Linköping Studies in Science and Technology Dissertations No. 990

## Structure-Based Design and Synthesis of Protease Inhibitors Using Cycloalkenes as Proline Bioisosteres and Combinatorial Syntheses of a Targeted Library

**Fredrik Thorstensson** 



## Linköpings universitet

Division of Chemistry Department of Physics, Chemistry and Biology Linköpings universitet, SE-581 83 Linköping, Sweden

Linköping 2005

**Cover art.** Connolly surface map of the X-ray structure of the -thrombincyclopentenedicarboxylic acid derivative complex (yellow) at 1.9 Å resolution superimposed on the X-ray structure of the -thrombin-PPACK complex (magenta) (Thorstensson, F.; et al. *J. Med. Chem.* **2003**, *46*, 1165-1179).

Thesis cover art was also on the cover of each issue of the Journal of Medicinal Chemistry beginning July 2003 and continuing through December 2003 with the motivation: The basis for our selection was a combination of esthetics and good science.

© 2005 Fredrik Thorstensson

ISBN 91-85457-78-7

ISSN 0345-7524

Printed in Sweden by LiuTryck

Linköping 2005

#### Abstract

Structure-based drug design and combinatorial chemistry play important roles in the search for new drugs, and both these elements of medicinal chemistry were included in the present studies. This thesis outlines the synthesis of protease inhibitors against thrombin and the HCV NS3 protease, as well as the synthesis of a combinatorial library using solid phase chemistry. In the current work potent thrombin inhibitors were generated based on the D-Phe-Pro-Arg motif incorporating cyclopentene and cyclohexene scaffolds that were synthesized by ringclosing metathesis chemistry. A structure-activity relationship study was carried out using the crystallographic results for one of the inhibitors co-crystallized with thrombin. HCV NS3 protease inhibitors comprising the proline bioisostere 4-hydroxy-cyclopent-2-ene-1,2dicarbboxylic acid were synthesized displaying low nanomolar activity. The stereochemistry and regiochemistry of the scaffolds were determined by NOESY and HMBC spectra, respectively. The final diastereometic target compounds were isolated and annotated by applying TOCSY and ROESY NMR experiments. Furthermore, a 4-phenyl-2-carboxypiperazine targeted combinatorial chemistry library was synthesized to be used early in the lead discovery phase. This was done using a scaffold that was synthesized by palladiumcatalyzed aromatic amination chemistry and subsequently derivatized with eight electrophiles and ten nucleophiles.

## Contents

Contents	iii
List of papers	v
Abbreviations	vii
1 Introduction	1
2 Synthesis of a combinatorial library (Paper I)	3
2.1 Synthesis of the scaffolds	4
2.2 Synthesis of the combinatorial library	6
2.3 Purity and yields	7
2.4 Conclusions	8
3 Serine proteases	9
4 Thrombin protease inhibitors (Paper II)	12
4.1 The development of thrombin inhibitors based on the D-Phe-Pro-Arg motif	14
4.2 Synthesis of the scaffolds	17
4.3 The olefin metathesis reaction	19
4.3.1 Metathesis trials	22
4.4 Synthesis of the final products	24
4.5 Structure-activity relationships	27
4.6 Conclusions	30
5 Inhibitors of the HCV NS3 protease (Paper III)	31
5.1 Development of NS5A/5B substrate-based inhibitors of the HCV NS3 protease	34
5.2 Synthesis of the scaffold	37
5.2.1 Structure determination of the scaffold using 2D-NMR	39
5.3 Synthesis of the final products	41
5.3.1 Chemical stability	45
5.4 Structure determination of the target compounds using 2D-NMR	46
5.5 Structure-activity relationships of the final products	50
5.6 Conclusions	52
6 Acknowledgments	53
7 References	55

## List of papers

This thesis is based on the following papers, which are referred to in the text by their Roman numerals:

I. Solid-phase synthesis of libraries generated from a 4-phenyl-2-carboxypiperazine scaffold.

Jonas W. Nilsson, <u>Fredrik Thorstensson</u>, Ingemar Kvarnström, Tudor Oprea, Bertil Samuelsson and Ingemar Nilsson. *J. Comb. Chem.*, **2001**, *3*, 546-553.

