depending on the inhibitor.<sup>[19]</sup> In case of a globular state blocker of the ZED1301 type a band at similar position as the inactive state should be observed.

Affinity measurements of TG2 inhibitors concerning a variant of TG2 that does not contain both  $\beta$ -barrel domains could provide an important clue concerning the involvement of the  $\beta$ -barrel 1 domain in terms of the formation of the  $\alpha$ -helix. In contrast to the wildtype, the affinity of a TG2a\* inhibitor should be lower with respect to the  $\beta$ -barrel domain deletion variant. This would mean that the formation of the five-stranded  $\beta$ -sheet out of the four-stranded  $\beta$ -sheet of the  $\beta$ -barrel 1 domain represents a precondition for the formation of the  $\alpha$ -helix out of strand III of the three-stranded  $\beta$ -sheet. A decreased affinity to the  $\beta$ -barrel domain deletion variant should not be observed by an TG2a-potent inhibitor of the ZED1301-type.

The preferred formation of the globular active state during increase of the calcium concentration could be investigated by affinity measurements of a linear state blocker depending on the calcium concentration. The affinity should decrease with increasing calcium concentration. This observation should not be observed with a globular state blocker.

Also, the hypothesis that the linear active state possesses mainly deamidation activity whereas the globular active state exhibits predominantly transamidation activity needs to be investigated.

For this purpose, the ratio of deamidation and transamidation could be analyzed. With increasing calcium concentration, the substrate should increasingly exhibit transamidation activity, as far as the assumption is correct that a preferred formation of the globular active state correlates with enhanced calcium concentration.

Furthermore, the ratio of deamidation and transamidation could be investigated for a TG2 variant that is only composed by the  $\beta$ -sandwich and the catalytic domain. The  $\beta$ -barrel-reduced TG2 variant should have a higher tendency towards transamidation compared to the wildtype, since the four-stranded  $\beta$ -sheet of the  $\beta$ -barrel 1 domain cannot displace strand III of the three-stranded  $\beta$ -sheet. Therefore, the tendency to form a  $\alpha$ -helix would be reduced. Thus, calcium binding site 2 could be occupied more easily and thus the catalytic dyad can be formed.

It would be also interesting to investigate if also other transglutaminases are able to adopt the linear active state in addition to the globular active state or if the linear state is unique for TG2.

Also, the function of the lacking hydrogen bond between the indole ring of Trp 370 (forming the hydrophobic pocket) and the highly conserved amino acid Gln 400 should be investigated. The crystal structure of FXIIIa<sup>o</sup> shows that occupancy of calcium binding site 2 orients Gln 400 in a way to enable formation of a hydrogen bond with the indole ring. Since the analysis of the crystal structures indicate that calcium coordination mainly prepares transglutaminases for transamidation, it can be assumed that the hydrogen bond fixes the indole ring in its “tunnel” position to prevent a premature opening of the hydrophobic tunnel resulting in the hydrolysis of the thioester. This assumption could be investigated by mutating the involving amino acids Gln 400 and Glu 434. Expectably, a mutation should result in a preferred deamidation.

## 8 Zusammenfassung und Ausblick (Transglutaminasen)

Thromboembolischen Ereignisse wie Schlaganfälle, Lungenembolien und Herzinfarkte begegnet man präventiv mit der Verabreichung antikoagulativ wirkender Substanzen wie Heparin, Vitamin-K-Antagonisten (z.B. Markumar) und den neuen oralen direkten Thrombin- und Faktor Xa-Blockern.<sup>[61]</sup> Diese im Volksmund als Blutverdünner bezeichneten Substanzen verringern die Gerinnung (Koagulation) des Blutes und somit auch das Thromboserisiko. Allerdings wird bei der Verabreichung solcher Substanzen auch das Risiko, eine lebensbedrohliche Blutung zu erleiden, beträchtlich erhöht. Zudem ist im Falle der Vitamin-K-Antagonisten eine zeitaufwendige Überwachung und Kontrolle notwendig, was die Lebensqualität der Patienten zusätzlich einschränkt.

Als ein überlegenes Antikoagulans mit einer geringeren Blutungsneigung könnte sich ein Wirkstoff erweisen, der das Enzym Faktor XIII (FXIII) der Blutgerinnungskaskade blockiert. FXIII hat den entscheidenden Vorteil, dass dessen Inhibierung im Gegensatz zu allen anderen Enzymen der Blutgerinnungskaskade den Plasmapegel von Thrombin (FII) nicht beeinträchtigt. Thrombin ist das zentrale Enzym der Blutgerinnung, das die Aktivierung von Fibrinogen zu Fibrin katalysiert, das dann polymerisiert und mit weiteren Bestandteilen des Blutes (wie Erythrozyten und Thrombozyten) das Blutgerinnsel bildet (*Abb. 3.1*). Neben der Aktivierung von Fibrinogen ist Thrombin zudem an der Aktivierung der primären Hämostase (Aggregation der roten Blutkörperchen) beteiligt und weist sogar eine antithrombotische Wirkung auf.<sup>[62]</sup>

FXIII gehört zur Proteinklasse der Transglutaminasen, deren Hauptfunktion die Quervernetzung von Proteinen durch die Ausbildung einer Isopeptidbindung zwischen Glutamin- und Lysinseitenketten darstellt.<sup>[7]</sup> FXIII ist das letzte Enzym der Blutgerinnungskaskade und verleiht durch Quervernetzung der Fibrinfasern dem Blutgerinnsel seine finale Stabilität (*Abb. 3.1*).<sup>[52]</sup> Dies bedeutet letztendlich, dass sich bei der Inhibierung von FXIII noch ein schwaches Blutgerinnsel ausbilden könnte, wodurch das Blutungsrisiko bei einem FXIII-Blocker geringer im Vergleich zu allen anderen Antikoagulantien sein sollte.

Primäres Ziel der vorliegenden Arbeit war es, die Kristallstruktur des Faktors XIII in der aktiven Konformation (FXIIIa) zu erhalten, um anschließend die Information bezüglich der chemischen Beschaffenheit des aktiven Zentrums für die Entwicklung von FXIIIa-Blockern nutzen zu können.

In Kooperation mit dem Biotech-Unternehmen Zedira GmbH konnte FXIIIa<sup>o</sup> im Komplex mit dem kovalent-gebundenen Inhibitor ZED1301 kristallisiert werden (Abb. 2.3).<sup>[60]</sup> Die Kristallstruktur zeigt, dass es beim Übergang vom inaktiven in den aktiven Zustand zu einer enormen konformativen Änderung kommt, wobei die beiden  $\beta$ -Barrel-Domänen zur Seite wegdrehen und so das aktive Zentrum für Inhibitoren bzw. Substrate zugänglich gemacht wird (Abb. 2.8). Der Inhibitor ZED1301 bindet kovalent über seine Kopfgruppe, einem  $\alpha,\beta$ -ungesättigter Carbonsäureester, an Cys 314 des aktiven Zentrums. Die Bindung des Liganden induziert durch Rotation des Indolrings von Trp 370 die Bildung eines hydrophoben Tunnels (Abb. 2.11), dessen natürliche Funktion vermutlich die Abschirmung einer hydrolytischen Spaltung des intermediär gebildeten Thioesters ist. Die Substrat-Bindungsstelle kann in drei Bereiche eingeteilt werden: Dem katalytischen Zentrum, einem Bereich, der sich N-terminal vom katalytischen Zentrum befindet und als  $\alpha$ -Raum bezeichnet wird sowie einem Bereich, der sich C-terminal vom katalytischen Zentrum befindet und als  $\beta$ -Raum bezeichnet wird (Abb. 2.7). Im  $\beta$ -Raum kommt es zur Bildung einer hydrophoben Tasche, die vermutlich durch den Tryptophan-Indolring des Inhibitors induziert wird.

In enger Zusammenarbeit mit Zedira wurde die Leitstruktur im  $\alpha$ -Raum optimiert. Als vielversprechend erwies sich ein Thiazol-Baustein (Abb. 3.5). Von dem ausgehend es gelungen ist, durch die Einführung eines Ethylesters die hydrophobe Tasche im  $\alpha$ -Raum zu adressieren (Abb. 3.9). Hierdurch konnte die Affinität um etwa eine Größenordnung gesteigert werden.

Die bisherigen Inhibitoren weisen aufgrund der peptischen chemischen Struktur eine geringe Plasmahalbwertszeit auf. Zudem sind die bis jetzt zur Verfügung stehenden Inhibitoren aufgrund ihres hohen Molekulargewichts nicht oral applizierbar. Daher sollte in Zukunft versucht werden, den Bereich zwischen katalytischem Zentrum und hydrophober Tasche im  $\beta$ -Raum durch einen nicht-peptidischen Baustein zu ersetzen, um so den Abbau durch Proteasen verhindern und das Molekulargewicht reduzieren zu können.



