Design, Synthesis and Characterization of Small Molecule Inhibitors and Small Molecule – Peptide Conjugates as Protein Actors

Jonas W. Nilsson

Division of Chemistry
Department of Physics and Measurement Technology, Biology and Chemistry
Linköpings Universitet, SE-581 83 Linköping, Sweden

Linköping 2005
Abstract

This thesis describes different aspects of protein interactions. Initially the function of peptides and their conjugates with small molecule inhibitors on the surface of Human Carbonic Anhydrase isoenzyme II (HCAII) is evaluated.

The affinities for HCAII of the flexible, synthetic helix-loop-helix motif conjugated with a series of spacered inhibitors were measured by fluorescence spectroscopy and found in the best cases to be in the low nM range. Dissociation constants show considerable dependence on linker length and vary from 3000 nM for the shortest spacer to 40 nM for the longest with a minimum of 5 nM for a spacer with an intermediate length. A rationale for binding differences based on cooperativity is presented and supported by affinities as determined by fluorescence spectroscopy. Heteronuclear Single Quantum Correlation Nuclear Magnetic Resonance (HSQC) spectroscopic experiments with $^{15}$N-labeled HCAII were used for the determination of the site of interaction.

The influence of peptide charge and hydrophobicity was evaluated by surface plasmon resonance experiments. Hydrophobic sidechain branching and, more pronounced, peptide charge was demonstrated to modulate peptide – HCAII binding interactions in a cooperative manner, with affinities spanning almost two orders of magnitude.

Detailed synthesis of small molecule inhibitors in a general lead discovery library as well as a targeted library for inhibition of α-thrombin is described. For the lead discovery library 160 members emanate from two $N_4$-arylpiperazine-2-carboxylic acid scaffolds derivatized in two dimensions employing a combinatorial approach on solid support.

The targeted library was based on peptidomimetics of the D-Phe-Pro-Arg showing the scaffolds cyclopropane-1R,2R-dicarboxylic acid and (4-amino-3-oxo-morpholin-2-y1)-acetic acid as proline isosters. Employing 4-aminomethyl-benzamidine as arginine mimic and different hydrophobic amines and electrophiles as D-phenylalanine mimics resulted in 34 compounds showing $IC_{50}$ values for α-thrombin ranging more than three orders of magnitude with the best inhibitor showing an $IC_{50}$ of 130 nM. Interestingly, the best inhibitors showed reversed stereochemistry in comparison with a previously reported series employing a 3-oxo-morpholin-2-yl-acetic acid scaffold.
This thesis is based on four scientific papers and one summary, discussed in the introductory text and enclosed as appendices:

**Paper I**  

**Paper II**  

**Paper III**  

**Paper IV**  

**Unpublished results**  
1 Introduction

Protein actors have a profound role in bioorganic chemistry. The comprehension of protein interactions are crucial for the understanding of cellular transport and communication, biosensing (diagnostics and screening) and drug design. The diverse set of tools available for the evaluation of protein interactions include techniques for biophysical characterization such as X-ray crystallography, NMR spectroscopy and surface plasmon resonance. Structure-activity relationships may be utilized as well as biomolecules capable of binding recognition and reporting. Biological or bio-mimicking constructs such as antibodies, aptamers and synthetic peptides may be used for the identification, quantification and inhibition of biomolecular interactions involving proteins.

Peptide conjugates come in many flavors and range from side chain labeled, small molecule derivatized and branched peptides to more complex macromolecular constructs such as conjugations of PNA/RNA/DNA to peptide oligomers. Peptides have the advantage of being accessible via routes that are completely synthetic, making directed modifications easy, selective and flexible.

Since more than 25 years instruments for automated peptide synthesis have been commercially available. Even though the automation initially might have been seen as an expensive luxury, instruments can now be afforded even by small research groups. The speed and ease with which even longer peptides can be synthesized by this tool is remarkable, and with an ever expanding set of available building blocks, protection groups and unnatural amino acids, peptide conjugates have been made available leading to a plethora of new molecules and applications.

In general most peptides with as many as 50 or even 60 amino acids can be synthesized. By employing orthogonal protection strategies and self catalyzed selective condensations, post-synthesis modifications in three, four or even five dimensions are achievable. The variability is immense. The synthesis of one molecule each of every peptide with 58 amino acids that can be constructed from the 20 natural amino acids would require 5 times the mass of the visible universe!
In sensing applications, modifications of any residue by the incorporation of fluorescense-, spin- and radio-labels can be used to probe biomolecular surfaces as well as other targets. Peptides can also be used as vehicles for drug delivery. Cytotoxins conjugated with specific peptides can target cancer cells to induce cell death\(^1\). PNA-peptide conjugates can be used to site specifically cleave or modify DNA\(^2,3\) or as antisense-based imaging agents when chelated with radioactive nuclei for the detection of pathological overexpression of certain mRNA\(^4\).

In this thesis I demonstrate affinity modulation of small molecule inhibitors conjugated to peptides. Not only can peptides be synthesized in any sequence, but linkage length, character as well as position of conjugation can easily be varied, to generate affinity probes suitable for almost any target. When peptides are combined with sensing labels, the affinities can be determined with high accuracy by direct measurement of single samples or in microtiter plate format for large-scale drug screening and drug target evaluation.

The prerequisite for drug screening, however, is a potential drug. In chapter 3 the rational design and synthesis of \(\alpha\)-thrombin inhibitors are described. \(\alpha\)-Thrombin is a key enzyme in the coagulation cascade, and modulation of its activity is crucial for patients suffering from thromboembolic disorders, the major cause of morbidity and mortality in the developed world.

The starting point for the design of a new enzyme inhibitor or other drug candidate is to gain as much information as possible about the dedicated target. This includes the deduction of key structural elements used in previously known inhibitors or substrates, lead library screening, structure activity relationships in combination with X-ray crystallography analysis of enzyme with and without inhibitor or natural substrate.

Lead libraries exist in two forms. General lead discovery libraries (Chapter 3.1) are drug like or lead like substances without any designated target. High throughput screening of thousands or millions of compounds can lead to the discovery of new leads unrelated to previously known inhibitors\(^5\). Targeted libraries (Chapter 3.2) are created as an expansion of a known lead where some of the key structural elements are predefined. Further optimization includes in vitro activity screening and X-ray crystallography analyses of inhibitors bound to the enzyme leading to the design of new targeted libraries. Thus, the circle is closed and this iterative structure based drug design process will hopefully lead to a preclinical drug.

This is, however, only the first step towards a new drug. Before reaching the market a series of evaluations must be performed. The first step being evaluation of bioavailability and toxic side effects most commonly performed on cell culture or animal. If no obstacles are found, tests on humans are performed in three phases. First on a small number of healthy humans to confirm that no side effects undetected in animals
occur (Phase I), secondly on a small number of patients (Phase II) and finally on a large number of patients (Phase III). Only a small number of the preclinical drugs actually reach the market.
2 Small molecule – peptide conjugates based on enzyme inhibitors

Interactions between biomacromolecules are required for the function of all biological systems. Protein assemblies working together as a single unit perform functions such as transcription and ATP synthesis. In vivo transport, signal transduction and pathogen recognition all depend on surface interactions with proteins or antibodies. Advances in proteomics demand a deeper understanding of the nature of the complicated framework of hydrogen bonds and electrostatic/hydrophobic interactions that are needed to match the intricate polymorph surface contours of proteins.

The large pool of available biomacromolecules such as antibodies, aptamers and proteins optimized by phage display is a valuable source of components for biomolecular recognition. The biological approach is a well-established and useful tool for the production of high affinity macromolecules. The structural control, however, is mostly left to the biological system and directed optimizations as well as the incorporation of non-natural building blocks might be difficult to assess.

Pharmaceutical Industry, working with small molecules as inhibitors can achieve high affinities, not by flexibility or complexity, but by cumbersome optimization of small, rigid lead substances. Systematic trial and error, together with computational aids such as molecular modeling and in silico docking are used to find the “perfect fit”.

This chapter describes a conjugate approach, which combines the affinity and specificity of known small molecule inhibitors with the variability of pre-organized polypeptide constructs. Since they are produced in vitro, the available modifications are almost unlimited, thereby giving us tools such as unnatural amino acid constructs and fluorescent reporter functionality. The conjugate consists of a polypeptide connected via a linker to a small molecule inhibitor that can be anchored into the active site of the target enzyme. By systematic variation of linker length or position as well as of peptide charge distribution or hydrophobic character it can be used to probe peptide – protein interactions and to optimize affinity.

Even though synthetic strategies employed in lead optimizations or in peptide conjugate syntheses limit the accessible size of libraries of screened species to several orders of
magnitude below those of aptamers and antibodies ($10^{12}$-$10^{14}$) the directed approach is expected to compensate for this. Intelligent selections replace random screening.

2.1 Polypeptide design
To be able to present a pre-organized surface, the peptide needs to adopt a defined secondary and supersecondary structure. The scaffold in this investigation is a 42 amino acid residue helix-loop-helix motif that dimerizes to form a four-helix bundle. This was achieved partly by choosing amino acids with high helix propensities and by introducing a helix breaking proline at position 21 to generate the loop motif.

In a perfect helix there are 3.6 residues per turn. In the folded state residues with approximately this interresidue distance, i.e. residues in positions three and seven of the heptad repeat pattern will form one side of the helix. Introduction of hydrophobic residues in these positions (e.g. pos 16, 12, 9, 5 upstream from loop and pos 27, 31, 35, 38 downstream) leads to a hydrophobic face that can interact with other helical hydrophobic faces and form helical bundle structures. If the other residues in the helix are polar and charged the helix is referred to as amphiphilic. When helical bundles are formed the hydrophobic faces from several amphiphilic helices generate a hydrophobic core and the polar residues ensure solubility. Water is excluded from the core and the free energy is minimized.

The dimerization not only stabilizes the construct, but in the system described in this thesis it also enhances the effect of reporter functions such as fluorescent probes. Upon binding to a host protein, hydrophobic and charged interactions between the monomer subunits are replaced by interactions between each monomer and the target protein, making it possible for the peptide to bind as a monomer. A fluorescent probe attached to the peptide may have a self-quenching effect due to the spatial proximity in the dimer. As this quenching is removed upon binding, a “light-up” effect will occur, making binding interactions directly measurable by fluorescence.

2.2 The peptide conjugate – HCAII system
Human Carbonic Anhydrase isoenzyme II (HCAII) is a well-characterized enzyme that has been used for studies of protein folding, ligand design and enzyme-template based inhibitor generation employing “click chemistry”. In human erythrocytes it catalyzes the equilibrium between carbon dioxide and water, forming bicarbonate and oxonium ion. It is a monomeric protein constituted of 259 amino acids. In its native form, a Zinc ion is coordinated by three histidine residues in the active site located in a 15 Å deep funnel. Several high resolution X-ray crystal structures of HCAII co-crystallized with high-affinity inhibitors have been reported.

One class of inhibitors is based on substituted benzenesulfonamides as their common motif. Benzenesulfonamide itself has been shown to have a $K_d$ of $1.5 \mu$M and the alkyl
substituted 4-sulfamoyl-benzamides show chain length dependent $K_d$ values ranging from approximately 80-150 nM (methyl) to 1.2-2.5 nM (heptyl or octyl)\textsuperscript{9,12} (Figure 2.1).

\[ \text{Figure 2.1. Benzenesulfonamide (top) and the 4-sulfamoyl-benzamide group (bottom) used as an anchoring unit into the active site of HCAII.} \]

X-ray crystallography derived structures\textsuperscript{13} strongly imply that the directional vector of the Zn – NH$_2$SO$_2$ bond as well as the position of the aromatic ring are well defined for all \textit{para} substituted benzenesulfonamides as well as for the \textit{ortho,ortho} and \textit{meta,meta} difluoro substituted ones (Figure 2.2). Deviations are found for 3-mercuri-4-amino-benzenesulfonamide (3CA2) and to some extent for 3-nitro-4-(2-oxo-pyrrolidin-1-yl)-benzenesulfonamide\textsuperscript{10,14} (1KWQ) which displays an aromatic ring dislocated from the plane defined by the other inhibitors.