II. Synthesis of Novel Thrombin Inhibitors. Use of Ring-Closing Metathesis Reactions for Synthesis of P2 Cyclopentene and Cyclohexene Dicarboxylic Acid Derivatives.

<u>Fredrik Thorstensson</u>, Ingemar Kvarnström, Djordje Musil, Ingemar Nilsson and Bertil Samuelsson. *J. Med. Chem.* **2003**, *46*, 1165-1179.

# III. Synthesis of Novel Potent Hepatitis C Virus NS3 Protease Inhibitors. Discovery of 4-Hydroxy-cyclopent-2-ene-1,2-dicarboxylic Acid as a L-Hydroxyproline Bioisostere in the P2-Position.

<u>Fredrik Thorstensson</u>, Fredrik Wångsell, Ingemar Kvarnström, Lotta Vrang, Elizabeth Hamelink, Anders Hallberg, Åsa Rosenquist, Bertil Samuelsson. Submitted to *J. Med. Chem.* 

## Abbreviations

BINAP	2,2'-bis(diphenylphosphino)-1,1'-binaphthyl			
Boc <sub>2</sub> O	di-tert-butyl dicarbonate			
Boc-ON	2-(tert-butoxycarbonyloxyimino)-2-phenylacetonitrile			
BOP-Cl	bis-(2-oxo-3-oxazolidinyl)-phosphinic-chloride			
СМ	cross metathesis			
dba	dibenzylideneacetone			
DCC	dicyclohexyl carbodiimide			
DCM	dichloromethane			
DIPEA	N,N'-diisopropylethylamine			
DMAP	4-dimethylaminopyridine			
DMPU	N,N'-dimethylpropyleneurea			
EC <sub>50</sub>	median effective concentration (required to induce a 50% effect in a cell			
	based assay)			
EDC	N-(3-dimethylaminopropyl)-N-ethylcarbodiimide hydrochloride			
Fmoc	9-fluorenylmethyl carbamate			
Fmoc-ONsu	N-(9-fluorenylmethoxycarbonyloxy)-succinimide			
HATU	O-(7-azabenzotriazol-1-yl)-N,N,N,N-tetramethyluronium			
	hexafluorophosphate			
HMBC	heteronuclear multiple bond correlation			
HMPA	hexamethylphosphoramide			
HOBt	1-hydroxy benzotriazole			
HPLC	high performance liquid chromtography			
HTS	high-throughput screening			
IC <sub>50</sub>	inhibitor concentration that causes a 50 % decrease in enzyme activity			
LC	liquid chromatography			
LDA	lithium diisopropylamide			
MS	mass spectroscopy			
NHC	N-heterocyclic carbene			
NMM	N-methyl morpholine			
NOESY	nuclear Overhauser effect spectroscopy			
Pab	<i>p</i> -amidino benzylamine			
p-TsOH	para-toluenesulfonic acid			
r.t.	room temperature			
RCM	ring-closing metathesis			
ROESY	rotational frame Noe spectroscopy			
ROM	ring-opening metathesis			

SAR	structure-activity relationship
TBTA	tert-butyl 2,2,2-tri-chlooacetimidate
TES	triethylsilane
TFA	trifluoroacetic acid
TFMSA	trifluoromethane sulfonic acid
TMEDA	<i>N</i> , <i>N</i> , <i>N</i> ', <i>N</i> '-tetramethylethylenediamine
TOCSY	total correlation spectroscopy
UV	ultraviolet
Ζ	benzyloxycarbonyl

## **1** Introduction

Drug discovery is a complex undertaking, and it is not possible to present straightforward simple flowcharts that adequately describe this process. The task includes intricate interplay between biology and chemistry, or more precisely, application of organic chemistry and the more specialized field called medicinal chemistry in combination with biochemical and biological tests (Figure 1.1).