Neben der Verwendung für die Struktur-basierte Entwicklung eines FXIIIa-Blockers gewährt die Kristallstruktur von FXIIIa<sup>o</sup> mechanistische Einblicke auf molekularer Ebene, die wiederum von fundamentaler Bedeutung für die Entwicklung weiterer Inhibitor von FXIIIa und anderer Transglutaminasen sind. Mit Hilfe der Kristallstruktur von FXIIIa<sup>o</sup> konnte erstmals verstanden werden, wie Kalziumionen in den Aktivierungsprozess involviert sind (*Abb. 2.10*). Durch Koordination an Kalziumionen kommt es zur Neuordnung von Schleifen-Regionen, die Teile der Cosubstrat-Bindungsstelle sind (*Abb. 2.12, Abb. 3.17*). Zudem löst die Koordination der Kalziumionen die Bildung einer katalytischen Diade aus, deren Funktion die Deprotonierung des Stickstoffatoms der Seitenkette des Lysin-Cosubstrats ist (*Abb. 2.2*). Im Vergleich zur inaktiven Konformation bildet der Indolring von Trp 370 des hydrophoben Tunnels eine Wasserstoffbrücke mit dem Carboxamid-Sauerstoff von Gln 400, dessen Carboxamid-NH wiederum mit dem Carboxylat-Sauerstoff von Glu 434 wechselwirkt (*Abb. 3.13*). Auch die konformative Umorientierung dieser beiden Aminosäuren wird durch die Komplexierung an Kalziumionen induziert (*Abb. 3.14*). Interessanterweise sind diese beiden Aminosäuren, sowie die Aminosäuren der katalytische Diade und der Kalziumbindungsstellen hoch konserviert innerhalb der Klasse der menschlichen Transglutaminasen (*Kapitel 2.6.1, Abb. 3.15*). Des Weiteren ist beim Übergang vom inaktiven in den aktiven Zustand eine rotierende Bewegung eines dreisträngigen  $\beta$ -Faltblatts zu beobachten, die durch die Kalzium-Koordination von Asp 351 unterstützt wird (*Abb. 2.10c*). Das dreisträngige  $\beta$ -Faltblatt ist Teil des  $\beta$ -Raums, so dass das Ausmaß der Rotation den für die Bindung eines Liganden zur Verfügung stehenden Platz des  $\beta$ -Raums bestimmt (*Abb. 4.12*).

FXIIIa<sup>o</sup> konnte zudem mit dem Inhibitor Mi0621 des akademischen Kooperationspartners der Arbeitsgruppe von Prof. Steinmetzer kristallisiert werden, der sich signifikant von den Inhibitoren vom ZED1301-Typ unterscheidet (*Abb. 4.3*). Mi0621 enthält weder die rigide Aminosäure Prolin noch ein Tryptophan zur Besetzung der hydrophoben Tasche im  $\beta$ -Raum. Stattdessen besitzt Mi0621 Isoleucin als Aminosäure mit einer unpolaren Seitenkette, die sich allerdings an der Position des Prolins befindet. Bei Betrachtung der Kristallstruktur von FXIIIa<sup>o</sup> im Komplex mit Mi0621 zeigt sich, dass die hydrophobe Tasche im  $\beta$ -Raum nicht ausgebildet ist und vermutlich durch den Indolring der Zedira-Inhibitoren erst induziert wird (*Abb. 4.6*). Jedoch wird der hydrophobe Bereich im  $\beta$ -Raum von der Isoleucin-Seitenkette adressiert. Interessanterweise nimmt das dreisträngige  $\beta$ -Faltblatt von FXIIIa<sup>o</sup> im Komplex mit Mi0621

eine Orientierung zwischen der des inaktiven Zustands und der in der Kristallstruktur von FXIIIa° im Komplex mit ZED1301 an.

Einschließlich FXIII gibt es acht menschliche katalytisch-aktive Transglutaminasen (FXIII, TG1-TG7).<sup>[1]</sup> Neben FXIII ist Transglutaminase 2 von erheblichem pharmakologischem Interesse und stellt ein vielversprechendes Zielprotein zur Behandlung von Zöliakie, Alzheimer und Diabetischer Nephropathie dar (Abb. 1.3).<sup>[2]</sup>

In der in 2007 publizierte Kristallstruktur der Transglutaminase 2 im Komplex mit einem kovalenten Inhibitor (als TG2a\* bezeichnet) nimmt das Enzym im aktiven Zustand keine, wie im Fall von FXIIIa°, globuläre, sondern eine lineare Konformation an (Abb. 5.3).<sup>[19]</sup> Auf Grundlage der neuen strukturellen und mechanistischen Erkenntnisse, die aus der Kristallstruktur von FXIIIa° gewonnen werden konnten und weiteren experimentellen Daten aus der Literatur, wurde versucht, das Auftreten zweier unterschiedlicher aktiver Konformationen bei FXIII und TG2 zu erklären.

Zunächst zeigt sich bei genauerer Betrachtung der Kristallstruktur von TG2a\*, dass Strang III des dreisträngigen  $\beta$ -Faltblatts von der ersten  $\beta$ -Barrel-Domäne verdrängt wird und ein viersträngiges  $\beta$ -Faltblatt der  $\beta$ -Barrel-1-Domäne mit Strang II des ursprünglichen dreisträngigen  $\beta$ -Faltblatts ein fünfsträngiges  $\beta$ -Faltblatt bildet (Abb. 5.4). Der verdrängte Strang III bildet ein  $\alpha$ -helikales Strukturelement im Bereich des  $\beta$ -Raums, wodurch dieser enorm verkleinert wird. Bemerkenswerterweise sind Aminosäuren, aus denen sich die  $\alpha$ -Helix bildet, Teil der Kalzium-Bindungsstelle 2 (Abb. 5.5c). Bei Besetzung der Kalzium-Bindungsstelle 2 ist demnach eine lineare Anordnung der Domänen nicht möglich, da die  $\beta$ -Barrel-1-Domäne mit dem dreisträngigen  $\beta$ -Faltblatt zusammenstoßen würde (Abb. 5.3). Da erst die Besetzung der Kalzium-Bindungsstelle 2 die Ausbildung der katalytischen Diade ermöglicht (Abb. 2.10c), sollte der lineare aktive Zustand der TG2 keine Transamidierungsaktivität besitzen.

Mit einem Homologiemodell von TG2a basierend auf der Kristallstruktur von FXIIIa° (Abb. 5.12 - Abb. 5.15), konnte gezeigt werden, dass TG2 eine zu FXIII in der aktiven Konformation äquivalente Konformation einnehmen kann. Hierfür spricht auch die Beobachtung, dass einige FXIII-Inhibitoren vom  $\alpha(\text{wh})\text{xxxPW}$ -Typ die gleiche Affinität zu FXIII und TG2 besitzen. Diese Inhibitoren sollten jedoch aufgrund ihrer Größe nicht in der

Lage sein, im aktiven Zentrum von TG2 im linearen Zustand (TG2a\*) zu binden, da sie mit der  $\alpha$ -Helix im  $\beta$ -Raum zusammenstoßen würden (*Abb. 5.6*). Inhibitoren vom  $\alpha(\text{wh})\text{xxxPW}$ -Typ binden vermutlich nur an die hier vorgeschlagene globuläre aktive Konformation der TG2.

Dementsprechend scheint der Inhibitor zu bestimmen, ob TG2 den linearen oder globulären aktiven Zustand annimmt (*Abb. 5.17*). Hierauf deuten auch FRET-Messungen der Arbeitsgruppe Keillor aus Kanada hin, die zeigen, dass je nach gebundenem Inhibitor das Enzym zwei Zustände annehmen kann (*Abb. 5.16*).<sup>[113]</sup> Der eine Zustand entspricht der linearen Konformation, bei dem zweiten Zustand nimmt das Enzym nach den Autoren einen kompakteren Zustand an. Die Messung der Abstände zwischen N- und C-Terminus der inaktiven Konformation sowie der globulären aktiven Konformation zeigt, dass die Distanz zwischen N- und C-Terminus beim Übergang in die globuläre aktive Konformation abnimmt und der kompaktere Zustand vermutlich der globulären aktiven Konformation entspricht.

Überträgt man diese Erkenntnisse auf die natürliche biochemische Funktion der Transglutaminase 2, so könnte die Komplementarität des Substrates zur linearen oder globulären aktiven Konformation darüber entscheiden, ob der lineare aktive Zustand gebildet und der Glutamin-Rest im katalytischem Zentrum somit hydrolysiert wird (Deamidierung) oder ob der globuläre aktive Zustand gebildet und der entsprechende Glutamin-Rest mit einem Lys-Rest eines Cosubstrats durch Ausbildung einer Isopeptidbindung quervernetzt wird (Transamidierung). Die Hypothese der Substrat-induzierten Chemoselektivität könnte eine Erklärung dafür liefern, dass simultan (also unter gleichen Bedingungen) bestimmte Glutamine eines Proteins zu Glutamat deamidiert werden, wohingegen andere Glutamine desselben Proteins transamidiert werden.<sup>[7, 114, 115]</sup>

Ob ein Substrat deamidiert oder transamidiert wird, kann zudem auch von der chemischen Umgebung (wie dem pH-Wert) beeinträchtigt werden.<sup>[97]</sup> Nach Analyse der Kristallstrukturen könnte auch die Kalziumkonzentration das Verhältnis von linearer (TG2a\*) und globulärer (TG2a) Form beeinträchtigen, da bei Besetzung der Kalzium-Bindungsstelle 2 die Bildung des linearen Zustands vermutlich nicht möglich wäre. Somit müsste es mit steigender Kalzium-Konzentration zu einer zunehmend bevorzugten Bildung des globulären aktiven Zustands kommen. Eine Aussage darüber, welcher der beiden Einflussfaktoren (Substrat-Komplementarität oder Kalziumkonzentration) entscheidend für die Chemoselektivität ist, lässt sich jedoch nicht treffen.