Employing the 4-sulfamoyl-benzamide as a basal connector into the active site ensures a well-defined positioning of the anchoring group. The conjugates were constructed by connecting the benzamide to the helix-loop-helix scaffold using an aliphatic linker. Interactions can be probed and tuned employing the conjugate principle. The linker extends the conjugate to the surface of the protein, placing the peptide in position for multiple hydrophobic/electrostatic interactions with HCAII (Figure 2.3). Subtle changes of linker length or position as well as changes in peptide charge or hydrophobic character as well as in the distribution of charges and hydrophobic groups within the peptide scaffold can affect both binding strength and protein distortion.
Figure 2.2. Aligned overlay showing a stick representation of nine inhibitors from the Protein Databank together with the Zinc coordinating histidines of HCAII at position 94, 96 and 119. The rest of HCAII is shown shaded in an overlay ribbon representation. The aromatic rings of the benzenesulfonamides are well aligned for the para substituted (1CNW; 1CNX; 1CNY; 1IF4; 1OQ5) and the symmetrically difluoro substituted (1IF5; 1IF6) inhibitors. The ortho mercury substituted (3CA2, green) and the meta nitro substituted (1KWQ, purple) show different positions of the aromatic ring.

Figure 2.3. Principle schematic of possible interactions between HCAII and a peptide – small molecule conjugate. The nature of the true interactions is not known in detail. HCAII structure and ligand position is based on X-ray crystallographic structure of HCAII co-crystallized with N-(2-{2-[2-(2-aminoacetylamino)-ethoxy]-ethoxy}-ethyl)-4-sulfamoyl-benzamide, pdb entry 1CNW.
2.3 Evaluation of linker length

A systematic variation of linker length was performed to evaluate the dependence of the distance between the active site Zinc ion of HCAII and the peptide helix-loop-helix motif [Paper I]. The linker needs a specific length to be able to reach down into the cleft of the active site without suffering the penalty of severely distorting the peptide or protein. However, if the linker is too long there may be an entropic loss due to residual degrees of freedom in the bound state leading to independent binding of the benzenesulfonamide and the peptide. Optimally, binding contributions from ligand, linker and peptide should work simultaneously and in synergy. The chelate effect can then generate cooperativity between the different binding events leading to a strong binder. Cooperativity arises when the pre-organization of the peptide resulting from the binding of the benzenesulfonamide into the active site is such that it is superior to that of a freely moving peptide.

2.3.1 Design and synthesis of a helix-loop-helix motif linked to a benzenesulfonamide group

Two 42 amino acid residue polypeptide templates were utilized (KE2, TA4). In solution these helix-loop-helix templates were found to be dimers at low µM concentration but it has been previously shown that the peptides are bound as monomers to HCAII\(^{15}\). A fluorophore was introduced to enable direct measurement of the strength of the interaction by standard fluorometry titrations.

The peptides were synthesized according to standard Fmoc chemistry protocols using acid labile protection groups. The side chain of the lysine in position 15 was protected with an allyloxycarbonyl protection group which after solid-phase synthesis of the peptide could be selectively removed by Pd\([PPh_3]_4\) and then sulfonylated with a dansyl probe for fluorometric assay. Cleavage of the peptide from the resin by trifluoroacetic acid was followed by purification on reversed phase HPLC.

The ligands were prepared from 4-sulfamoyl-benzoic acid and unbranched primary amino acids using carbodiimide chemistry (Scheme 2.1).

![Scheme 2.1. Synthesis of the activated benzenesulfonamide ligands.](image-url)
The helix-loop-helix motif (KE2-DNS(15)) was selectively modified at position 34 using the site selective self-modification reaction, a valuable tool for easy access to modified peptides\textsuperscript{16}. Previous studies have shown that the histidine in position 11 preferably directs acylation to a lysine in position 15 in similar peptides. When this position does not contain an amine, a lysine in position 34 can be selectively targeted. This would in principle be applicable to this system and both the dansyl probe and the benzenesulfonamide group could be introduced in the correct positions in aqueous solution by sequential addition of the corresponding reagents to a solution of the peptide scaffold. However, due to solubility issues, the dansyl probe was introduced on solid support employing a selective protection group strategy.

Selective modification of lysine 34 was achieved in hepes buffer, pH 8 using the \textit{N}-hydroxy succinimide activated ligand to yield the fluorophore labeled conjugates KE2-DNS(15)-C\textit{n}(34), \(n=6; 7; 8; 11\) (Scheme 2.2). The alternative conjugate TA4-DNS(15)-C8(34) was produced in an analogous manner.

![Scheme 2.2. Selective side chain modification with linked benzenesulfonamide.](image)

**2.3.2 Conjugate – HCAII interactions**

All peptide conjugates were screened for affinity to HCAII. This was done in three modes. Initially, direct titration of a 0.2 \(\mu\)M solution of conjugate with HCAII was used to determine the dissociation constant of the conjugate – HCAII complex. However for direct titration to be a reliable method, the conjugate concentration should be in the same order of magnitude as the dissociation constant. Due to limitations in signal intensities and surface adhesion problems at low concentrations, additional measurements utilizing indirect determination of dissociation constants with a competitive inhibitor (acetazolamide) were carried out. Solutions containing 2 \(\mu\)M
peptide conjugate and 1 µM HCAII was titrated with acetazolamide. The determined affinities was dependent on whether the wavelength of maximum fluorescence intensity or the absolute value of the intensity was used, and thus both will be used in the following discussion.

It is evident that for short linkers (0-4 methylene groups) the affinities are all low and surprisingly indifferent to linker length (Table 2.1). This might illustrate a state in which severe distortions to the peptide or protein structure is needed to accommodate binding. By introducing a fifth methylene group however (KE2-DNS(15)-C6(34)) the affinity increased markedly showing Kd values of 20; 25 and 16 nM when determined by the three modes of measurement. The consistent values and good empirical fit indicates that the 1:1 binding model is valid. For the longest linkers however (KE2-DNS(15)-C8(34) and KE2-DNS(15)-C11(34)) there are some discrepancies between the different modes of measurement. Also in the case of C11 the fluorescence intensity increases with increasing acetazolamide concentration, which is opposite to the expected behavior. This might indicate a more complex way of binding to the protein.

<table>
<thead>
<tr>
<th>Conjugate</th>
<th>No. CH₂ in linker</th>
<th>Kd / nM</th>
<th>[a]</th>
<th>[b]</th>
<th>[c]</th>
</tr>
</thead>
<tbody>
<tr>
<td>KE2-DNS(15)-C0(34)*</td>
<td>-</td>
<td>3000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C2(34)</td>
<td>1</td>
<td>1000</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C4(34)</td>
<td>3</td>
<td>700</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C5(34)</td>
<td>4</td>
<td>800</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C6(34)</td>
<td>5</td>
<td>20</td>
<td>23±2</td>
<td>19±2</td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C7(34)</td>
<td>6</td>
<td>10</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C8(34)</td>
<td>7</td>
<td>4</td>
<td>29±8</td>
<td>6±2</td>
<td></td>
</tr>
<tr>
<td>KE2-DNS(15)-C11(34)</td>
<td>10</td>
<td>24±4</td>
<td>22±4</td>
<td>53±13</td>
<td></td>
</tr>
<tr>
<td>TA4-DNS(15)-C8(34)</td>
<td>7</td>
<td>70</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

[a] As determined from direct titration of conjugate with HCAII. [b] and [c] As determined from titration of HCAII – conjugate complex with a competitive inhibitor using fluorescence intensities [b] or wavelength maxima [c] to extract data. Data is presented as duplicates ± one standard deviation as determined by the Levenberg-Marquardt curve-fitting algorithm. Dissociation constants for entries 1-4 are from ref. 15.

Table 2.1. Affinities of conjugates for HCAII recognition and binding.

Measures were taken to determine the nature of these complexes. A quencher peptide (KE2-DAB(15)) in which the dansyl probe was replaced by 4-(4-dimethylamino-phenylazo)-benzoate (dabcyl) was utilized. By titrating the conjugate – HCAII complexes with KE2-DAB(15) the dissociation constants of the ternary complex HCAII – KE2-DNS(15)-Cn(34) ↔ KE2-DAB(15) was determined to be 290; 110 and 90 µM for C6; C8 and C11 respectively. Even though longer linkers show greater

* KE2-DNS(15)-C0(34) has no linker. The 4-sulfamoyl-benzoic acid is directly connected to Lys34.
tendencies for aggregation the dissociation constants are much higher than the concentration used for fluorometric assay (2 µM). Arguing that the ternary complex studied should have similar dissociation constants as that of HCAII – $\text{KE2-DNS(15)-Cn(34)} \leftrightarrow \text{KE2-DNS(15)-Cn(34)}$ the population of ternary complexes from HCAII and two polypeptide conjugates is low under the experimental conditions and thus no conclusive model could be presented.

In general, affinities based on measurements of wavelengths of maximum fluorescence should be considered more reliable than those based on fluorescence intensities due to self-quenching phenomena that are not linearly related to the dissociation equilibrium. Thus based on the entries in columns [a] and [c] in table 2.1 the dissociation constants are approximately 20; 10; 5 and 57 nM for conjugates having linkers with 5, 6, 7 and 10 methylene groups respectively. This indicates cooperativity with an optimum at 7 methylene groups. Shorter linkers exert excessive strain on the system and longer linkers indicate an entropic loss due to discrete behavior of peptide and benzenesulfonamide, both phenomena leading to an affinity decrease.

Two different peptide scaffolds with the C8 linkers were used. Previous studies showed that the KE2 scaffold is essentially a non-binder of HCAII by itself. In contrast, TA4 was shown to bind in the µM range. Surprisingly, this is not at all reflected in the conjugates where $\text{KE2-DNS(15)-C8(34)}$ shows more than an order of magnitude stronger binding to HCAII than $\text{TA4-DNS(15)-C8(34)}$ does. To understand the nature of the peptide – HCAII interactions in more detail a NMR spectroscopy investigation was initiated.

HSQC spectroscopic experiments with $^{15}$N-labeled HCAII in the presence and absence of conjugate were undertaken. The chemical shifts of HCAII are very conserved under the experimental conditions. Seven different samples of HCAII were analyzed for perturbations of the amide proton – nitrogen shifts and all cross peaks were reproduced within +/- 1.5 Hz in the $^{15}$N dimension and +/- 2.5 Hz in the $^1$H dimension. When the environment of an amide is moderated, either by close proximity to the conjugate or by structural perturbations due to binding, the chemical shifts are affected. To distinguish natural variations from real interactions, shifts exceeding 10 Hz in either the $^1$H or $^{15}$N dimension were considered significant.

The NMR spectrum has been assigned for HCAII$^{17}$, which enables visualization of the polypeptide conjugate binding interactions with HCAII. Figure 2.4 reveals a plausible explanation for why TA4 and KE2 show reversed behavior when binding independently or as conjugates. Unmodified KE2 did not affect the chemical shifts of HCAII due to its low affinity, an observation confirmed by SPR measurements. Even though residues around the active site of HCAII were affected by TA4 as well as by the conjugate TA4-DNS(15)-C8(34), binding regions were significantly different.
Figure 2.4. Residues of HCAII affected (dark) by TA4 (left) and TA4-DNS(15)-C8(34) (right).