In the **target identification** step, researchers determine the therapeutic focus, e.g. a specific enzyme or receptor that plays a vital role in the disease or condition in question. Many pathophysiological conditions and disease states are due to over expression or enhanced activity of a specific enzyme, *e.g.* protease, and this activity can be attenuated by administration of a specific protease inhibitor. In the design of such protease inhibitors one strategy utilizes the knowledge of the natural substrates from which **substrate-based** protease

Target identification         Paper I: Several targets chosen         Paper II: The serine protease thrombin         Paper II: The serine protease thrombin         Paper III and IV: The serine protease HCV NS3						
Substrate based analogs Paper II and III		HTS on libraries Paper I				
In Vitro tests						
Hit to lead identification Paper II: Thrombin. D-Phe-Pro-Arg Paper I: Combinatorial chemistry. Confidential screening data Paper III: HCV. Chg-tLeu-Hyp(7-methoxy-2-phenylquinolin-4-ol)-Acca						
Lead compound Paper III	In vitro	o tests	Computer aided refinement and prediction. (X-Ray)			
Lead optimization Paper II						
Biological tests						





Figure 1.2. The development of a candidate drug to a marketable medicine.

inhibitors can be designed (papers II-III). These molecules mimic the natural substrate of the target, and in most cases the starting point of synthesis is a truncated polypeptide (e.g., a hexapeptide) in which all amino acid positions are systematically varied to gain increased knowledge of SAR, *i.e.* specificity and selectivity. Another approach that attracted much attention in the 1990s involves the synthesis of large compound libraries (comprising several million members) by the use of combinatorial chemistry. These libraries are tested (screened) against numerous targets utilizing HTS (High Throughput Screening) techniques to deliver active compounds, Hits (paper I). Promising Hit compounds identified are then subjected to additional *in vitro* tests, and if still promising are optimized by chemical modifications based on the preliminary SAR established (e.g., by synthesis of a new library or some new polypeptides) (hit to lead identification phase). The goal of this iteration between synthesis and *in vitro* tests is to generate a **lead compound**. A lead compound shall exhibit high affinity for its target, have good metabolism and pharmacokinetic properties and include the basic pharmacophore. The pharmacophore is the correct spatial arrangement of chemical groups necessary to achieve the desired interaction with the target. The basic pharmacophore can be looked upon as a crude blueprint, in other words its structure includes vital features but is far from perfect. Refinements in the Hit-to-Lead phase is greatly aided by **X-ray** crystal data of an inhibitor in complex with the target protease. In the subsequent lead optimization phase the synthesis of new compounds is aided by additional SAR gained from extensive biological testings, and when a compound performs well in all biological tests, it can be nominated to a candidate drug and large scale synthesis of the active ingredient is performed before testing in humans can begin (Figure 1.2). In the first stage in human subjects, called **Phase I**, the drug is tested on healthy volunteers, and, if no complications occur, the testing is continued in **Phase II** studies, which include a few hundred patients with the disorder of interest. The last step entails the **Phase III** investigations, which include several thousand patients. If the compound exhibits better clinical properties than drugs that are already on the market, it is time to submit a **new drug application**, and hopefully the novel substance will be approved for clinical use by the appropriate institution (e.g., the Food and Drug Administration (FDA) in the United States). On average, the entire process from selecting a target disease to market analysis and the release of an approved drug takes about 10 to 15 years and costs approximately 8 billion SEK.<sup>1</sup>

#### 2 Synthesis of a combinatorial library (Paper I)

Potential lead compounds can be discovered by synthesizing numerous well-defined combinatorial libraries and then using HTS aimed against many different targets (*e.g.*, HIV protease, thrombin, and HCV NS3 protease). The objective of such screening is hit identification, that is, to discover one or several compounds in the synthesized library that exhibit activity against one of the chosen targets and can hopefully be refined to a lead.

In a project originated at AstraZeneca the generic scaffold S shown in Figure 2.1 was considered to be a therapeutically interesting template that was amenable to library exploration in the lead identification phase.<sup>2</sup> Although members of S already exist that have been shown to display antibacterial and oxytocin receptor-antagonistic activities, scaffold S



Figure 2.1. The generic scaffold S and the two specific scaffolds a and b.

donors or more than 10 hydrogen bond acceptors, nor should they have a molecular weight exceeding 500 g/mol or a calculated Log P value  $\geq 5$  in octanol/water. A compound can have good absorptive properties if it meets three of the four criteria, but probably not if it satisfies only one or two. The strategy used to synthesize the library was to prepare the regioisomeric scaffolds **1a** and **1b** by standard solution-phase still represent a interesting framework for pharmaceutical development.<sup>3,4</sup> From **S**, we selected the two specific scaffolds **a** and **b** for synthesis of a small library of drug-like compounds that could be used in biological evaluations in the early lead discovery phase. It is important that such a library follows the empirical "Lipinski's rule of five" concerning permeation and solubility.<sup>5</sup> It postulates that drug-like molecules should not have more than five hydrogen bond