Eine Untersuchung der Transamidierung von FXIII-Substraten unterschiedlicher Größe von Siebenlist und Mitarbeitern zeigt jedoch, dass mit zunehmenden Molekulargewicht, eine geringere Kalziumkonzentration für den Substratumsatz nötig ist.<sup>[137]</sup> Demnach scheint bei makromolekularen Substraten die Kalziumkonzentration für die Bindung an FXIII von geringerer Bedeutung zu sein. Dies ist ein Indiz dafür, dass Kalzium das Gleichgewicht vom inaktiven Zustand zu einer prä-aktiven Spezies, bei der durch konformative Änderung der beiden  $\beta$ -Barrel-Domänen das aktive Zentrum exponiert wird, verschiebt. Bei hoher Kalziumkonzentration könnten somit auch niedermolekulare Substrate mit geringerer Affinität an FXIII binden. An dieser Stelle sei bemerkt, dass FXIII im inaktiven Zustand als Dimer vorliegt und somit Kalzium möglicherweise auch das Dimer-Monomer-Gleichgewicht verschieben könnte.

Wie bereits erwähnt wurde, könnte das Substrat je nach struktureller und chemischer Beschaffenheit die Bildung des linearen oder globulären aktiven Zustands induzieren. Umgekehrt, könnte somit auch die vorliegende Konformation der aktiven TG2 (TG2a\* oder TG2a) festlegen, welche Inhibitoren affin gebunden werden. Sollte das Enzym in Abhängigkeit bestimmter Einflussfaktoren bevorzugt den globulären Zustand annehmen, so könnte ein potentieller Inhibitor des linearen aktiven Zustands im Organismus unter bestimmten Bedingungen eine deutlich geringere Affinität besitzen als es das in-vitro Experiment suggeriert hat. Demnach wäre es sinnvoll, sowohl TG2-Blocker des linearen als auch des globulären Zustands zu entwickeln. Somit hätte man Wirkstoffe an der Hand, die bei Krankheiten eingesetzt werden können, bei denen entweder die Deamidaseaktivität oder Transamidaseaktivität von pathogener Relevanz ist, wodurch letztendlich die Nebenwirkungen minimiert werden könnten.

Letztendlich sind weitere experimentelle Untersuchungen notwendig, um die Hypothese zu bekräftigen, dass TG2 in Abhängigkeit von der Komplementarität des Substrats und der lokalen Kalziumkonzentration eine linearen Deamidierungszustand oder einen globulären Transamidierungszustand annehmen kann. Im Folgenden soll ein kurzer Ausblick bezüglich weiterer Untersuchungen gegeben werden, die die in der vorliegenden Arbeit aufgestellte Hypothese bestätigen könnte.

Die Annahme, dass TG2 neben der linearen auch in der globulären Form vorliegen kann, muss durch weitere Experimente bestätigt werden.

Ein klarer Beweis wäre die Kristallisation von TG2 im Komplex mit einem Inhibitor vom ZED1301-Typ. In der entsprechenden Kristallstruktur sollte TG2 die globuläre Konformation annehmen. Hierzu müsste jedoch eine neue Kristallisationsbedingung gefunden werden. Allerdings könnte die Durchführung des FRET-Experiments mit TG2-affinen Inhibitoren vom ZED1301-Typ ebenfalls einen Hinweis auf die Existenz des globulären aktiven Zustands liefern. Bei diesem Experiment sollte das FRET-Signal die gleiche Intensität haben, wie bei den Inhibitoren, bei denen nach dem FRET-Experiment der Abstand zwischen  $\beta$ -Sandwich und  $\beta$ -Barrel-2-Domäne abgenommen hat.

Auch die native Gelelektrophorese könnte Aufschluss über die induzierte Konformation von TG2 in Abhängigkeit vom Inhibitor geben.<sup>[19]</sup> Bei einem *Globular-State Blocker* vom ZED1301-Typ sollte, im Gegensatz zu einem *linear-State Blocker*, eine Bande auf Höhe der inaktiven Form zu sehen sein.

Die Messung der Affinität der TG2-Inhibitoren bezüglich einer Variante der TG2, die nicht die beiden  $\beta$ -Barreldomänen enthält, könnte ein wichtiges Indiz für die Beteiligung der  $\beta$ -Barrel-1-Domäne an der Bildung der  $\alpha$ -Helix im aktiven Zentrum liefern. Die Affinität eines TG2a\*-Inhibitors sollte bezüglich der um die  $\beta$ -Barreldomänen-verkürzten TG2-Variante geringer im Vergleich zum Wildtyp sein. Dies würde bedeuten, dass die Ausbildung des fünfsträngigen  $\beta$ -Faltblattes aus dem viersträngigen  $\beta$ -Faltblatt der  $\beta$ -Barrel-1-Domäne eine Voraussetzung für die Bildung der  $\alpha$ -Helix aus Strang III des dreisträngigen  $\beta$ -Faltblattes wäre. Eine verringerte Affinität gegenüber der  $\beta$ -Barreldomänen-verkürzten TG2-Variante sollte hingegen bei einem TG2-affinen Inhibitor vom ZED1301-Typ nicht zu beobachten sein.

Die bevorzugte Bildung der globulären aktiven Konformation bei Zunahme der Kalziumkonzentration könnte mit Hilfe von Affinitätsmessungen eines *Linear-State Blockers* in Abhängigkeit von der Kalzium-Konzentration untersucht werden. Mit steigender Kalzium-Konzentration sollte die Affinität abnehmen. Diese Beobachtung sollte man bei einem *Globular-State Blocker* nicht machen.

Auch muss die Hypothese, dass der lineare aktive Zustand überwiegend Deamidierungsaktivität und der globuläre aktive Zustand vermehrt Transamidierungsaktivität besitzt, experimentell untersucht werden.

Hierzu könnte man das Deamidierungs-/Transamidierungsverhältnis in Abhängigkeit von der Kalzium-Konzentration untersuchen. Mit zunehmender Kalzium-Konzentration sollte es vermehrt zur Transamidierung kommen, sofern sich eine bevorzugte Bildung des globulären Zustands ab einer bestimmten Kalzium-Konzentration bewahrheitet.

Des Weiteren könnte man das Deamidierungs-/Transamidierungsverhalten der TG2-Variante untersuchen, die nur aus der  $\beta$ -*Sandwich*-Domäne und der katalytischen Domäne besteht. Die  $\beta$ -*Barreldomänen*-verkürzten TG2-Variante sollte im Vergleich zum Wildtyp eine höhere Tendenz zur Transamidierung aufweisen, da das viersträngige  $\beta$ -Faltblatt der  $\beta$ -*Barrel*-1-Domäne den Strang III des dreisträngigen  $\beta$ -Faltblattes nicht verdrängen kann. Dementsprechend ist die Tendenz zur Bildung der daraus resultierenden  $\alpha$ -Helix geringer. Als Konsequenz kann Kalzium-Bindungsstelle 2 leichter besetzt werden und sich somit die katalytische Diade ausbilden.

Es wäre des Weiteren interessant zu untersuchen, ob auch andere Transglutaminasen in der Lage sind, neben dem globulären Zustand den linearen Zustand einzunehmen oder ob der lineare Zustand einzigartig für TG2 ist.

Auch die Funktion der in der Kristallstruktur von TG2a\* nicht ausgebildeten Wasserstoffbrücke zwischen dem Indolring von Trp 370, der den hydrophoben Tunnel bildet, und der hochkonservierten Aminosäure Gln 400 sollte experimentell untersucht werden.

Nach der Kristallstruktur von FXIIIa<sup>o</sup> wird durch die Besetzung der Kalzium-Bindungsstelle 2 Gln 400 derart orientiert, dass die Wasserstoffbrücke zu dem Indolring ausgebildet werden kann. Da die Analyse der Kristallstrukturen zeigen, dass die Kalzium-Koordination

Transglutaminasen bezüglich der Transamidierung vorbereiten, liegt die Vermutung nahe, dass die Wasserstoffbrücke den Indolring in seiner „Tunnel“-Position fixiert, um eine verfrühte Öffnung des hydrophoben Tunnels und die damit verbundene Hydrolyse des Thioesters zu verhindern. Diese Annahme könnte durch Mutationsstudien der beiden involvierten Aminosäuren Gln 400 und Glu 434 untersucht werden. Erwartungsgemäß sollte es bei einer Mutation bevorzugt zur Deamidierung kommen.