Apparently, simultaneous binding of the scaffold to its medium-affinity binding site and sulfonamide to the Zinc ion was not possible. However, the area affected by TA4 seems to be located within reach of a linker originating in the active site. Redesigning the conjugate to adapt to the preferred location of the peptide would be an interesting challenge that might be achieved by systematic variations of linker length and hydrophobicity in combination with its attachment point to the peptide (see chapter 4.1).

For the KE2 – conjugate series differential perturbations of the HCAII surface were observed. No shifts were observed for KE2-DNS(15)-C5(34) possibly due to the inability of the bound peptide to retain its helical structure when the linkers are too short. It is unlikely that HCAII can accommodate a folded helical peptide within its binding cleft, but possibly an unordered peptide chain. This would lead to a large number of randomized conformations incapable of distinct interactions with HCAII. For longer linkers however, several perturbed amide protons were identified (Figure 2.5). This is also reflected by KE2-DNS(15)-C5(34) having a 40-fold lower affinity than KE2-DNS(15)-C6(34). The fact that the affinities for even shorter linkers are largely constant supports a model where the dominating binding event is the interaction between the aryl-sulfonamide and HCAII active site.

Figure 2.5. Residues of HCAII affected (dark) by KE2-DNS(15)-C6(34) (left), KE2-DNS(15)-C8(34) (middle) and KE2-DNS(15)-C11(34) (right).

For KE2-DNS(15)-C8(34) the interactions are significantly fewer than for KE2-DNS(15)-C6(34). In this complex 30% of the NMR intensity is lost suggesting some
degree of aggregation and the observation of fewer interactions between scaffold and HCAII remain unexplained.

For the longest linker, KE2-DNS(15)-C11(34), the observed interactions are very strong which might be due to a more complex binding model. Even though the nature of this complex is uncertain, a simple 1:1 interaction can be ruled out. Attempts to crystallize conjugate – HCAII complexes have so far failed, but NMR indicates a high molecular weight aggregate that might contain two or more HCAII. This would result in a molecular weight of >65kD for which the NMR signals would be broadened beyond detection and when looking at the total intensity of the NMR signal only 30% remains after addition of the conjugate. Thus, the majority of the complex ensembles are undetectable by NMR.

2.3.3 Conclusions

For optimum binding to HCAII no less than 5 methylene groups should be present in a polypeptide – benzenesulfonamide conjugate such as KE2-DNS(15)-Cn(34). Furthermore, 5 methylene groups are optimal if minimum translational freedom of the peptide is imperative. This however severely restricts the accessible binding area of HCAII, and puts rigorous constraints on the electrostatic contours of the peptide. In fact, a slightly longer linker might be preferred as this flexibility relaxes the demands on the peptide. This is also supported by the fact that maximum affinity was observed for a conjugate with 7 methylene groups.

If very long hydrophobic linkers are used, unwanted aggregation occurs. If increased hydrophobicity alone is the reason for the aggregation, hydrophilic elements might be introduced in linkers of similar length to enhance the probability that the scaffold binds in an optimal fashion to the target protein. It has been shown, however that polar polypeptide linkers provide poor binding, and thus the hydrophilic elements should be sparse keeping the linker character mainly hydrophobic.

2.4 Evaluation of scaffold properties

To deepen the understanding of the peptide properties that influence HCAII binding, a set of new peptides was synthesized [Paper II]. Starting from a non-binder (KE2) and a µM binder (TA4) as initial templates, two new peptides were designed to have intermediate properties between the two (Figure 2.6). The differences between the peptides were reduced to two parameters: peptide charge (-0.9 and +1.1 for KE2 and TA4 respectively) and hydrophobic branching (3 leucines and 3 norleucines for KE2 and TA4 respectively). The peptide JWN-H8 was designed to

<table>
<thead>
<tr>
<th>Table 2.2. Properties of scaffold peptides</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peptide</td>
</tr>
<tr>
<td>----------</td>
</tr>
<tr>
<td>KE2</td>
</tr>
<tr>
<td>JWN-H8</td>
</tr>
<tr>
<td>JWN-H9</td>
</tr>
<tr>
<td>TA4</td>
</tr>
<tr>
<td>TA4+</td>
</tr>
</tbody>
</table>

<sup>a</sup>Based on typical pKa values of amino acid residues in short peptides.
have the charge of TA4 and the hydrophobic branching of KE2 and vice versa for JWN-H9 (Table 2.2). Initial analysis indicated charge as the more important parameter and thus TA4+ was designed by extrapolation of this attribute to probe the validity of this conclusion.

KE2  Ac-NAADLEAAIRHLAEKKAARGPVDAAQQLEQLAKKFQAFARAG-5
JWN-H8 Ac-NAADLEAKIRHLAEKKAARGPVDAAQQLEQLAKKFQAFARAG-NH2
JWN-H9 Ac-NAADJEAAIRHLAEKKAARGPVDAQQLEQLAKKFQAFARAG-5
TA4  Ac-NAADJEAKIRHLAEKKAARGPVDAQQLEQLAKKFQAFARAG-NH2
TA4+ Ac-NAADJEAKIRHLAEKKAARGPRDAAQJAEQLARKFQAFARAG-NH2

Figure 2.6. Representation of differences in charge and side chain branching between peptide scaffolds. Residue charges at experimental conditions are shown in white with border and black for negative and positive respectively. Hydrophobic branching are shown in light gray (leucines (L); 2-branched) and darker gray (norleucines (J); linear). Peptide net charge at pH=7.4 changes from –0.9 (KE2; JWN-H9) to +1.1 (JWN-H8; TA4). Hydrophobic branching is reduced going from leucines (KE2; JWN-H8) to norleucines (JWN-H9; TA4). TA4+ is based on TA4 but has 3 additional positive charges.

2.4.1 Affinity evaluation by SPR

Surface plasmon resonance spectrometry (SPR) is a powerful technique for observing binding phenomena at surfaces. It is a highly sensitive method for detecting refractive index changes in a small space above a gold-surface. On this surface a thin layer of carboxymethylated dextran matrix is immobilized.* The matrix functions as a convenient handle for attachment of proteins, peptides or other species. By activating the carboxyl groups by EDC / NHS chemistry and then flushing the surface with HCAII some of the carboxyl groups from the matrix react with free amines from the protein forming amide bonds and thus irreversible covalent links to the protein.

The sensing range is approximately 150 nm from the surface and the sensitivity can be high enough to detect changes as small as 0.0001° in angle of incidence, corresponding to approximately $10^6$ refractive index units. This is called one response or resonance unit (RU) and is proportional to the adsorbed mass. For typical biomolecules one RU corresponds to approximately 1 pg/mm².

For the experiments discussed in this thesis the typical amount of immobilized HCAII was 2300 RU. This would correspond to approximately 47600 HCAII molecules within the sensor range (150 nm) per µm² or typically 2.5 % volume coverage by protein (Figure 2.7).

---

* This is one type of commercial surface sold by Biacore Sweden; Rapsgatan 7; SE-754 50 Uppsala and is referred to as CM5-type sensor chip.
By flowing peptide solutions at different concentrations over the surface, interactions between HCAII and peptide could be quantified. The response parameter \( r \) is the relative amount of HCAII complexed with peptide and can be calculated from equation 2.1 using the experimentally determined number of response units (RU\text{\_response}). Due to the covalent nature of immobilization some of the HCAII will be nonnative. The fraction of native protein \( n \) and the dissociation constant were used as fitting parameters, employing the Leverberg-Marquadt algorithm on equation 2.2 to search for minimum \( \chi^2 \).

Using the strongest binder \( \text{TA4}^+ \) as a surface quality descriptor and searching for the best fit of equation 2.2 to the experimental results led to the values of \( K_d=227 \ \mu\text{M} \) and \( n=0.9 \) indicating that only 90 % of the immobilized HCAII was available for binding (Figure 2.8). Under the assumption that this is a valid global surface quality descriptor, the peptide – HCAII dissociation constants for all peptides were calculated (Table 2.3).
Equation 2.1\(^a\).  
\[
r = \frac{\text{RU}_{\text{response}}}{\frac{M_{\text{peptide}}}{\text{RU}_{\text{HCAII immobilized}}}} - \frac{M_{\text{HCAII}}}{M_{\text{HCAII}}}
\]
\(^a\)Response parameter (r) calculated from refractive index responses and molecular weights.

Equation 2.2\(^b\).  
\[
K_d = \frac{[E][P]}{[EP]}
\]
\[
E_{\text{act}} = n \times E_{\text{tot}}
\]
\[
\Rightarrow r = \frac{n[P]}{K_d + [P]}
\]
\(\Rightarrow\) Dissociation constant (K\(_d\)) and fraction of native HCAII (n) were used as parameters in fitting function. Enzyme (E) is HCAII and peptide (P) is KE2, JWN-H8, JWN-H9, TA4 or TA4+.

Figure 2.8. Response parameter (r) as function of TA4+ concentration (+). The solid line represents the best fit of equation 2.2 with K\(_d\)=227\(\mu\)M and n=0.9.

**Variable descriptions:** r (the unitless response parameter); RU\(_{\text{response}}\) (response from peptide in refractive units); RU\(_{\text{HCAII immobilized}}\) (amount of immobilized HCAII in refractive units); M\(_{\text{peptide}}\) (peptide molecular weight); M\(_{\text{HCAII}}\) (HCAII molecular weight); K\(_d\) (dissociation constant); [E] (concentration of free enzyme); [P] (concentration of free peptide); [EP] (concentration of peptide – enzyme complex); E\(_{\text{act}}\) (total concentration of active enzyme); E\(_{\text{tot}}\) (total concentration of enzyme); n (native fraction of enzyme).

The results show that neither charge nor hydrophobic branching alone explains the affinity differences. The weak binder KE2 has a K\(_d\) of 12 \(\mu\)M. Replacing the branched L by the linear J leads to JWN-H9 and an observed change in affinity by less than a factor of two, while the addition of two charges to KE2 leads to JWN-H8 and an almost four-fold increase in affinity. The combined effect of changing both charge and hydrophobic branching, leads to TA4 for which the effect is even more pronounced and an affinity increase of a factor of 26 was observed.

<table>
<thead>
<tr>
<th>Table 2.3. HCAII Dissociation constants.</th>
<th>Peptide</th>
<th>K(_d) (mM)</th>
<th>Charge</th>
<th>(\gamma)-branched leucines</th>
</tr>
</thead>
<tbody>
<tr>
<td>KE2</td>
<td>12</td>
<td>-0.9</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>JWN-H8</td>
<td>3.2</td>
<td>+1.1</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>JWN-H9</td>
<td>7.6</td>
<td>-0.9</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>TA4</td>
<td>0.46</td>
<td>+1.1</td>
<td>No</td>
<td></td>
</tr>
<tr>
<td>TA4+</td>
<td>0.23</td>
<td>+4.1</td>
<td>No</td>
<td></td>
</tr>
</tbody>
</table>

Expressing the values in \(\Delta\Delta G^\circ\) clearly indicates that the parameters work cooperatively. Results show 8.0 kJ/mol for the combined effect of charge and branching, markedly
more than the sum of the individual effects being 1.1 and 3.2 kJ/mol for hydrophobic branching and charge respectively.

TA4+ was designed by extrapolation based on the conclusion drawn from the measurements described above that charge was the dominant reason for the higher affinity of TA4 in comparison with that of KE2. The design was based on the sequence of TA4 in which A13, V22 and A37 were replaced by arginines to give a charge of +4.1, three more than in the parent peptide. Even though this peptide was an even better binder of HCAII than all other peptides, the effect was not as pronounced as expected. For the series JWN-H9, TA4 and TA4+ displaying the charges –0.9, +1.1 and +4.1 respectively the corresponding dissociation constants were 7.6, 0.46 and 0.23 mM. Thus changing the charge by two units from –0.9 to +1.1 gave more than a 16-fold increase in affinity while the change of three units from +1.1 to +4.1 gave only an additional two-fold increase. We concluded that with regards to charge magnitude, the peptide TA4+ is nearly optimized and that other modification such as change of charge distribution or hydrophobic character are needed to further increase its affinity.