## References

- [1] L. Lorand, R. M. Graham, *Transglutaminases: Crosslinking enzymes with pleiotropic functions*. Nat Rev Mol Cell Bio **4**, 140-156 (2003).
- [2] D. Aeschlimann, V. Thomazy, *Protein crosslinking in assembly and remodelling of extracellular matrices: The role of transglutaminases*. Connect Tissue Res **41**, 1-+ (2000).
- [3] E. Candi, S. Oddi, A. Paradisi, A. Terrinoni, M. Ranalli, P. Teofoli, G. Citro, S. Scarpato, P. Puddu, G. Melino, *Expression of transglutaminase 5 in normal and pathologic human epidermis*. J Invest Dermatol **119**, 670-677 (2002).
- [4] E. Candi, S. Oddi, A. Terrinoni, A. Paradisi, M. Ranalli, A. Finazzi-Agro, G. Melino, *Transglutaminase 5 cross-links loricrin, involucrin, and small proline-rich proteins in vitro*. J Biol Chem **276**, 35014-35023 (2001).
- [5] S. Y. Kim, S. I. Chung, K. Yoneda, P. M. Steinert, *Expression of Transglutaminase-1 in Human Epidermis*. J Invest Dermatol **104**, 211-217 (1995).
- [6] K. Laki, L. Lorand, *On the Solubility of Fibrin Clots*. Science **108**, 280-280 (1948).
- [7] S. Gundemir, G. Colak, J. Tucholski, G. V. W. Johnson, *Transglutaminase 2: A molecular Swiss army knife*. Bba-Mol Cell Res **1823**, 406-419 (2012).
- [8] J. S. K. Chen, K. Mehta, *Tissue transglutaminase: an enzyme with a split personality*. Int J Biochem Cell B **31**, 817-836 (1999).
- [9] S. Mishra, L. J. Murphy, *Tissue transglutaminase has intrinsic kinase activity - Identification of transglutaminase 2 as an insulin-like growth factor-binding protein-3 kinase*. J Biol Chem **279**, 23863-23868 (2004).
- [10] S. Mishra, A. Saleh, P. S. Espino, J. R. Davie, L. J. Murphy, *Phosphorylation of histones by tissue transglutaminase*. J Biol Chem **281**, 5532-5538 (2006).
- [11] G. Hasegawa, M. Suwa, Y. Ichikawa, T. Ohtsuka, S. Kumagai, M. Kikuchi, Y. Sato, Y. Saito, *A novel function of tissue-type transglutaminase: protein disulphide isomerase*. Biochem J **373**, 793-803 (2003).
- [12] P. G. Mastroberardino, M. G. Farrace, I. Viti, F. Pavone, G. M. Fimia, G. Melino, C. Rodolfo, M. Piacentini, *"Tissue" transglutaminase contributes to the formation of disulphide bridges in proteins of mitochondrial respiratory complexes*. Bba-Bioenergetics **1757**, 1357-1365 (2006).
- [13] G. E. Begg, L. Carrington, P. H. Stokes, J. M. Matthews, M. A. Wouters, A. Husain, L. Lorand, S. E. Iismaa, R. M. Graham, *Mechanism of allosteric regulation of transglutaminase 2 by GTP*. P Natl Acad Sci USA **103**, 19683-19688 (2006).
- [14] J. W. Zhang, M. Lesort, R. P. Guttman, G. V. W. Johnson, *Modulation of the in situ activity of tissue transglutaminase by calcium and GTP*. J Biol Chem **273**, 2288-2295 (1998).
- [15] X. Jin, J. Stamnaes, C. Klock, T. R. DiRaimondo, L. M. Sollid, C. Khosla, *Activation of Extracellular Transglutaminase 2 by Thioredoxin*. J Biol Chem **286**, 37866-37873 (2011).
- [16] J. Stamnaes, D. M. Pinkas, B. Fleckenstein, C. Khosla, L. M. Sollid, *Redox Regulation of Transglutaminase 2 Activity*. J Biol Chem **285**, 25402-25409 (2010).
- [17] V. C. Yee, L. C. Pedersen, I. Letrong, P. D. Bishop, R. E. Stenkamp, D. C. Teller, *3-Dimensional Structure of a Transglutaminase - Human Blood-Coagulation Factor-Xiii*. P Natl Acad Sci USA **91**, 7296-7300 (1994).
- [18] B. Ahvazi, H. C. Kim, S. H. Kee, Z. Nemes, P. M. Steinert, *Three-dimensional structure of the human transglutaminase 3 enzyme: binding of calcium ions changes structure for activation*. Embo J **21**, 2055-2067 (2002).
- [19] D. M. Pinkas, P. Strop, A. T. Brunger, C. Khosla, *Transglutaminase 2 undergoes a large conformational change upon activation*. Plos Biol **5**, 2788-2796 (2007).



- [20] J. W. Keillor, C. M. Clouthier, K. Y. P. Apperley, A. Akbar, A. Mulani, *Acyl transfer mechanisms of tissue transglutaminase*. *Bioorg Chem* **57**, 186-197 (2014).
- [21] J. M. Wodzinska, *Transglutaminases as targets for pharmacological inhibition*. *Mini-Rev Med Chem* **5**, 279-292 (2005).
- [22] L. Lorand, *Factor XIII and the clotting of fibrinogen: from basic research to medicine*. *J Thromb Haemost* **3**, 1337-1348 (2005).
- [23] L. Lorand, P. T. Velasco, S. N. P. Murthy, P. Lefebvre, D. Green, *Autoimmune antibody in a hemorrhagic patient interacts with thrombin-activated factor XIII in a unique manner*. *Blood* **93**, 909-917 (1999).
- [24] A. M. Sulic, K. Kurppa, T. Rauhavirta, K. Kaukinen, K. Lindfors, *Transglutaminase as a therapeutic target for celiac disease*. *Expert Opin Ther Tar* **19**, 335-348 (2015).
- [25] W. Dieterich, T. Ehnis, M. Bauer, P. Donner, U. Volta, E. O. Riecken, D. Schuppan, *Identification of tissue transglutaminase as the autoantigen of celiac disease*. *Nat Med* **3**, 797-801 (1997).
- [26] M. Lesort, J. Tucholski, M. L. Miller, G. V. W. Johnson, *Tissue transglutaminase: a possible role in neurodegenerative diseases*. *Prog Neurobiol* **61**, 439-463 (2000).
- [27] M. V. Karpuj, H. Garren, H. Slunt, D. L. Price, J. Gusella, M. W. Becher, L. Steinman, *Transglutaminase aggregates huntingtin into nonamyloidogenic polymers, and its enzymatic activity increases in Huntington's disease brain nuclei*. *P Natl Acad Sci USA* **96**, 7388-7393 (1999).
- [28] D. J. Selkoe, C. Abraham, Y. Ihara, *Brain Transglutaminase - Invitro Crosslinking of Human Neurofilament Proteins into Insoluble Polymers*. *P Natl Acad Sci-Biol* **79**, 6070-6074 (1982).
- [29] L. H. Huang, J. L. Haylor, Z. Hau, R. A. Jones, M. E. Vickers, B. Wagner, M. Griffin, R. E. Saint, I. G. C. Coutts, A. M. El Nahas, T. S. Johnson, *Transglutaminase inhibition ameliorates experimental diabetic nephropathy*. *Kidney Int* **76**, 383-394 (2009).
- [30] R. L. Eckert, M. T. Kaartinen, M. Nurminskaya, A. M. Belkin, G. Colak, G. V. W. Johnson, K. Mehta, *Transglutaminase Regulation of Cell Function*. *Physiol Rev* **94**, 383-417 (2014).
- [31] J. Sohn, T. I. Kim, Y. H. Yoon, J. Y. Kim, S. Y. Kim, *Novel transglutaminase inhibitors reverse the inflammation of allergic conjunctivitis*. *J Clin Invest* **111**, 121-128 (2003).
- [32] L. Huang, A. M. Xu, W. Liu, *Transglutaminase 2 in cancer*. *Am J Cancer Res* **5**, 2756-2776 (2015).
- [33] A. J. Cassidy, M. A. M. van Steensel, P. M. Steijlen, M. van Geel, J. van der Velden, S. M. Morley, A. Terrinoni, G. Melino, E. Candi, W. H. I. McLean, *A homozygous missense mutation in TGM5 abolishes epidermal transglutaminase 5 activity and causes acral peeling skin syndrome*. *Am J Hum Genet* **77**, 909-917 (2005).
- [34] W. J. Guan, K. D. Xia, Y. T. Ma, Y. T. Liu, Y. T. Shi, H. Jiang, L. Shen, K. Xia, J. D. Li, B. S. Tang, J. L. Wang, *Transglutaminase 6 interacts with polyQ proteins and promotes the formation of polyQ aggregates*. *Biochem Bioph Res Co* **437**, 94-100 (2013).
- [35] M. Hadjivassiliou, P. Aeschlimann, A. Strigun, D. S. Sanders, N. Woodroffe, D. Aeschlimann, *Autoantibodies in gluten ataxia recognize a novel neuronal transglutaminase*. *Ann Neurol* **64**, 332-343 (2008).
- [36] W. G. Jiang, R. J. Albin, A. Douglas-Jones, T. E. Mansel, *Expression of transglutaminases in human breast cancer and their possible clinical significance*. *Brit J Cancer* **88**, S46-S46 (2003).
- [37] L. L. Pan, Y. M. Huang, M. Wang, X. E. Zhuang, D. F. Luo, S. C. Guo, Z. S. Zhang, Q. Huang, S. L. Lin, S. Y. Wang, *Positional cloning and next-generation sequencing identified a TGM6 mutation in a large Chinese pedigree with acute myeloid leukaemia*. *Eur J Hum Genet* **23**, 218-223 (2015).
- [38] J. L. Wang, X. Yang, K. Xia, Z. M. Hu, L. Weng, X. Jin, H. Jiang, P. Zhang, L. Shen, J. F. Guo, N. Li, Y. R. Li, L. F. Lei, J. Zhou, J. A. Du, Y. F. Zhou, Q. A. Pan, J. A. Wang, J. Wang, R. Q. Li, B. S. Tang, *TGM6 identified as a novel causative gene of spinocerebellar ataxias using exome sequencing*. *Brain* **133**, 3510-3518 (2010).