An alternative model for interpretation of the data reproduced the differences as described in paper II. And thus, we are confident that the applied interpretation model gives us accurate relative affinities of the peptides for HCAII as well as significant differences between the binding abilities of the peptides. The accuracies of the absolute values of dissociation constants are, however, limited and the values should be considered approximate. This has however no implications regarding the $\Delta \Delta G^\text{‡}$ values and cooperativity reasoning.

### 2.4.2 Secondary structure

The helical contents of all peptides were evaluated by circular dichroism spectroscopy (CD) (Table 2.4). The mean residue ellipticity at 222 nm ($[\Theta]_{222}$) is considered a quantifier of helical content where a fully developed helix has $[\Theta]_{222} = -35 700 \text{ deg cm}^2 \text{ dmol}^{-1}$ based on helical poly-L-lysine at pH 11.1\(^\text{18}\).

**Table 2.4.** Mean residue ellipticities of peptide scaffolds.

<table>
<thead>
<tr>
<th>Peptide</th>
<th>$[\Theta]_{222}$ (deg cm(^2) dmol(^{-1}))</th>
<th>Conc. (µM)</th>
<th>Charge</th>
<th>$\gamma$-branched leucines</th>
</tr>
</thead>
<tbody>
<tr>
<td>KE2</td>
<td>-18 600</td>
<td>100(^a)</td>
<td>-0.9</td>
<td>Yes</td>
</tr>
<tr>
<td>JWN-H8</td>
<td>-20 700</td>
<td>100(^b)</td>
<td>+1.1</td>
<td>Yes</td>
</tr>
<tr>
<td>JWN-H9</td>
<td>-24 400</td>
<td>100(^b)</td>
<td>-0.9</td>
<td>No</td>
</tr>
<tr>
<td>TA4</td>
<td>-20 000</td>
<td>250(^c)</td>
<td>+1.1</td>
<td>No</td>
</tr>
<tr>
<td>TA4+</td>
<td>-15 000</td>
<td>100(^b)</td>
<td>+4.1</td>
<td>No</td>
</tr>
</tbody>
</table>

\(^a\)In 10 mM phosphate buffer at pH 7.4 from reference 19. \(^b\)In 20 mM Trizma\(^\text{®}\) buffer at pH 7.4. \(^c\)In 50 mM phosphate buffer at pH 7.4 from reference 20.

For the peptides in which the total charge is small the mean residue ellipticities are close to –20 000 deg cm\(^2\) dmol\(^{-1}\) with JWN-H9 as an exception, possibly due to higher helical
propensity of the unbranched norleucines. This effect however is cancelled in TA4 for no obvious reason. Not surprisingly, TA4+ shows considerably lower helical content than all the other peptides. Since peptides in this series are believed to be helical only when dimerized, a low helical content suggests that TA4+ is dissociated to a high degree. An overall large positive charge might destabilize the structure due to electrostatic repulsion. Furthermore, two alanines, usually considered highly helix inducing, were replaced by arginines.

2.5 Conclusions
To design high affinity conjugates many parameters must be taken into consideration. Only a few of all possible variations have been evaluated so far. Chapter 2.3 concluded that for each peptide scaffold there should exist an optimal length of the linker that might also differ between different templates. A set of linkers should therefore be evaluated for each peptide template. Linker character might also have an impact on both complex stability and morphology.

Peptide charge and hydrophobic branching were evaluated in chapter 2.4 and charge was seen to have a great impact on affinity. To further optimize binding, charge distribution would be a good candidate. There are also many hydrophobic amino acids of different length, branching and type (aliphatic, aromatic) that should be introduced in various positions and configurations to match the complex contour of the host protein.
3 Drug candidate identification and generation

3.1 Generating a lead

In the process of producing drug candidates there are generally a set of lead substances that show potential for evolving into active substances. When starting with a new target, or when the target is unknown it can be advantageous to have access to general lead discovery screening libraries of lead-like or drug-like compounds. To be able to distinguish between drug-like as opposed to nondrug-like substances we have tried to find general structural motifs that seem to lead to biological activity for a broad range of possible targets.

Drug-like properties as defined by Lipinski’s “rule of 5”\(^\text{22}\) states that the molecular weight should be less than 500 g/mol, the calculated logarithm of the octanol-water partition coefficient (logP) should be less than 5, the number of hydrogen-bond donor atoms should be less than five and the hydrogen bond acceptors should be less than 10. It is anticipated that both molecular weight and lipophilicity increases during the drug optimization. In order to make “room” for these modifications a lead-like substance could be defined as having low molecular weight (100-350 g/mol) and relatively low lipophilicity (logP=1-3)\(^\text{23}\).

By searching in the MDDR database\(^\text{24}\) for templates showing structural similarities attributed to biological activity, we have come up with three similar templates that was found in 3084 different substances (Figure 3.1) [Paper III].

![Diagram of templates](image)

673 Hits 2271 Hits 140 Hits

**Figure 3.1.** An overview of hits found in the MDDR database.
These three templates were reduced to an initial set of two scaffolds, A1 and A2, both having 3 points of diversity.

3.1.1 Synthetic strategy of the piperazine scaffolds

The scaffolds A1 and A2 were both synthesized in solution from racemic piperazine-2-carboxylic acid dihydrochloride. Since piperazine-2-carboxylic acid has some adverse properties related to solubility in organic solvents and several points of reaction, orthogonal protection of the carboxylic acid and the two amines was achieved in 78% in a three-step procedure as described by Bigge25. The protected amino acid 1 thus formed can easily be mono deprotected at any point by using different deprotection strategies i.e. alkali, acid or catalytic hydrogenation. By treatment with 4.5 M hydrochloric acid in dioxane / water the BOC protection group was easily removed in quantitative yield to form 2 (Scheme 3.1).

Scheme 3.1. Synthesis of piperazine scaffolds.

\[ \text{NH} \quad \text{HN} \quad \text{OH} \quad \text{O} \quad \text{x 2HCl} \quad \text{i-iii} \quad \text{Z} \quad \text{N} \quad \text{O} \quad \text{BOC} \quad \text{iv} \quad \text{N} \quad \text{N} \quad \text{O} \quad \text{Z} \quad \text{v or vi} \quad \text{O} \quad \text{O} \quad \text{NO}_2 \quad \text{vii} \quad \text{N} \quad \text{HN} \quad \text{O} \quad \text{O} \quad \text{NO}_2 \quad \text{CF}_3\text{SO}_3\text{H} \]

\( \text{1} \) \quad \text{(quant.)} \quad \text{2} \quad \text{3a (o-NO}_2\text{)} \quad \text{3b (m-NO}_2\text{)} \quad \text{4a (o-NO}_2, 87\% \text{)} \quad \text{4b (m-NO}_2, 66\% \text{)}

\[ \text{Reagents and conditions: (i) BOC-ON, pH 11, dioxan/H}_2\text{O; (ii) benzyl chloroformate, pH 9.5; (iii) methyliodide, aqueous sodium bicarbonate, DCM, Adogen 464; (iv) HCl, dioxane; (v) 2-fluoronitrobenzene, TEA, 60°C (for 3a); (vi) Pd(OAc)_2, BINAP, CsCO}_3, 3-bromonitrobenzene, toluene, 100°C (for 3b); (vii) TFMSA, anisole, DCM.} \]

The o-nitro-phenyl group was introduced by nucleophilic aromatic substitution employing 2-fluoronitrobenzene. Several methods were tried including refluxing in aqueous ethanol with sodium hydrogen carbonate and heating in DMF (150°C). However the most successful procedure was stirring a mixture of 1 eq. amine, 2 eq. of 3-fluoronitrobenzene and 2 eq. of triethylamine for 5 days at 60°C or 20 days at rt. The desired aniline scaffold 4a was achieved after acidic deprotection of the Z protection group employing TFMSA in 87% total yield.

The m-nitro-phenyl group could not be introduced in the same manner since 3-fluoronitrobenzene is much less prone to nucleophilic attack than 2-fluoronitrobenzene. However over the last decade, advances in palladium catalyzed aryl aminations mainly achieved by the Buchwald26 and Hartwig27 groups have given us valuable tools to attain these transformations.

Numerous factors, such as choice of phosphine ligand and base excessively affects the outcome. In figure 3.2 a typical catalytic cycle employing BINAP as a bidentate phosphine ligand and cesium carbonate as base is shown. The cycle is initiated by an active monodentate palladium(0) complex (4) that can be formed in several ways from
both palladium(0) and palladium(II) sources. Starting from palladium(II) acetate (1) an initial addition of the phosphine ligand occurs (2) after which the species is reduced (3). An equilibrium exists between the resting bidentate species (3) and the active monodentate palladium(0) species (4). The arylbromide is then oxidatively added to the active species to form complex 5. When a primary or secondary amine is present the complex reorganizes (6) and a base is used to remove a proton from the amine and the bromine ligand from palladium (7). The fate of complex 7 can be either the usually unwanted β-hydride elimination pathway (8) or reductive elimination forming the product. The active complex (4) is thus recycled and ready for another catalytic round.

**Figure 3.2.** The catalytic cycle of palladium catalyzed aryl aminations. For detailed explanations see text.

The palladium catalyzed aryl amination scheme above was used to achieve aniline scaffold 4b in 66% total yield after deprotection.

### 3.1.2 Derivatization of the piperazine scaffolds

The combination of a combinatorial approach with chemistry on solid support is a valuable tool that can produce a diverse set of hundreds or thousands of spatially separated compounds\(^{28,29}\) as well as mixtures of millions of compounds\(^{30,31}\). Hence, 4a and 4b were connected to a Merrifield resin via an activated Wang linker (Figure 3.3) and then reduced to have the first handle ready for derivatization (6a/6b) (Scheme 3.2).
Figure 3.3. Merrifield PS/DVB copolymer resin bead with \( p \)-nitrophenyl activated Wang-linker.

Scheme 3.2\(^a\). Connection to solid support and activation of the scaffold.

\[
\begin{align*}
\text{N} & \text{H} \quad \text{NH}_2 \\
\text{O} & \text{O} \\
\text{CF}_3\text{SO}_3\text{H} & \text{NO}_2 \\
\text{viii} & \quad \text{ix} \\
4a/4b & \quad 5a(0.57 \text{ mmol/g, 81 % load eff.}) \\
& \quad 5b(0.43 \text{ mmol/g, 61 % load eff.}) \\
& \quad 6a/6b
\end{align*}
\]

\(^a\) Reagents and conditions: (viii) DIPEA, DMF, activated PS/DVB-Wang resin; (ix) SnCl\(_2\)·H\(_2\)O, DMF.

The two different solid phase-supported scaffolds (6a, 6b) were each put in 80 spatially separated wells in an 8 rows by 10 columns configuration. To each of the rows a unique electrophile was added, together with pyridine / DCM (Scheme 3.3). When the reaction was complete, the fluid was rinsed off and the next point of diversity was exposed by ester hydrolysis employing potassium trimethylsilanoate in THF. Each column was treated with a unique amine, together with the coupling reagent pyBOP and the base NMM in DMF. The library members were cleaved from the solid support by treatment with trifluoroacetic acid in DCM to yield the final products 10a and 10b.

Scheme 3.3\(^a\). Reactions involved in derivatization.

\[
\begin{align*}
\text{N} & \text{H} \quad \text{NH}_2 \\
\text{O} & \text{O} \\
\text{CF}_3\text{SO}_3\text{H} & \text{NO}_2 \\
\text{x} & \quad \text{xi} \quad \text{xii} \quad \text{xiii} \\
6a/6b & \quad 7a\{1-8\} \\
& \quad 7b\{1-8\} \\
& \quad 8a\{1-8\} \\
& \quad 8b\{1-8\} \\
& \quad 9a\{1-8;1-10\} \\
& \quad 9b\{1-8;1-10\} \\
& \quad 10a\{1-8;1-10\} \\
& \quad 10b\{1-8;1-10\}
\end{align*}
\]

\(^a\) Reagents and conditions: (x) XCl, Pyridine, DCM; (xi) TMSOK, THF; (xii) HY, PyBOP, NMM, DMF; (xiii) TFA, 30 % in DCM. Typical yields for the total reaction 5a or 5b to 10a or 10b was 50-70 %.

The 2 different regioisomers combined with 8 electrophiles and 10 nucleophiles produced a total of 160 compounds, still having one point of diversity left for further substitution possibly for use in lead to drug optimization. (Figure 3.4)
3.2 Rational design of enzyme inhibitor

By studying existing inhibitors of a given enzyme, and by using existing knowledge of enzyme-inhibitor interactions, a rationale for the design of new targeted libraries of potential inhibitors can be generated. Synthesis of these libraries and evaluation of theoretical models in comparison to empirical results further adds to the total knowledge and can be used for further optimization. This completes the cycle that eventually leads to better understanding of general and specific interactions and hopefully more efficient inhibitors.

3.2.1 \(\alpha\)-Thrombin inhibitors

\(\alpha\)-Thrombin, also known as Factor IIa is a serine protease having a central role in the coagulation cascade where the proteolytic capability is used to cleave soluble fibrinogen into fibrin. Furthermore it activates Factor XIII resulting in fibrin cross-linking and it induces platelet activation and aggregation that enables a blood clot to form. In a healthy organism the purpose is to protect against excessive blood loss upon injury. However in several pathological conditions such as hypertension and artery plaque formation unwanted clot formation can lead to lethal conditions such as deep venous thrombosis, stroke, pulmonary embolism and myocardial infarction. For these patients regulation of blood coagulation is imperative, and administration of \(\alpha\)-thrombin inhibitors is a common treatment.

Inhibition/activity is often shown by short peptide sequences that display similar motifs as natural substrates i.e. have a similar peptide sequence as the natural substrate at some key points. These key points include active sites and cofactor binding sites. Bioavailability and metabolic rates of short peptides often limits their suitability as drugs. Oral availability, which is considered crucial, is poor, and metabolic rates are often too high, to be of therapeutic interest. The tri-peptide D-Phe-Pro-Arg (Figure 3.5) that mimics the \(\alpha\)-thrombin sensitive region \((\text{Phe-(AA)₅-Gly-Val-Arg-})\)\(^{32}\) in a loop of the pro-protein fibrinogen has shown inhibitory activity against \(\alpha\)-thrombin.
Knowledge of short peptides showing activity or knowledge of natural substrate configuration in enzyme binding site is a valuable starting point for inhibitor development. Peptidomimetics are small organic molecules that mimic the important structural motifs of these peptide sequences. The advantages over peptides are many, as synthetic constructs introduce many possibilities for structural diversity as well as enhanced bioavailability and metabolic rates.

Known α-thrombin inhibitors include PPACK\textsuperscript{33,34}, L-372.460\textsuperscript{39} and Melagatran\textsuperscript{35} (Figure 3.6) for which the similarities with the tri-peptide D-Phe-Pro-Arg are apparent. Two amide bonds connect three fragments named P3, P2 and P1 from left to right having a proline or proline-like scaffold in the P2 position. The P1-position all have positively charged guanidine, amine or amidine in coherence with the tri-peptide and in the P3-position hydrophobic elements are present.

Amide bonds are a common motif seen in many pharmaceutical substances. They work as both hydrogen bond acceptors and donors, which is an important factor in inhibitor binding\textsuperscript{36,37}. The chemistry around amide bonds is also well known, and there are
numerous commercially available chiral and achiral amines, carboxylic acids and amino acids to use for structural variation.

Analysis of a crystal structure of PPACK bound to α-thrombin active site from the protein databank shows three important binding pockets in the active site of α-thrombin (Figure 3.7). Key amino acids in α-thrombin active site are shown in gray together with PPACK inhibitor (white). The interacting fragments are presented in black. The S2 and D pockets both bind the P2 and P3 fragment of the inhibitor respectively. The interactions are mainly hydrophobic in character and include the side chains of Tyr60A and Trp60D in the S2 pocket as well as Ile174, Trp215 and Leu99 in the D pocket. The S1 / P1 interaction is a very strong ionic bond between the side chain amidine moiety of the inhibitor arginine-chloromethylketone and the Asp189 carboxylic function. Hydrogen bonds are also established between inhibitor and α-thrombin, e.g. Phe-NH₂ ↔ O=C-Gly216 (HB1), Arg-NH ↔ O=C-Ser214 (HB2) and His57-N ↔ CH₂C(Arg)(OH) ↔ O-Ser195 (HB3, a hemiacetal formed by two covalent bonds with the arginine chloromethylketone).

Figure 3.7. Schematic of PPACK bound to α-thrombin active site. For details see text.

* A search in the MDL® Available Chemicals Directory (http://www.chemweb.com/databases) reveals that there are approximately 4500 primary amines, 2500 secondary amines, 9500 carboxylic acids and 500 acid chlorides with a molecular weight below 250 g/mol that are commercially available. Using only 10 % of these will make a possible library of above half a million or half a billion compounds for 2 or 3 points of diversity respectively.
3.2.2 The leads

By reducing structural information from known inhibitors into pseudo-structures that display the important key features, new leads can be generated (Figure 3.8). From these pseudo-structures new targeted libraries can emanate whose potential will be evaluated and added to the total knowledge base for a specific target. This iterative process, if successful, eventually generates inhibitors with drug potential. Out of the eight proline mimicking scaffolds presented, detailed analysis of the cyclopropane scaffold (B) and the $N^4$-amino-morpholinone scaffolds (C1 and C2) with numerous substituents are included in this thesis. Design, synthesis and activity for the templates D1 and D2 has been discussed by Dahlgren at al.\textsuperscript{39} and templates E, F and G has been discussed by Thorstensson et al.\textsuperscript{40}

![Figure 3.8. Deduction of six scaffolds from two general pseudo-structures. The boxes show which scaffolds that are discussed in this thesis.](image)

3.2.3 Mimetics of $d$-Phe-Pro-Arg

The choice of arginine mimetic is restricted by the very strong S1 / P1 bidentate electrostatic interaction, that is crucial for the function of this class of inhibitors. Indeed P1 in known $\alpha$-thrombin inhibitors often include an amidine type structure\textsuperscript{41}. In this design 4-aminomethyl-benzamidine (Pab)\textsuperscript{42} (Figure 3.9) was used.

The proline isoster should enable interaction with the hydrophobic S2 pocket in $\alpha$-thrombin. Three distinct structural motifs as discussed in chapter 3.2.2 were chosen. For each scaffold, numerous hydrophobic P3 substituents were selected to interact with the distal pocket (D) and for their ability to obtain hydrogen bonds in analogy with $d$-phenylalanine in the tripeptide.
3.3 Initial attempts – the cyclopropane scaffold

1,2-Disubstituted cyclopropanes have a very rigid framework lacking the flexibility displayed by larger ring systems. Preliminary studies showed potential for this scaffold with \( \alpha \)-thrombin activities in the low \( \mu \)M range and therefore a more thorough investigation was onset with the production of a targeted library.

3.3.1 Synthesis of the cyclopropane scaffold

Enantiomerically pure cyclopropane-1,2-dicarboxylic acid can be generated by asymmetric alkylation of succinic acid and demonstrates the usefulness of a chiral auxiliary [Unpublished results]. \( \alpha \)-Alkylation of a succinic acid template would give a racemic mixture of preferably trans diacid (Scheme 3.4). However, a chiral auxiliary predominately produces one of the diastereomers\(^{43} \) that can be recrystallized to optical purity of above 99 %.

**Scheme 3.4.**
The use of chiral auxiliary\(^a \) to produce enantiomerically enriched material is shown.

\[
\text{No chiral auxiliary, } X \quad \xrightarrow{\text{Alkylation}} \quad \text{1:1 mixture of trans enantiomers plus some meso-cis form.}
\]

\[
\text{Chiral auxiliary, } X_c=(+)-Menthyl \quad \xrightarrow{\text{Alkylation}} \quad \text{96:4 mixture of trans diastereomers plus some cis form.}
\]

\(^a\)Chiral auxiliary, \( X_c=(+)-Menthyl \)

Succinic acid anhydride was condensed with (+)-menthol in high yield (95 %) using a Dean Stark apparatus (Scheme 3.5). The dienolate was formed by treatment with 2 equivalents of lithium 2,2,6,6-tetramethylpiperidide.

**Scheme 3.5\(^a \).** Synthesis of the cyclopropane scaffold.

\[
\text{11} \quad \xrightarrow{\text{i}} \quad \text{12 (95 %)} \quad \xrightarrow{\text{ii}} \quad \text{13 (26 %)}
\]

\(^a\)Reagents and conditions: (i) (+)-Menthol, (±)-10-camphorsulfonic acid, reflux. toluene; (ii) LTMP, \( \text{CH}_3\text{BrCl} \)
Upon treatment with the electrophile bromochloromethane, addition was achieved on the side with least steric shielding which in the case of the dienolate of (+)-menthylester predominately results in the cyclopropane with the R,R configuration.

3.3.2 Derivatization of cyclopropane scaffolds

Ideally derivatization is achieved with minimum effort starting from a common platform. When a library of similar compounds is needed, the total effort will be less if divergence can be introduced late in the reaction scheme. Since reaction steps might differ in yield, ease and purity there isn’t an absolute correlation between minimum effort and number of steps from the common platform. In this case two different approaches was implemented (Scheme 3.6).

**Method A.** First the di-(+)-menthylester 13 was monohydrolysed using sodium hydroxide in hot isopropanol. The carboxylate was used after purification as the common starting point for the diverse set of substances. Creation of the first amide bond using HATU / DIPEA / DMF and a diverse set of amines was followed by the second ester hydrolysis. After purification the PabZ was coupled using HATU and the synthesis converged with 17 of method B.

**Method B.** The di-(+)-menthyester (13) was first transesterfied to the dimethylester followed by mild mono hydrolysis using 0.9 eq. of lithium hydroxide in dioxane / water which after extraction resulted in a mixture of the diacid and the monomethylester (18). PabZ was coupled using HATU after which a gentle hydrolysis employing lithium hydroxide achieved 20 as the common platform for diversity. The second amide coupling with a diverse set of amines formed the converging products 17.

The main difference between method A and B is the ester used. Method A uses the (+)-menthyl ester, which makes it easy to visualize on tlc and easy to purify by column chromatography. The ester hydrolysis conditions however, are too harsh to be performed in the presence of the Z-protected amidine. Therefore the first amide coupling must be the diverse, and the total steps from the common platform to final products are three.

Method B however employs the methyl ester that can be hydrolyzed under mild conditions. Therefore the second amide coupling can be the diverse, and the number of steps from the common platform is reduced to one. Drawbacks to this scheme are that no effective purification strategy was found for the steps 13-19, and thus the yield of this path is somewhat erratic. The physical properties of 18, 19 and 20 are also adverse being very polar and hard to purify using the straight phase column chromatography, which for many organic chemists can be seen as the purification method of choice.