- [39] W. G. Jiang, L. Ye, A. J. Sanders, F. Ruge, H. G. Kynaston, R. J. Ablin, M. D. Mason, *Prostate transglutaminase (TGase-4, TGaseP) enhances the adhesion of prostate cancer cells to extracellular matrix, the potential role of TGase-core domain*. J Transl Med **11**, (2013).
- [40] <https://de.dreamstime.com/stockfoto-menschliche-anatomie-image29075830>.
- [41] L. Muszbek, Z. Bereczky, Z. Bagoly, I. Komaromi, E. Katona, *Factor Xiii: A Coagulation Factor with Multiple Plasmatic and Cellular Functions*. Physiol Rev **91**, 931-972 (2011).
- [42] L. Lorand, J. M. Jeong, J. T. Radek, J. Wilson, *Human Plasma Factor-Xiii - Subunit Interactions and Activation of Zymogen (Reprinted from Biochemistry, Vol 32, Pg 3527-3534, 1993)*. Proteolytic Enzymes in Coagulation, Fibrinolysis, and Complement Activation, Part A **222**, 22-35 (1993).
- [43] J. T. Radek, J. M. Jeong, J. Wilson, L. Lorand, *Association of the  $\alpha$ -Subunits of Recombinant Placental Factor-Xiii with the Native Carrier B-Subunits from Human Plasma*. Biochemistry-US **32**, 3527-3534 (1993).
- [44] T. J. Hornyak, J. A. Shafer, *Role of Calcium-Ion in the Generation of Factor-Xiii Activity*. Biochemistry-US **30**, 6175-6182 (1991).
- [45] S. D. Lewis, T. J. Janus, L. Lorand, J. A. Shafer, *Regulation of Formation of Factor-Xiiia by Its Fibrin Substrates*. Biochemistry-US **24**, 6772-6777 (1985).
- [46] C. S. Greenberg, C. C. Miraglia, *The Effect of Fibrin Polymers on Thrombin-Catalyzed Plasma Factor-Xiiia Formation*. Blood **66**, 466-469 (1985).
- [47] T. J. Janus, S. D. Lewis, L. Lorand, J. A. Shafer, *Promotion of Thrombin-Catalyzed Activation of Factor-Xiii by Fibrinogen*. Biochemistry-US **22**, 6269-6272 (1983).
- [48] E. F. Luscher, *[Fibrin-stabilizing factor from thrombocytes]*. Schweizerische medizinische Wochenschrift **87**, 1220-1221 (1957).
- [49] L. Muszbek, R. Adany, G. Szegedi, J. Polgar, M. Kawai, *Factor-Xiii of Blood-Coagulation in Human-Monocytes*. Thromb Res **37**, 401-410 (1985).
- [50] L. Muszbek, G. Haramura, J. Polgar, *Transformation of Cellular Factor-Xiii into an Active Zymogen Transglutaminase in Thrombin-Stimulated Platelets*. Thromb Haemostasis **73**, 702-705 (1995).
- [51] L. Muszbek, J. Polgar, Z. Boda, *Platelet Factor-Xiii Becomes Active without the Release of Activation Peptide during Platelet Activation*. Thromb Haemostasis **69**, 282-285 (1993).
- [52] L. Muszbek, V. C. Yee, Z. Hevessy, *Blood coagulation factor XIII: Structure and function*. Thromb Res **94**, 271-305 (1999).
- [53] C. L. Nikolajsen, T. F. Dyrland, E. T. Poulsen, J. J. Enghild, C. Scavenius, *Coagulation factor XIIIa substrates in human plasma: identification and incorporation into the clot*. J Biol Chem **289**, 6526-6534 (2014).
- [54] A. Inbal, L. Muszbek, *Coagulation factor deficiencies and pregnancy loss*. Semin Thromb Hemost **29**, 171-174 (2003).
- [55] A. Inbal, A. Lubetsky, T. Krapp, D. Castel, A. Shaish, G. Dickneite, L. Modis, L. Muszbek, A. Inbal, *Impaired wound healing in factor XIII deficient mice*. Thromb Haemostasis **94**, 432-437 (2005).
- [56] R. Dardik, J. Loscalzo, A. Inbal, *Factor XIII (FXIII) and angiogenesis*. J Thromb Haemost **4**, 19-25 (2006).
- [57] G. Klebe, *Virtual ligand screening: strategies, perspectives and limitations*. Drug discovery today **11**, 580-594 (2006).
- [58] R. Macarron, M. N. Banks, D. Bojanic, D. J. Burns, D. A. Cirovic, T. Garyantes, D. V. Green, R. P. Hertzberg, W. P. Janzen, J. W. Paslay, U. Schopfer, G. S. Sittampalam, *Impact of high-throughput screening in biomedical research*. Nat Rev Drug Discov **10**, 188-195 (2011).
- [59] P. J. Hajduk, J. Greer, *A decade of fragment-based drug design: strategic advances and lessons learned*. Nat Rev Drug Discov **6**, 211-219 (2007).

- [60] M. Stieler, J. Weber, M. Hils, P. Kolb, A. Heine, C. Buchold, R. Pasternack, G. Klebe, *Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants*. *Angew Chem Int Edit* **52**, 11930-11934 (2013).
- [61] D. Garcia, E. Libby, M. A. Crowther, *The new oral anticoagulants*. *Blood* **115**, 15-20 (2010).
- [62] J. H. Griffin, *Blood coagulation. The thrombin paradox*. *Nature* **378**, 337-338 (1995).
- [63] I. Komaromi, Z. Bagoly, L. Muszbek, *Factor XIII: novel structural and functional aspects*. *J Thromb Haemost* **9**, 9-20 (2011).
- [64] V. Schroeder, H. P. Kohler, *New developments in the area of factor XIII*. *J Thromb Haemost* **11**, 234-244 (2013).
- [65] S. Finney, L. Seale, R. T. Sawyer, R. B. Wallis, *Tridegin, a new peptidic inhibitor of factor XIIIa, from the blood-sucking leech Haementeria ghilianii*. *Biochem J* **324**, 797-805 (1997).
- [66] R. J. Shebuski, G. R. Sitko, D. A. Claremon, J. J. Baldwin, D. C. Remy, A. M. Stern, *Inhibition of Factor-XIIa in a Canine Model of Coronary-Thrombosis - Effect on Reperfusion and Acute Reocclusion after Recombinant Tissue-Type Plasminogen-Activator*. *Blood* **75**, 1455-1459 (1990).
- [67] L. Lorand, R. B. Credo, T. J. Janus, *Factor-Xiii (Fibrin-Stabilizing Factor)*. *Methods in enzymology* **80**, 333-341 (1981).
- [68] Y. Sugimura, M. Hosono, F. Wada, T. Yoshimura, M. Maki, K. Hitomi, *Screening for the preferred substrate sequence of transglutaminase using a phage-displayed peptide library - Identification of peptide substrates for TGase 2 and Factor XIIIa*. *J Biol Chem* **281**, 17699-17706 (2006).
- [69] L. Muszbek, R. A. Ariens, A. Ichinose, *Factor XIII: recommended terms and abbreviations*. *J Thromb Haemost* **5**, 181-183 (2007).
- [70] M. Andersen, J. Faber, *Conformational changes of recombinant FXIII upon both proteolytic and non-proteolytic activation studied by HX-MS*. *Haemophilia* **16**, 42-42 (2010).
- [71] R. T. Woofter, M. C. Maurer, *Role of calcium in the conformational dynamics of factor XIII activation examined by hydrogen-deuterium exchange coupled with MALDI-TOF MS*. *Arch Biochem Biophys* **512**, 87-95 (2011).
- [72] A. Mary, K. E. Achyuthan, C. S. Greenberg, *The Binding of Divalent Metal-Ions to Platelet Factor-Xiii Modulates Its Proteolysis by Trypsin and Thrombin*. *Arch Biochem Biophys* **261**, 112-121 (1988).
- [73] M. S. Weiss, H. J. Metzner, R. Hilgenfeld, *Two non-proline cis peptide bonds may be important for factor XIII function*. *Febs Lett* **423**, 291-296 (1998).
- [74] B. A. Fox, V. C. Yee, L. C. Pedersen, I. Le Trong, P. D. Bishop, R. E. Stenkamp, D. C. Teller, *Identification of the calcium binding site and a novel ytterbium site in blood coagulation factor XIII by X-ray crystallography*. *J Biol Chem* **274**, 4917-4923 (1999).
- [75] B. Ahvazi, K. M. Boeshans, W. Idler, U. Baxa, P. M. Steinert, *Roles of calcium ions in the activation and activity of the transglutaminase 3 enzyme*. *J Biol Chem* **278**, 23834-23841 (2003).
- [76] J. M. Hettasch, C. S. Greenberg, *Analysis of the catalytic activity of human factor XIIIa by site-directed mutagenesis*. *J Biol Chem* **269**, 28309-28313 (1994).
- [77] A. Heil, J. Weber, C. Buchold, R. Pasternack, M. Hils, *Differences in the inhibition of coagulation factor XIII-A from animal species revealed by Michael Acceptor- and thioimidazol based blockers*. *Thromb Res* **131**, e214-222 (2013).
- [78] R. B. Credo, C. G. Curtis, L. Lorand, *Ca<sup>2+</sup>-Related Regulatory Function of Fibrinogen*. *P Natl Acad Sci USA* **75**, 4234-4237 (1978).
- [79] U. Mueller, N. Darowski, M. R. Fuchs, R. Forster, M. Hellmig, K. S. Paithankar, S. Puhlinger, M. Steffien, G. Zocher, M. S. Weiss, *Facilities for macromolecular crystallography at the Helmholtz-Zentrum Berlin*. *J Synchrotron Radiat* **19**, 442-449 (2012).