The conclusion is that neither of method A or B could be seen as more advantageous than the other. Thus both were used as appropriate.
Scheme 3.6. Derivatization of cyclopropane scaffolds.

\[
\begin{align*}
\text{Method B} & \quad \text{Method A} \\
\text{v} & \quad \text{vi, vii} \\
\text{vii} & \quad \text{i} \\
18 & \quad 13 \\
& \quad 14 \\
(60\%) & \quad (14\%) \\
(14\%) & \quad (79\%)
\end{align*}
\]

\[
\begin{align*}
19 & \quad 20 \\
& \quad 17a-h \\
(57\%) & \quad (100\%) \\
& \quad (70-90\%)
\end{align*}
\]

\[X_c = (+)-methyl\]

8Reagents and conditions: (i) NaOH, reflux. iPrOH; (ii) HATU, DIPEA, NHR\(_1\)R\(_2\), DMF; (iii) NaOH, hot iPrOH / water; (iv) HCl(aq); (v) HATU, DIPEA, PabZ×2HCl, DMF; (vi) reflux. MeOH, H\(_2\)SO\(_4\); (vii) LiOH in dioxane / water.

The final targets were synthesized from 17a-h by hydrogenation (Scheme 3.7). After purification using preparative HPLC eight compounds (21a-h) were isolated as acetic acid complexes. Method A was used for 21a-e and method B for 21f-h.

Scheme 3.7. Release of final products.

\[
\begin{align*}
\text{viii} & \quad \text{a} \\
\text{b; c} & \quad \text{d} \\
\text{e} & \quad \text{f; g; h; i; j}
\end{align*}
\]

8Reagents and conditions: (viii) Pd-C (10\%), EtOH, H\(_2\), preparative HPLC MeOH / H\(_2\)O / AcOH.

3.3.3 Inhibitor evaluations

All compounds were screened for \(\alpha\)-thrombin activity. It was disappointing to see that none of the inhibitors reproduced the low \(\mu\)M activity. The best substrate was 21b having an IC\(_{50}\) value of 9\(\mu\)M. Reinvestigation by in silico docking revealed a plausible explanation for the lack of activity (Figure 3.10). In comparison with proline the directional vectors of the two bonds protruding from the ring are quite different in cyclopropane. This renders it almost impossible for the P3 hydrophobic fragment to make favorable interactions with the distal pocket. Furthermore it limits the possibility for a hydrogen bond between the P2-P3 linkage and Gly216 to be established.
This is in agreement with other published α-thrombin inhibitors that utilizes rigid P2-motifs and can be summarized by the P2-directional vector axiom that states that for a potent α-thrombin inhibitor the directional vector of the P2-framework should obey the following rules:

*An sp3 atom or occasionally in certain series an sp2 atom for the group moving into the S1 pocket.*

*An sp2 atom for the group moving into the D pocket.*

*In some cases an sp2/sp2 combination can be viable but only when substituents are in a meta-type correlation and never for an ortho-type correlation.*

Also reevaluation of the initial low μM activity showed that results were misinterpreted due to solubility issues. The actual activity should be interpreted as >1.3μM and thus, as the potential for the cyclopropane scaffold in α-thrombin inhibitors was considered limited, evaluation of other scaffolds were prioritized.

### 3.4 The story continues – the N4-amino-morpholinone scaffolds

By expanding the set of scaffolds with N4-amino-morpholinones we were hoping to reestablish the lost hydrogen bond with Gly216 by introduction of a lone pair located on the hydrazone N4 nitrogen [Paper IV]. The directional vectors are also more in alignment with what is expected to fit nicely into the binding pockets of α-thrombin.
3.4.1 Synthesis of N4-amino-morpholinone scaffolds

Natural L-malic acid (S-form), and the unnatural D-malic acid (R-form) are commercially available chemicals. Starting with the very cheap (S)-malic acid the (S)-form of the scaffold (C2, p.38) will result. Initially esterification was achieved using thionylchloride in methanol, followed by O-allylation employing allylbromide in Ag2O / toluene to afford 24S in excellent yields (Scheme 3.8). Diol formation of the olefin was achieved with osmiumtetraoxide. The diol was then oxidatively cleaved using periodate to afford the aldehyde 25S in 82 % yield. Reductive alkylation of the aldehyde with BOC-protected hydrazine in two steps was followed by lactamization to afford 28S. Worth mentioning is that only extremely slow (days) lactamization was seen using refluxing toluene with or without triethylamine. When hot DMF was used as solvent some unexplored side reaction was dominating. Surprisingly however is that when hot or refluxing water was used for lactamization, not only did the reaction take place in high yield and in a short period of time (hours), it was also accompanied by the desirable BOC-group deprotection.

Scheme 3.8\textsuperscript{a}. Synthesis of N4-amino-morpholinone scaffold.

\[ \text{HO OH} \]
\[ \text{O O} \]
\[ \text{OH O} \]
\[ \text{O O} \]
\[ \text{22S (91\%)} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{23S (91\%)} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{24S (84\%)} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{25S (82 \%)} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{26S (78 \%)} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{27S} \]
\[ \text{O O} \]
\[ \text{O O} \]
\[ \text{28S (89 \%, two steps)} \]

\textsuperscript{a}Reagents and conditions: (i) SOCl\textsubscript{2}, MeOH; (ii) allyl bromide, silver(I) oxide, toluene; (iii) osmium(VIII) oxide, N-methyl morpholine-N-oxide, THF / H\textsubscript{2}O 3:1; (iv) sodium periodate, THF / H\textsubscript{2}O 3:1; (v) hydrazine carboxylic acid tert-butyl ester, toluene, 65 \textdegree C; (vi) H\textsubscript{2} / Pd-C, THF; (vii) reflux H\textsubscript{2}O.

The R-form of the scaffold (28R, C1) was synthesized using the same procedure, but starting from the unnatural (R)-malic acid.

3.4.2 Derivatization of N4-amino-morpholinone scaffolds

The N4-amino-morpholinone nucleophilic amine of scaffolds 28S and 28R were reacted with different electrophiles (Scheme 3.9). Reaction with bromobenzene and 3-bromonitrobenzene was achieved employing palladium-catalyzed aryl aminations as described in chapter 3.1.1 (Table 3.1). Reactions with acylchlorides and sulfonylchlorides were performed in pyridine to achieve the amides and sulfonamides respectively, and reactions with phenylisocyanate and aldehydes were performed in
toluene, forming the urea and imines correspondingly. The imines were further reduced by catalytic hydrogenation to form the alkylated products.

**Scheme 3.9.** Derivatization, general strategy.

![Scheme 3.9 diagram]

Table 3.1. Derivatization of the N-4-amino-morpholinone scaffolds.

<table>
<thead>
<tr>
<th>Cond.</th>
<th>Reagent</th>
<th>R1</th>
<th>R2</th>
<th>Product</th>
</tr>
</thead>
<tbody>
<tr>
<td>viii</td>
<td>3-Bromonitrobenzene</td>
<td>3-Nitrophenyl</td>
<td>H</td>
<td>29a</td>
</tr>
<tr>
<td>viii</td>
<td>Bromobenzene</td>
<td>Phenyl</td>
<td>H</td>
<td>29b</td>
</tr>
<tr>
<td>ix</td>
<td>Benzylic chloride</td>
<td>Benzoyl</td>
<td>H</td>
<td>(R)-29d</td>
</tr>
<tr>
<td>ix</td>
<td>Benzenesulfonyl chloride</td>
<td>Benzenesulfonyl</td>
<td>H</td>
<td>(R)-29f</td>
</tr>
<tr>
<td>ix</td>
<td>Phenylacetyl chloride</td>
<td>Phenylacetyl</td>
<td>H</td>
<td>(R)-29g</td>
</tr>
<tr>
<td>ix</td>
<td>Phenylmethanesulfonyl chloride</td>
<td>Phenylmethanesulfonyl</td>
<td>H</td>
<td>(R)-29h</td>
</tr>
<tr>
<td>x</td>
<td>Phenylisocyanate</td>
<td>N-Anilinocarbonyl</td>
<td>H</td>
<td>(R)-29i</td>
</tr>
<tr>
<td>ix</td>
<td>Isopropyl chloroformate</td>
<td>Isopropoxycarbonyl</td>
<td>H</td>
<td>(R)-29j</td>
</tr>
<tr>
<td>x, xi</td>
<td>Phenyl-acetaldehyde</td>
<td>Phenethyl</td>
<td>H</td>
<td>(R)-29k</td>
</tr>
<tr>
<td>x, xi</td>
<td>3-Phenyl-propionaldehyde</td>
<td>3-Phenyl-propyl</td>
<td>H</td>
<td>(R)-29l</td>
</tr>
<tr>
<td>x, iv</td>
<td>Benzaldehyde + Benzaldehyde dimethylacetal</td>
<td>Benzyl</td>
<td>H</td>
<td>(R)-29m</td>
</tr>
<tr>
<td>xii</td>
<td>Benzylbromide</td>
<td>Benzyl</td>
<td>H</td>
<td>(S)-29o</td>
</tr>
<tr>
<td>ix</td>
<td>2,5-Dimethoxybenzenesulfonyl chloride</td>
<td>2,5-Dimethoxybenzenesulfonyl</td>
<td>H</td>
<td>(S)-29p</td>
</tr>
<tr>
<td>ix</td>
<td>2,4-Difluorobenzenesulfonyl chloride</td>
<td>2,4-Difluorobenzenesulfonyl</td>
<td>H</td>
<td>(S)-29q</td>
</tr>
<tr>
<td>ix</td>
<td>4-Chloro-2,5-dimethylbenzenesulfonyl chloride</td>
<td>4-Chloro-2,5-dimethylbenzenesulfonyl</td>
<td>H</td>
<td>(S)-29r</td>
</tr>
<tr>
<td>ix</td>
<td>2,3-Dihydro-benzofuran-5-sulfonyl chloride</td>
<td>2,3-Dihydro-benzofuran-5-sulfonyl</td>
<td>H</td>
<td>(S)-29t</td>
</tr>
<tr>
<td>x, xi</td>
<td>p-Tolyl acetaldehyde</td>
<td>2-p-Tolyl-ethylamino</td>
<td>H</td>
<td>(S)-29u</td>
</tr>
</tbody>
</table>

Reagents and conditions: (viii) Pd(II) acetate, XANTPHOS, toluene; (ix) pyridine; (x) toluene; (xi) H2 / Pd-C, THF; (xii) DIPEA, NaHCO3, LiI, DMF. Scaffold 28S was used for racemic and S configured products. Scaffold 28R was used for R configured products.

Ester hydrolysis employing lithium hydroxide in methanol / water was followed by coupling with PabZ using HATU and DIPEA in DMF (Scheme 3.10). Deprotection by hydrogenolysis was followed by purification using preparative HPLC to isolate the acetate salt.
Scheme 3.10*. Introduction of P1-moety in N4-amino-morpholinone derivatives.

\[ \text{Scheme 3.10*} \]

Reagents and conditions (xiii) 3 eq. LiOH in MeOH / H2O; (xiv) PabZ, HATU, DIPEA, DMF; (xv) H2 / Pd-C, preparative HPLC in MeOH / H2O / AcOH. *(R)-29e was synthesized by other strategy, described in paper IV. *The aromatic nitro-group of 31a is reduced to amine in this step. *Chlorine in (S)-31r is partially cleaved by hydrogenation leading to (S)-32r and (S)-32s.

The resulting library contained 26 different compounds as potential α-thrombin inhibitors.

### 3.4.3 Inhibition data

In parallel with synthesis of new derivatives, enzyme-binding data was acquired. This is crucial to be able to direct synthetic efforts in an efficient manner and to make the best use of available resources. Therefore inhibitory properties of all compounds directed to α-thrombin inhibition were continuously analyzed. A summary of α-thrombin IC\textsubscript{50}-values for N4-morpholinone substances are shown in table 3.2.