- [80] Z. Otwinowski, W. Minor, *Processing of X-ray diffraction data collected in oscillation mode*. Method Enzymol **276**, 307-326 (1997).
- [81] *XPREP Software for Data Preparation & Reciprocal Space Exploration, Version 2013/1, Bruker AXS Inc., Madison, WI, USA.*
- [82] S. Bailey, *The Ccp4 Suite - Programs for Protein Crystallography*. Acta Crystallogr D **50**, 760-763 (1994).
- [83] A. J. McCoy, R. W. Grosse-Kunstleve, P. D. Adams, M. D. Winn, L. C. Storoni, R. J. Read, *Phaser crystallographic software*. J Appl Crystallogr **40**, 658-674 (2007).
- [84] A. Sali, T. L. Blundell, *Comparative Protein Modeling by Satisfaction of Spatial Restraints*. J Mol Biol **234**, 779-815 (1993).
- [85] J. M. Pei, N. V. Grishin, *PROMALS: towards accurate multiple sequence alignments of distantly related proteins*. Bioinformatics **23**, 802-808 (2007).
- [86] G. Langer, S. X. Cohen, V. S. Lamzin, A. Perrakis, *Automated macromolecular model building for X-ray crystallography using ARP/wARP version 7*. Nat Protoc **3**, 1171-1179 (2008).
- [87] P. Emsley, B. Lohkamp, W. G. Scott, K. Cowtan, *Features and development of Coot*. Acta Crystallogr D **66**, 486-501 (2010).
- [88] P. D. Adams, P. V. Afonine, G. Bunkoczi, V. B. Chen, I. W. Davis, N. Echols, J. J. Headd, L. W. Hung, G. J. Kapral, R. W. Grosse-Kunstleve, A. J. McCoy, N. W. Moriarty, R. Oeffner, R. J. Read, D. C. Richardson, J. S. Richardson, T. C. Terwilliger, P. H. Zwart, *PHENIX: a comprehensive Python-based system for macromolecular structure solution*. Acta Crystallogr D **66**, 213-221 (2010).
- [89] *SYBYL-X 2.0, Tripos International, 1699 South Hanley Rd., St. Louis, Missouri, 63144, USA*
- [90] R. A. Laskowski, M. W. Macarthur, D. S. Moss, J. M. Thornton, *Procheck - a Program to Check the Stereochemical Quality of Protein Structures*. J Appl Crystallogr **26**, 283-291 (1993).
- [91] F. A. Momany, R. Rone, *Validation of the General-Purpose Quanta(R)3.2/Charmm(R) Force-Field*. J Comput Chem **13**, 888-900 (1992).
- [92] <http://www.who.int/mediacentre/factsheets/fs310/en/>.
- [93] L. Muszbek, Z. Bagoly, Z. Bereczky, E. Katona, *The involvement of blood coagulation factor XIII in fibrinolysis and thrombosis*. Cardiovascular & hematological agents in medicinal chemistry **6**, 190-205 (2008).
- [94] F. Duckert, E. Jung, D. H. Shmerling, *A Hitherto Undescribed Congenital Haemorrhagic Diathesis Probably Due to Fibrin Stabilizing Factor Deficiency*. Thromb Diath Haemost **5**, 179-186 (1960).
- [95] L. Hsieh, D. Nugent, *Factor XIII deficiency*. Haemophilia **14**, 1190-1200 (2008).
- [96] B. Ahvazi, K. M. Boeshans, W. Idler, U. Baxa, P. M. Steinert, F. Rastinejad, *Structural basis for the coordinated regulation of transglutaminase 3 by guanine nucleotides and calcium/magnesium*. J Biol Chem **279**, 7180-7192 (2004).
- [97] B. Fleckenstein, Y. Molberg, S. W. Qiao, D. G. Schmid, F. von der Mullbe, K. Elgstoen, G. Jung, L. M. Sollid, *Gliadin T cell epitope selection by tissue transglutaminase in Celiac disease - Role of enzyme specificity and pH influence on the transamidation versus deamidation reactions*. J Biol Chem **277**, 34109-34116 (2002).
- [98] W. Kabsch, *Xds*. Acta crystallographica. Section D, Biological crystallography **66**, 125-132 (2010).
- [99] O. S. Smart, T. O. Womack, A. Sharff, C. Flensburg, P. Keller, W. Paciorek, C. Vonnrhein, G. Bricogne, *Grade Web Server; Global Phasing Limited: Cambridge, UK.* (2012).
- [100] P. G. Doiphode, M. V. Malovichko, K. N. Mouapi, M. C. Maurer, *Evaluating factor XIII specificity for glutamine-containing substrates using a matrix-assisted laser desorption/ionization time-of-flight mass spectrometry assay*. Anal Biochem **457**, 74-84 (2014).

- [101] Schechte.I, A. Berger, *On Size of Active Site in Proteases .I. Papain*. Biochem Bioph Res Co **27**, 157-& (1967).
- [102] K. Hardes, M. Z. Hammamy, T. Steinmetzer, *Synthesis and characterization of novel fluorogenic substrates of coagulation factor XIII-A*. Anal Biochem **442**, 223-230 (2013).
- [103] J. W. Keillor, K. Y. P. Apperley, A. Akbar, *Inhibitors of tissue transglutaminase*. Trends Pharmacol Sci **36**, 32-40 (2015).
- [104] J. C. Powers, J. L. Asgian, O. D. Ekici, K. E. James, *Irreversible inhibitors of serine, cysteine, and threonine proteases*. Chem Rev **102**, 4639-4750 (2002).
- [105] K. Fickenscher, A. Aab, W. Stuber, *A Photometric Assay for Blood-Coagulation Factor-Xiii*. Thromb Haemostasis **65**, 535-540 (1991).
- [106] J. J. Gorman, J. E. Folk, *Structural features of glutamine substrates for human plasma factor XIIIa (activated blood coagulation factor XIII)*. J Biol Chem **255**, 419-427 (1980).
- [107] D. de Sanctis, A. Beteva, H. Caserotto, F. Dobias, J. Gabadinho, T. Giraud, A. Gobbo, M. Guijarro, M. Lentini, B. Lavault, T. Mairs, S. McSweeney, S. Petitdemange, V. Rey-Bakaikoa, J. Surr, P. Theveneau, G. A. Leonard, C. Mueller-Dieckmann, *ID29: a high-intensity highly automated ESRF beamline for macromolecular crystallography experiments exploiting anomalous scattering*. J Synchrotron Radiat **19**, 455-461 (2012).
- [108] P. Emsley, K. Cowtan, *Coot: model-building tools for molecular graphics*. Acta Crystallogr D **60**, 2126-2132 (2004).
- [109] N. W. Moriarty, R. W. Grosse-Kunstleve, P. D. Adams, *electronic Ligand Builder and Optimization Workbench (eLBOW): a tool for ligand coordinate and restraint generation*. Acta crystallographica. Section D, Biological crystallography **65**, 1074-1080 (2009).
- [110] H. Tatsukawa, Y. Furutani, K. Hitomi, S. Kojima, *Transglutaminase 2 has opposing roles in the regulation of cellular functions as well as cell growth and death*. Cell Death Dis **7**, (2016).
- [111] I. K.-S. Máté Á. Demény, László Fésüs, *Structure of Transglutaminases: Unique Features Serve Diverse Functions*. (Springer, 2016), pp. 1-41.
- [112] S. P. Liu, R. A. Cerione, J. Clardy, *Structural basis for the guanine nucleotide-binding activity of tissue transglutaminase and its regulation of transamidation activity*. P Natl Acad Sci USA **99**, 2743-2747 (2002).
- [113] N. S. Caron, L. N. Munsie, J. W. Keillor, R. Truant, *Using FLIM-FRET to Measure Conformational Changes of Transglutaminase Type 2 in Live Cells*. Plos One **7**, (2012).
- [114] S. Boros, E. Ahrman, L. Wunderink, B. Kamps, W. W. de Jong, W. C. Boelens, C. S. Emanuelsson, *Site-specific transamidation and deamidation of the small heat-shock protein Hsp20 by tissue transglutaminase*. Proteins **62**, 1044-1052 (2006).
- [115] J. Stamnaes, B. Fleckenstein, L. M. Sollid, *The propensity for deamidation and transamidation of peptides by transglutaminase 2 is dependent on substrate affinity and reaction conditions*. Bba-Proteins Proteom **1784**, 1804-1811 (2008).
- [116] MOE 2014.2009 - Manual.
- [117] *Molecular Operating Environment (MOE)*, 2014.2009; Chemical Computing Group Inc., 1010 Sherbooke St. West, Suite #2910, Montreal, QC, Canada, H2013A 2012R2017, **2015**.
- [118] P. Labute, *The generalized Born/volume integral implicit solvent model: Estimation of the free energy of hydration using London dispersion instead of atomic surface area*. J Comput Chem **29**, 1693-1698 (2008).
- [119] W. L. DeLano, J. W. Lam, *PyMOL: A communications tool for computational models*. Abstr Pap Am Chem S **230**, U1371-U1372 (2005).
- [120] D. A. Erlanson, R. S. McDowell, T. O'Brien, *Fragment-based drug discovery*. J Med Chem **47**, 3463-3482 (2004).

- [121] D. C. Rees, M. Congreve, C. W. Murray, R. Carr, *Fragment-based lead discovery*. *Nat Rev Drug Discov* **3**, 660-672 (2004).
- [122] O. Cala, I. Krimm, *Ligand-Oriented Based Fragment Selection in STD NMR Screening*. *J Med Chem* **58**, 8739-8742 (2015).
- [123] J. K. Kranz, C. Schalk-Hihi, *Protein thermal shifts to identify low molecular weight fragments*. *Methods in enzymology* **493**, 277-298 (2011).
- [124] P. Linke, K. Amaning, M. Maschberger, F. Vallee, V. Steier, P. Baaske, S. Duhr, D. Breitsprecher, A. Rak, *An Automated Microscale Thermophoresis Screening Approach for Fragment-Based Lead Discovery*. *Journal of biomolecular screening*, (2015).
- [125] S. Perspicace, D. Banner, J. Benz, F. Muller, D. Schlatter, W. Huber, *Fragment-Based Screening Using Surface Plasmon Resonance Technology*. *Journal of biomolecular screening* **14**, 337-349 (2009).
- [126] N. Radeva, S. G. Krimmer, M. Stieler, K. Fu, X. J. Wang, F. R. Ehrmann, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, J. Schiebel, A. Heine, G. Klebe, *Experimental Active-Site Mapping by Fragments: Hot Spots Remote from the Catalytic Center of Endothiapepsin*. *J Med Chem* **59**, 7561-7575 (2016).
- [127] N. Radeva, J. Schiebel, X. J. Wang, S. G. Krimmer, K. Fu, M. Stieler, F. R. Ehrmann, A. Metz, T. Rickmeyer, M. Betz, J. Winkquist, A. Y. Park, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, *Active Site Mapping of an Aspartic Protease by Multiple Fragment Crystal Structures: Versatile Warheads To Address a Catalytic Dyad*. *J Med Chem* **59**, 9743-9759 (2016).
- [128] J. Schiebel, S. G. Krimmer, K. Rower, A. Knorlein, X. J. Wang, A. Y. Park, M. Stieler, F. R. Ehrmann, K. Fu, N. Radeva, M. Krug, F. U. Huschmann, S. Glockner, M. S. Weiss, U. Mueller, G. Klebe, A. Heine, *High-Throughput Crystallography: Reliable and Efficient Identification of Fragment Hits*. *Structure* **24**, 1398-1409 (2016).
- [129] J. Schiebel, N. Radeva, S. G. Krimmer, X. J. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe, *Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study*. *Acs Chem Biol* **11**, 1693-1701 (2016).
- [130] R. D. Bach, O. Dmitrenko, C. Thorpe, *Mechanism of thiolate-disulfide interchange reactions in biochemistry*. *J Org Chem* **73**, 12-21 (2008).
- [131] R. L. Rich, D. G. Myszka, *Advances in surface plasmon resonance biosensor analysis*. *Curr Opin Biotech* **11**, 54-61 (2000).
- [132] R. E. Kontermann, *Strategies for extended serum half-life of protein therapeutics*. *Curr Opin Biotech* **22**, 868-876 (2011).
- [133] M. J. Roberts, M. D. Bentley, J. M. Harris, *Chemistry for peptide and protein PEGylation*. *Adv Drug Deliver Rev* **64**, 116-127 (2012).
- [134] J. W. Lichtman, J. A. Conchello, *Fluorescence microscopy*. *Nat Methods* **2**, 910-919 (2005).
- [135] R. Yuste, *Fluorescence microscopy today*. *Nat Methods* **2**, 902-904 (2005).
- [136] A. Mero, M. Schiavon, F. M. Veronese, G. Pasut, *A new method to increase selectivity of transglutaminase mediated PEGylation of salmon calcitonin and human growth hormone*. *Journal of controlled release : official journal of the Controlled Release Society* **154**, 27-34 (2011).
- [137] K. R. Siebenlist, D. A. Meh, M. W. Mosesson, *Protransglutaminase (factor XIII) mediated crosslinking of fibrinogen and fibrin*. *Thromb Haemostasis* **86**, 1221-1228 (2001).

## Abbreviations

Å	Angstrom ( $1\text{Å} = 10^{-10}\text{ m}$ )
Ac	Acetate
Ala	Alanine
Arg	Arginine
Asn	Asparagine
Asp	Aspartic acid
<i>B</i> -factor	Debye-Waller factor
c	cellular
Ca	Calcium
Cys	Cysteine
Da	Dalton
DIPEA	Diisopropylethylamine
DMF	Dimethylformamide
DON	6-Diazo-5-oxo-L-norleucine
ESR	Electron spin resonance
F	Fragment
FBLD	Fragment-based lead discovery
$F_c$	Calculated structure amplitudes
FII	Factor II
Fmoc	Fluorenylmethyloxycarbonyl
$F_o$	Observed structure amplitudes
FRET	Förster resonance energy transfer
FX	Factor X
FXIII	Factor XIII
FXIII-A	A subunit of FXIII
FXIIIa°	Active FXIII-A with uncleaved activation peptide
FXIII-B	B subunit of FXIII
GDP	Guanosine diphosphate
Gln	Glutamine

Glu	Glutamic acid
Gly	Glycine
G-protein	Guanine nucleotide-binding protein
GTP	Guanosine triphosphate
HATU	1-[Bis(dimethylamino)methylene]-1H-1,2,3-triazolo[4,5-b]pyridinium 3-oxid hexafluorophosphate
H-bond	Hydrogen bond
His	Histidine
HOBt	Hydroxybenzotriazole
HTS	High-throughput screening
IC <sub>50</sub>	Half maximal inhibitory concentration
Ile	Isoleucine
ITC	Isothermal microcalorimetry
k	Kilo
K	Kelvin
LC	Liquid chromatography
Leu	Leucin
Lys	Lysine
MA	Michael acceptor
Met	Methionine
MS	Mass spectrometry
MST	Micro thermophoresis
MW	Molecular weight
Nle	Norleucine
NMR	Nuclear magnetic resonance
p	Plasmatic
PDB	Protein data bank
PDB ID	Protein data bank identifier
PEG	Polyethylene glycol
pH	Potentialis hydrogenii
Phe	Phenylalanine



pK <sub>a</sub>	Logarithmic acid dissociation constant
PKC	Protein kinase C
Pro	Proline
R-factor	Reliability factor
RMSD	Root mean square deviation
Ser	Serine
SPR	Surface plasmon resonance
TBTU	N,N,N',N'-Tetramethyl-O-(benzotriazol-1-yl)uronium tetrafluoroborate
tBu	Tert-butyl
TG(s)	Transglutaminase(s)
TG2a*	Transglutaminase 2 in the active linear state
TG3a <sup>c</sup>	Transglutaminase 3 with enzymatically cleaved β-barrel domains
Thr	Threonine
TLS	Translation/libration/screw-motion
Tris	Tris(hydroxymethyl)aminomethane
Trp	Tryptophan
TSA	Thermal-shift assay
Tyr	Tyrosine
V <sub>0</sub>	Void volume
Val	Valine
V <sub>e</sub>	Elution volume
X-ray	Röntgen radiation

## One- and three-letter code of proteinogenic amino acids

The one-letter code for amino acids was used in the figures of the present thesis. The following table contains the one-/three-letter code and the name of the corresponding amino acid.