<table>
<thead>
<tr>
<th>Comp. identifier</th>
<th>Structure\textsuperscript{a}</th>
<th>IC\textsubscript{50} (μM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>32a</td>
<td><img src="image" alt="Structure 32a" /></td>
<td>&gt;32</td>
</tr>
<tr>
<td>32b</td>
<td><img src="image" alt="Structure 32b" /></td>
<td>34.4</td>
</tr>
<tr>
<td>32c</td>
<td><img src="image" alt="Structure 32c" /></td>
<td>4.9</td>
</tr>
<tr>
<td>(R)-32d</td>
<td><img src="image" alt="Structure (R)-32d" /></td>
<td>105.7</td>
</tr>
<tr>
<td>(R)-32e</td>
<td><img src="image" alt="Structure (R)-32e" /></td>
<td>54.7</td>
</tr>
<tr>
<td>(R)-32f</td>
<td><img src="image" alt="Structure (R)-32f" /></td>
<td>37.7</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>(S)-32f</td>
<td><img src="image" alt="Structure (S)-32f" /></td>
<td>2.89</td>
</tr>
<tr>
<td>(R)-32g</td>
<td><img src="image" alt="Structure (R)-32g" /></td>
<td>&gt;133</td>
</tr>
<tr>
<td>(R)-32h</td>
<td><img src="image" alt="Structure (R)-32h" /></td>
<td>36.8</td>
</tr>
<tr>
<td>Comp. identifier</td>
<td>Structure</td>
<td>$IC_{50}$ $\mu$M</td>
</tr>
<tr>
<td>------------------</td>
<td>-----------</td>
<td>-----------------</td>
</tr>
<tr>
<td>(S)-32k</td>
<td><img src="image1" alt="Structure" /></td>
<td>1.79</td>
</tr>
<tr>
<td>(R)-32l</td>
<td><img src="image2" alt="Structure" /></td>
<td>45.4</td>
</tr>
<tr>
<td>(S)-32l</td>
<td><img src="image3" alt="Structure" /></td>
<td>4.74</td>
</tr>
<tr>
<td>(R)-32m</td>
<td><img src="image4" alt="Structure" /></td>
<td>36.2</td>
</tr>
<tr>
<td>(R)-32n</td>
<td><img src="image5" alt="Structure" /></td>
<td>13.6</td>
</tr>
<tr>
<td>(S)-32n</td>
<td><img src="image6" alt="Structure" /></td>
<td>10.6</td>
</tr>
<tr>
<td>(S)-32o</td>
<td><img src="image7" alt="Structure" /></td>
<td>19.8</td>
</tr>
</tbody>
</table>

$\text{P}_{\text{ab}} = \text{NH NH NH}_2$.  

### 3.4.4 Structure activity relationships

The N4-amino-morpholinone-scaffold based inhibitors first synthesized (32a through (R)-32m) were all based on the assumption that the preferred stereochemistry would be (R) as this configuration showed the highest affinities for a previously reported morpholinone scaffold\(^{39}\)(See D1 in figure 3.8 p.38). Since the additional nitrogen in the new series was believed to introduce possible hydrogen bond interactions with the C=O of Gly216 in $\alpha$-thrombin it was disappointing to see that the new compounds with comparable length of the P3 side chain showed only modest affinity and in no case was the affinity significantly improved over those of the previous series.

However, docking studies using the Glide\(^{44,45}\) and MacroModel\(^{46}\) modeling software indicated that structures with the (S) stereochemistry would fit in the active site at least as well as the (R) form. Thus some (S) form inhibitors ((S)-32{f,h,k,l}) were synthesized and the affinities were significantly improved showing 4-20 times higher affinity than the corresponding (R) form.

Based on this a new set of 141 potential inhibitors were enumerated using commercially available aldehydes and benzenesulfonyl chlorides. Based on virtual docking employing the Glide software a set of six compounds ((S)-32{p-u}) were selected for synthesis.
together with two compounds (\(\text{(R)}-32\text{n}\) and \(\text{(S)}-32\text{n}\)) that were selected for comparison with the previous scaffold\(^{39}\).

Out of the resulting analogs, \(\text{(S)}-32\text{p}\) having a 2,5-dimethoxybenzenesulfonyl in the P3 position, showed an affinity improvement by more than an order of magnitude giving an IC\(_{50}\) of 0.130 \(\mu\)M. As one can see from the IC\(_{50}\) values, the lipophilic substitution of benzenesulfonamide derivatives is well accommodated in the distal pocket of \(\alpha\)-thrombin giving more than an order of magnitude higher IC\(_{50}\) values for \((\text{(S)}-32\text{p,r,s)}\); IC\(_{50}=0.130; 0.164; 0.247\mu\)M \) compared to that of the unsubstituted \((\text{(S)}-32\text{f; IC}_{50}=2.89\mu\)M). The more polar 2,4-difluoro benzenesulfonyl seems to make less favorable interactions with the hydrophobic pocket \((\text{(S)}-32\text{q; IC}_{50}=8.14\mu\)M).

To gain more insight on the binding of this series of compounds to the active site of \(\alpha\)-thrombin, the most potent inhibitor \((\text{(S)}-32\text{p)}\) was cocrystallized in complex with \(\alpha\)-thrombin and subjected to X-ray analysis. The Connolly surface map of the \(\alpha\)-thrombin – \((\text{S)}-32\text{p}\) complex and an outline of the important interactions between the two are shown in figure 3.11. Analysis of X-ray data show the expected strong salt-bridge between Asp189 and the strongly basic P1-\(p\)-amidinobenzyl with N-O distances of 2.70 and 2.74\(\text{Å}\). There is also a weak hydrogen bond between the N-H of the P1-P2 amide linkage and Ser214 with an N-O distance of 3.15\(\text{Å}\). Moreover the morpholinone ring moves deep into the S2 pocket leading to a significant movement of Tyr60A and Trp60D compared to that of similar inhibitors\(^{47,48,49,50}\). This makes it impossible for the NH of the P2-P3 linkage to reach the C=O of Gly216 and as a result no hydrogen bond could be established. This may also be caused by a potential deprotonation of the NH leading to unfavorable electrostatic repulsion between the negatively charged nitrogen and the electronegative C=O of Gly216. Structural elements as these do show depressed pK\(_{a}\) values e.g. \(N'\)-acetyl-\(N'\)-methyl-benzenesulfonohydrazide have a pK\(_{a}\) of 7.85. The 2,5-dimethoxy benzenesulfonyl group fits nicely into the hydrophobic distal pocket with the 5-methoxy group in a small cavity between Leu99 and Tyr60A.
3.5 Conclusions

The initial library containing cyclopropane scaffolds showed low potency as expected from the lack of previously reported ortho-sp3-sp3 scaffolds as α-thrombin inhibitors. The N4-amino-morpholinone scaffold however resulted in several sub-μM binders and further optimizations might result in compounds with preclinical drug potential.

Even though an additional nitrogen was introduced compared to that of the previously reported morpholinone scaffold no hydrogen bond with Gly216 could be established for (S)-32p. If this is due to the partially negatively charged nitrogen, successful reestablishment of the hydrogen bond might be possible with substrates showing higher pKa. Also substrates with a distinctly different P3-conformation might not move the Tyr60A and Trp60D as much, making the Gly216 accessible for hydrogen bonding. An
interesting substrate is \textbf{(S)-32u} showing an IC$_{50}$ of 0.961µM and probably obeying the two above-mentioned features. If in fact the hydrogen bond is present for this substrate, further optimizations might result in high affinity inhibitors.
4 Future prospects and unpublished results

This is a chapter dedicated to some speculations about what could be done and why. This is my personal thoughts and its scientific plausibility and completeness might be argued. To reflect this, the language is chosen in many cases to be presented in first person.

4.1 A possible route to optimizing a peptide conjugate

In one of our less successful projects I was aiming at self-catalyzed selective modifications of peptides using peptide active esters as substrates. Simple substrates such as $p$-nitrophenyl fumarate can be introduced in this way\textsuperscript{16}, but the success of peptide active esters was limited. Thinking that this might be a sequence dependent problem, an idea arose to generate a library of peptides, and then to find the most appropriate by some “fishing out” strategy. Even though the concept failed due to the fact that the problem wasn’t sequence dependent the method of synthesis is still valid and could easily be adapted to conjugate optimization as described in chapter 2.3.2. The idea was that by a slight reprogramming of the peptide synthesis machine to be able to produce a large number of similar peptides in a combinatorial way.

The original sequence was RYESYGQ with Q being attached to the resin by its side chain rather than the C-terminus. This allowed post synthetic activation forming a $p$-nitrophenyl ester on the C-terminus.

The synthesis protocol was modified so that instead of supplying the peptide synthesis machine with Y, S, E, Y for the 3\textsuperscript{rd}, 4\textsuperscript{th}, 5\textsuperscript{th} and 6\textsuperscript{th} synthesis step I supplied it with equimolar mixtures of the amino acid and alanine. In this way there was not only one peptide, but actually $2^4=16$ peptides in equimolar amount with amino acid 2, 3, 4, 5 being either alanine or Y, S or E. Usually, to perform an amino acid coupling cycle, a four-fold excess of Fmoc protected amino acid is used. This was seen as a potential problem due to the fact that different amino acids couple at different rates. Thus, instead of four equivalents, initially only one equivalent was used (0.5 equivalent of each amino acid). After a prolonged coupling time the resin was rinsed and to complete coupling
another four equivalents (two of each amino acid) was introduced without prior Fmoc deprotection.

The protocol was very successful and all expected fragments could be identified on MALDI. Surprisingly, a simple reverse phase HPLC purification allowed spatial separation of almost every individual fragment (Figure 4.1).

![Figure 4.1](image)

**Figure 4.1.** HPLC chromatogram using a standard reversed phase C8-column. All peaks were identified on MALDI. For each peak a label indicating which amino acids that are substituted for alanine is shown. Numbers are theoretical masses (u) and within parenthesis are experimental deviations (u). It is clear that the substitution Tyr→Ala has greatest impact on reversed phase HPLC mobility. Next is Ser→Ala, and Glu→Ala has the smallest impact. Also isomers of singly Tyr→Ala substituted peptides are separated.

Optimization of the peptide conjugates could be conceived by an adaptation of the protocol mentioned above by starting from a TA4 variant, having the sequence of TA4 with the exception of lysine at position 34 being exchanged for alanine. Consider selecting 6 key points, including position 34 as attachment points for the linker. Introduction of Mtt protected lysines at 1/6th equivalents in these positions leaving 5/6th as the original sequence amino acid could be done by the method described above. This
would in theory create \((1 - \frac{1}{6})^6 = 33\%\)

* of the unmodified mutant (TA4-K34A), \((1 - \frac{1}{6})^5 = 40\%\) of mono modified species (6.7\% of each peptide) and the rest (26\%), would be poly modified. After selective deprotection of the Mtt groups, benzenesulfonamides attached to linkers could be introduced on solid phase. Adding 6 different linkers in equimolar amount first in deficit, then after prolonged reaction time, in excess, would in principle generate \(6 \times 6 = 36\) different mono modified conjugates available in equimolar amounts (approx. 1 \(\mu\)mol each for 100 \(\mu\)mol peptide synthesis scale). As the ligands are mainly hydrophobic it would be plausible to separate the mono modified species from the rest by reversed phase HPLC.

To what use could this conjugate mix be? There are several fishing-out strategies for finding the strongest binders that spring to mind. Two of them are:

* **Mix excess conjugate with HCAII. Separate by centrifugation with a 10000 MW cutoff filter or by gel filtration chromatography.**

* **Use polymer bound HCAII to enrich the most potent inhibitor.**

And to identify the HCAII – conjugate complex one could use:

* **MALDI in combination with proteolytic digestion.**

* **MS/MS fragmentation studies.**

This strategy could also be used to further expand the search for the “perfect fit”. Choosing the best candidate from the initial ligand length/position optimization one can start working on peptide structural variations by the same means.