One-letter code	Three-letter code	Amino acid
A	Ala	Alanine
C	Cys	Cysteine
D	Asp	Aspartate
E	Glu	Glutamate
F	Phe	Phenylalanine
G	Gly	Glycine
H	His	Histidine
I	Ile	Isoleucine
K	Lys	Lysine
L	Leu	Leucine
M	Met	Methionine
N	Asn	Asparagine
P	Pro	Proline
Q	Gln	Glutamine
R	Arg	Arginine
S	Ser	Serine
T	Thr	Threonine
V	Val	Valine
W	Trp	Tryptophan
Y	Tyr	Tyrosine

## Publications and Patents

M. Stieler, J. Weber, M. Hils, P. Kolb, A. Heine, C. Buchold, R. Pasternack, G. Klebe *Structure of Active Coagulation Factor XIII Triggered by Calcium Binding: Basis for the Design of Next-Generation Anticoagulants*. Angew Chem Int Edit **52**, 11930-11934 (2013).

M. Hils, R. Pasternack, C. Büchold, J. Weber, A. Heine, G. Klebe, M. Stieler: *Crystal structure of blood coagulation factor xiiia*, WO 2006056575 A1.

N. Radeva, J. Schiebel, X. Wang, S. G. Krimmer, K. Fu, M. Stieler, F. R. Ehrmann, A. Metz, T. Rickmeyer, M. Betz, J. Winquist, A. Y. Park, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe *Active Site Mapping of an Aspartic Protease by Multiple Fragment Crystal Structures: Versatile Warheads To Address a Catalytic Dyad*. J Med Chem **59**, 9743-9759 (2016).

N. Radeva, S. G. Krimmer, M. Stieler, K. Fu, X. Wang, F. R. Ehrmann, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, J. Schiebel, A. Heine, and G. Klebe *Experimental Active-Site Mapping by Fragments: Hot Spots Remote from the Catalytic Center of Endothiapepsin*. J Med Chem **59**, 7561-7575 (2016).

J. Schiebel, S. G. Krimmer, K. Röwer, A. Knörlein, X. Wang, A. Y. Park, M. Stieler, F. R. Ehrmann, K. Fu, N. Radeva, M. Krug, F. U. Huschmann, S. Glöckner, M. S. Weiss, U. Mueller, G. Klebe, A. Heine *High-Throughput Crystallography: Reliable and Efficient Identification of Fragment Hits* Structure **24**, 1398-1409 (2016).

J. Schiebel, N. Radeva, S. G. Krimmer, X. Wang, M. Stieler, F. R. Ehrmann, K. Fu, A. Metz, F. U. Huschmann, M. S. Weiss, U. Mueller, A. Heine, G. Klebe *Six Biophysical Screening Methods Miss a Large Proportion of Crystallographically Discovered Fragment Hits: A Case Study*. ACS Chem Biol **11**, 1693-1701 (2016).

## Conference Presentations

Stieler M. *Crystal Structure of Blood Coagulation Factor XIII in the Active Conformation*. **Talk** at Gordon Research Conference - Transglutaminases in Human Disease Processes, Lucca, Italy, 2014.

Stieler M., Weber J., Kolb P., Heine A., Büchold C., Hils M., Pasternack R., Klebe G. *Crystal Structure of Blood Coagulation Factor XIII: Template for the Design of a Novel Anticoagulant*. **Poster** at Annual Meeting of the German Pharmaceutical Society awarded with Lesmüller Poster Award, Düsseldorf, Germany, 2014.

Stieler M., Weber J., Hils M., Kolb P., Heine A., Büchold C., Pasternack R., Klebe G. *The Next Generation Anticoagulant: Structure-Based Design of a Factor XIII Blocker*. **Talk** at Annual Meeting of the Society of Thrombosis and Hemostasis Research, Düsseldorf, Germany, 2015.

## Danksagung

Solch ein Vorhaben, wie die Dissertation, ist nur unter der Mithilfe vieler Menschen möglich, denen ich nachfolgend danken möchte.

Mein Dank gilt zunächst meinem Doktorvater Herrn **Prof. Dr. Gerhard Klebe**, der immer ein offenes Ohr für seine Doktoranden hat. Besonders habe ich die eigenständige Arbeitsweise in unserer Arbeitsgruppe geschätzt, was mir die Möglichkeit gegeben hat, mich während meiner Doktorarbeit mit diversen Fragestellungen meines Projektes zu beschäftigen. Auch für die Möglichkeit der Teilnahme an mehreren internationalen Konferenzen möchte ich mich bei Herrn Prof. Klebe bedanken.

Des Weiteren möchte ich mich bei meinem Zweitgutachter Herrn **Prof. Dr. Peter Kolb** bedanken, der mich während meiner Doktorarbeit im Bereich der Computerchemie unterstützt hat.

Ein besonderer Dank gilt auch Herrn Dr. Ralf Pasternack, Herrn Dr. Martin Hils, Herrn Johannes Weber, Herrn Dr. Andreas Heil und Herrn Dr. Christian Büchold der Firma **Zedira GmbH** für die konstruktive und freundschaftliche Zusammenarbeit sowie der Zurverfügungstellung des Proteins und der Inhibitoren.

Herrn **Prof. Dr. Thorsten Steinmetzer** und Herrn **Dr. Zouhir Hammamy** danke ich für die Synthese des Inhibitors Mi0621.

Herrn **Prof. Dr. Andreas Heine** möchte ich für die geduldige Unterstützung im Bereich der Proteinkristallographie danken.

Die Promotionszeit werde ich als eine sehr schöne Zeit in Erinnerung behalten, was nicht zuletzt an dem angenehmen kollegialen Umfeld lag, in dem sich viele Freundschaften entwickelt haben. Ich möchte daher all meinen **Kollegen** danken, insbesondere Chris, Fredi, Phong, Michael, Stefan, Nicole, Manu, Jonathan, Corey, Florent, Felix, Neli, Denis, Barbara, Adam, Inna, Alex, Namir, Timo, Jakub und Tobias.

Für die Administration des Computernetzwerkes und der Organisation sowie der Unterstützung im Laborbereich gilt mein Dank Felix, Michael, Thomas, Lydia, Christian und Steffi.

Auch meinen **Freunden** und **Verwandten** möchte ich für deren moralischen Beistand und Verständnis danken.

Letztendlich danke ich von ganzem Herzen meinen **Eltern** und meiner gesamten **Familie** für die jahrelange bedingungslose Unterstützung.

## Lebenslauf

	<b>Persönliche Daten</b>
Name	Martin Stieler
Geburtsdatum	22.01.1981
Geburtsort	Marburg
Staatsangehörigkeit	deutsch
	<b>Beruflicher Werdegang</b>
12/2017 – heute	<b>Wissenschaftlicher Mitarbeiter</b> Zedira GmbH, Darmstadt
	<b>Akademischer Werdegang</b>
11/2011 – heute	<b>Doktorand</b> Philipps-Universität Marburg, Pharmazeutische Chemie, AG Klebe
04/2005 – 10/2011	<b>Diplom, Chemie</b> Philipps-Universität Marburg
09/2001 – 12/2004	<b>Vordiplom, Chemie</b> Parallel: Erlangung der fachgebundenen Hochschulreife Gesamthochschule Siegen
	<b>Berufsausbildung</b>
09/1999 – 06/2001	<b>Chemisch-Technischer-Assistent</b> Adolf-Reichwein-Schule, Marburg
	<b>Schulische Laufbahn</b>
09/1997 – 06/1999	<b>Fachoberschule, Fachrichtung Chemietechnik</b> Adolf-Reichwein-Schule, Marburg
09/1991 – 06/1997	<b>Realschule</b> Gesamtschule Ebsdorfergrund, Heskem

## **Eidesstattliche Erklärung**

(gemäß §10 der Promotionsordnung)

Ich versichere, dass ich meine Dissertation

### **Structure-Based Design of a Blood Coagulation Factor XIII Blocker**

selbständig ohne unerlaubte Hilfe angefertigt und mich dabei keiner anderen als der von mir ausdrücklich bezeichneten Quellen bedient habe. Alle vollständig oder sinngemäß übernommenen Zitate sind als solche gekennzeichnet.

Die Dissertation wurde in der jetzigen oder einer ähnlichen Form noch bei keiner anderen Hochschule eingereicht und hat noch keinen sonstigen Prüfungszwecken gedient.

Marburg, den 08.01.2017



---

(Martin Stieler)