### 4.2 Selective peptide cleavage N-terminal of cysteine

It would be a great benefit to expand the range of proteases and chemical agents that selectively cleave peptides or proteins. Well, I did just that – mostly by accident.

In an effort to produce peptides with an ability to transfer acyl groups to a designated target (such as a protein), we designed a peptide named JWN-CYS1. The peptide is a 42 amino acid residue helix-loop-helix motif containing one lysine residue with an ability to carry an anchoring group for a specific protein and one cysteine residue with a potentially volatile thioester that can transfer acyl-groups to the target by transesterification/transamidation. This doubly acylated species, JWN-CYS1[diBSAC6], is depicted in figure 4.2. When absence of function was observed, I tried to hydrolyze

---

* A general formula of the fraction of product with \(x\) modifications when \(n\) points are modified to the fraction of \(p\) is \((1 - p)^{n-x} p^x \frac{n!}{x!(n-x)!}\). If one modification is desired \((x=1)\) the formula is reduced to \((1 - p)^{n-1} p \cdot \frac{n!}{x!(n-x)!}\) and maximal yield is given at the function maximum which is \(p=1/n\). For the example given where \(n=6\), the optimal fraction of introduced Mtt protected lysine should thus be 1/6\(^{th}\).
the formed thioester. Remarkably it withstand both hydroxide (pH=10) and even hot nucleophiles such as hydrazine at pH=9 without being cleaved.

![Figure 4.2. The original peptide JWN-CYS1[diBSAC6].](image1)

However when the pH was raised to 12.2, things started to happen after 30 minutes. Some of the thioester was cleaved as indicated on MALDI, but the major products were cleaved peptide containing two fragments whose mass was identified to be cleavage products from cysteine N-terminal cleavage (Figure 4.3). Evidently the found fragments both contained the original sidechain modification. However it is likely that it has migrated to the terminal amine forming an amide for fragment carrying the N-terminal cysteine.

<table>
<thead>
<tr>
<th>Fragment</th>
<th><img src="image2" alt="Fragment" /></th>
</tr>
</thead>
<tbody>
<tr>
<td>Formula</td>
<td>C_{62}H_{91}N_{19}O_{16}S_{2}</td>
</tr>
</tbody>
</table>

![Figure 4.3. Cleavage fragments of peptide JWN-CYS1[diBSAC6].](image3)

The applications of this might seem limited but it is an unexpected and interesting feature. Experiments indicate that the cleavage will not occur unless the cysteine is acylated. This might be taken advantage of to further specify cleavage sites. Concerning the generality of this procedure, it has been shown to work for at least one other peptide.
At these high pH values there will probably be some amino acid racemization but when used as an analysis tool this might not be important.
5 Notations and abbreviations

5.1 Peptide notations
In general, peptides are named as NAME-M(pos)-M(pos)

NAME is the peptide abbreviation, which can be one of:

- KE2 (Ac-NAADLEAARHLAELAARGPVDAAQLAEQQLAKKFEAFARAG-OH)
- TA4 (Ac-NAADJEAIKHLAEKJAARGPVDAAQLAEQQLARKFERFAgar-NH3)
- TA4+ (Ac-NAADJEAIKHLREKJAARGPRDAAQJAEQQLARKFERFERFARAG-NH3)
- JWN-H8 (Ac-NAADLEAIKHLAEKLAARGPVDAAQLAEQQLAKKFEAFARAG-NH3)
- JWN-H9 (Ac-NAADJEAIKHLAEKJAARGPVDAAQLAEQQLAKKFEAFARAG-OH)

M(pos) indicates sidechain modification at position pos with modifier M, where M can be:

- DNS (Dansylation)
- DAB (Dabcylation)
- C0 (4-Sulfamoyl-benzoylation)
- C2 (2-(4-Sulfamoyl-benzoylamino)-acetylation)
- C4 (4-(4-Sulfamoyl-benzoylamino)-butyrylation)
- C6 (6-(4-Sulfamoyl-benzoylamino)-hexanoylation)
- C7 (7-(4-Sulfamoyl-benzoylamino)-heptanoylation)
- C8 (8-(4-Sulfamoyl-benzoylamino)-caprylation)
- C11 (11-(4-Sulfamoyl-benzoylamino)-undecanoylation)

Modifications of peptides KE2 and TA4 are mentioned in this thesis. The modifications are DNS or DAB (position 15, shaded in light gray) and C0, C2, C4, C6, C7, C8 or C11 (position 34, shaded in dark gray).

Peptide notations – Quick reference
KE2 (Ac-NAADLEAIKHLAEKLAARGPVDAAQLAEQQLAKKFEAFARAG-OH)
TA4 (Ac-NAADJEAIKHLAEKJAARGPVDAAQLAEQQLARKFERFAgar-NH3)
TA4+ (Ac-NAADJEAIKHLREKJAARGPRDAAQJAEQQLARKFERFERFARAG-NH3)
JWN-H8 (Ac-NAADLEAIKHLAEKLAARGPVDAAQLAEQQLAKKFEAFARAG-NH3)
JWN-H9 (Ac-NAADJEAIKHLAEKJAARGPVDAAQLAEQQLAKKFEAFARAG-OH)

Not. Modification Substructure
C0 4-Sulfamoyl-benzoylation
C2 2-(4-Sulfamoyl-benzoylamino)-acetylation
C4 4-(4-Sulfamoyl-benzoylamino)-butyrylation
C6 6-(4-Sulfamoyl-benzoylamino)-hexanoylation
C7 7-(4-Sulfamoyl-benzoylamino)-heptanoylation
C8 8-(4-Sulfamoyl-benzoylamino)-caprylation
C11 11-(4-Sulfamoyl-benzoylamino)-undecanoylation

DNS (Dansylation)
DAB (Dabcylation)

Sample Conjugate

Ac-NAADLEAARHLAELAARGPVDAAQLAEQQLAKKFEAFARAG-OH
KE2-DNS(15)-C6(34)
5 Notations and abbreviations

5.1 Peptide notations

In general peptides are named as NAME-M(pos)-M(pos)

NAME is the peptide abbreviation, which can be one of:

KE2  \(\text{Ac-NADJEAKIRHLAEKLAARGPVDAAQJAEQLARKFEAFARAG-NH}_2\)

TA4  \(\text{Ac-NAADJEAKIRHLAEKLAARGPVDAAQJAEQLARKFEAFARAG-OH}\)

T4A+  \(\text{Ac-NAADJEAKIRHLAEKLAARGPVDAAQJAEQLARKFARAG-NH}_2\)

JWN  \(\text{Ac-NAADJEAAIRHLAEKLAARGPVDAAQJAEQLARKFEAFARAG-OH}\)

JWN-H9  \(\text{Ac-NAADJEAKIRHLAEKLAARGPVDAAQJAEQLARKFEAFARAG-OH}\)

M(pos) indicates sidechain modification at position pos with modifier M, where M can be:

DNS (Dansylation)

BOC (Boc-ON)

DCC (N,N-Dicyclohexycarbodiimide)

DCM (Dichloromethane CH2Cl2)

DMSO (Dimethyl sulfoxide)

DMF (N,N-Dimethylformamide)

DMPU (1,3-Dimethyl-3,4,5,6-tetrahydro-2(1H)-pyrimidinone)

DNS (Dansyl group)

HATU (O-(7-Azabenzotriazol-1-yl)-N,N,N',N''-tetramethyluronium hexafluorophosphate)

LDA (Lithium diisopropylamide Li+ -N(iPr)2)

LTMP (Lithium 2,2,6,6-tetramethylpiperidide)

NMM (N-Methyl morpholine)

Pab (4-Aminomethyl-benzamidine also known as β-amino benzylamine)

PabZ (4-Aminomethyl-benzamidine also known as β-amino benzylamine)

PS/DVB (Polystyrene/divinylbenzene copolymer)

pyBOP (Benzotriazol-1-yl)-N,N,N',N''-tetramethyluronium hexafluorophosphate)

TEA (Triethylamine)

TFA (Trifluoroacetic acid)

TMSA (Trifluoromethanesulfonic acid)

THF (Tetrahydrofuran)
<table>
<thead>
<tr>
<th>Abbrev.</th>
<th>Chemical Name</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>TMSOK</td>
<td>Potassium trimethylsilanolate</td>
<td><img src="TMSOK.png" alt="Structure" /></td>
</tr>
<tr>
<td>XANTPHOS</td>
<td>9,9-dimethyl-4,5-bis(diphenyl-phosphino)xanthene</td>
<td><img src="XANTPHOS.png" alt="Structure" /></td>
</tr>
<tr>
<td>Z</td>
<td>Benzyloxy carbonyl protection group</td>
<td><img src="Z.png" alt="Structure" /></td>
</tr>
</tbody>
</table>
6 References

13 RCSB Protein Data Bank. Entries 1CNW; 1CNX; 1CNY; 1IF4; 1IF5; 1IF6; 1KWQ; 1OKM; 1OKN; 1OQ5 and 3CA2. http://www.rcsb.org/pdb/.


The Protein Data Bank is operated by the Research Collaboratory for Structural Bioinformatics (RCSB). Database information: PDB ID: 1ABJ; http://www.rcsb.org/pdb/.


Acknowledgements

Thank you…

Lars Baltzer, my main supervisor during the last three years, for sharing your knowledge and enthusiasm with me, and for the wonderful espresso.

Ingemar Kvarnström my main supervisor during the “thrombin era” (1998-2002), for sharing your great knowledge in organic synthesis with me.

Ingemar Nilsson, for all help, and for giving me of your time when you had little to spend. Thank you for your support and helpfulness when I was stationed at AstraZeneca in Mölndahl.

Bertil Samuelsson, for novel ideas about synthetic strategies.

For participation in joint projects with me…

Fredrik Thorstensson for sharing my first PhD time and publication, for being a friend and for your fierce persuasion into hanging out on bars instead of staying home with my family.

Karin Enander, our group’s fluorescence pioneer, for sharing paper I with me and for all the interesting discussions.

Martin Lundquist, NMR in paper I, and for my whisky connoisseur training.

Gunnar Dolphin, for sharing paper I with me and for always being so willing to help.

Tess Andersson for sharing paper I with me, and your positive attitude.

Nalle Jonsson for biochemistry discussions.

Djordje Musil, for doing the X-ray crystallography in paper IV.

Tudor Oprea, for your input on paper III.
Special thanks for helps on particular matters…

Shenghua Huang and Elisabeth Sauer-Eriksson (conjugate – HCAII complex crystallization attempts)

Johanna Deinum, Karolina Åkesson and Olle Karlsson (Thrombin screening)

Karl-Erik ”Kirre” Karlsson (High Resolution MS)

Shalini Andersson (chiral HPLC)

Stefan Svensson and Per Hammaström (for oral discussions)

Thanks to present and former colleagues not mentioned above. For sharing discussions, knowledge and free time activities with me…

Alina, Andreas, Cissi, Gunnar H., Helena, Janosch, Jesus, Johan R., Johan V., Kerstin, Klas, Laila, Lotta, Maria, Patrik, Sofia and Susanne in the espresso corridors.

Anders, Daniel, David, Per-Ola, Åsa in the old organic chemistry.

I have the great opportunity to share vacations, and weekends with my wonderful family…

My parents, Kenneth and Gunnel and my siblings Fredrik and Jenny and their families.

And finally for your love and support, and for giving my life purpose…

My beloved sons Adam and Erik and my Lotta.

Thank you all!