**Figure 2.2.** Immobilized scaffold **b** with the two diversification points indicated.

chemistry (Scheme 2.1). These scaffolds were subsequently immobilized on a solid phase and combinatorially diversified with eight electrophiles and ten nucleophiles to create a library consisting of 80 x 2 = 160 compounds. The solid phase consists of a polymer (resin) with attached linkers to which a scaffold can be coupled, and we chose to use carboxy-activated wang resin (Figure 2.2). Solid-phase chemistry allows time-efficient purification, hence many

compounds can be synthesized. The easy purification procedure, which consists of sequential rinsing of the resin with different solvents, makes it possible to use excessive amounts of reagents (10–30 equivalents) and thereby increase the reaction rates and raise the yields. The drawback of solid-phase chemistry is that commonly employed monitoring techniques, such as TLC and HPLC/MS, can't be used, and it is instead necessary to perform cleavage of the scaffold followed by solution-phase analyses, or to conduct NMR experiments directly on the resin. This disadvantage is a problem only when trying out the reaction sequence, not when a successful protocol has been worked out.

#### 2.1 Synthesis of the scaffolds

The acid 2 amino was selectively monoprotected with Boc-ON (2-(tertbutoxycarbonyloxyimino)-2-phenylacetonitrile) in dioxane/water, and the pH of the solution was adjusted to 11 with aqueous sodium hydroxide (Scheme 2.1). The pH was then lowered to 9.5 with hydrochloric acid, and Z-Cl (benzyl chloroformate) was added. Thereafter, the diprotected piperazine-2-carboxylic acid was esterified with iodomethane and aqueous sodium bicarbonate under phase-transfer conditions in DCM (dichloromethane), which gave the fully protected piperazine-2-carboxylic acid 3 in 78% yield.<sup>6</sup> The tert-butoxycarbonyl group was cleaved to yield 4 by treating 3 with hydrochloric acid in dioxane. Reacting Zprotected piperazine 4 with o-fluoro-nitrobenzene for five days gave 5a in 95% yield.

Scheme 2.1. Synthesis of scaffolds 1a and 1b.<sup>a</sup>



<sup>a</sup> **Reagents and conditions:** (a) Boc-ON, pH 11, dioxane/H<sub>2</sub>O 1:1; (b) Z-Cl, pH 9.5; (c) CH<sub>3</sub>I, NaHCO<sub>3</sub> (aq), adogen 464, DCM; (d) HCl, dioxane; (e) 2-fluoronitrobenzene, TEA, 60°C; (f) Pd(OAc)<sub>2</sub>, BINAP, CsCO<sub>3</sub>, 3-bromonitrobenzene, toluene, 100 °C; (g) TFMSA, anisole, DCM.

However, as expected, direct nucleophilic aromatic substitution using the corresponding *m*-fluoro-nitrobenzene was unsuccessful, because the *m*-nitro substituent has a less prominent deactivating effect. Using the palladium chemistry described below, **5b** was synthesized in 78% yield, and Z-deprotection of **5a** and **5b** using TFMSA (trifluoromethane sulfonic acid) in DCM gave the free amines **1a** and **1b** in 92% and 85% yield, respectively.

Advances in palladium-catalyzed aromatic aminations achieved by Stephen L. Buchwald and John F. Hartwig over the last decade have enabled chemists to synthesize compounds that can't be synthesized by direct nucleophilic aromatic substitutions.<sup>7,8</sup> There are many viable palladium sources (*e.g.*, PdCl<sub>2</sub>, Pd<sub>2</sub>(dba)<sub>3</sub>, and Pd(Ph<sub>3</sub>)<sub>4</sub>), and there is also a multitude of different phosphines (*e.g.*, PPh<sub>3</sub>, Xantphos, Potol<sub>3</sub> and PBu<sub>3</sub>). The catalytic cycle using Pd(OAc)<sub>2</sub> as a source of Pd(0) generated *in situ* with BINAP as a phosphine ligand is illustrated in Figure 2.3 (Phase 1). The monodentate Pd(0) species achieves an oxidative addition to the aryl bromide, creating the (L)Pd(Ar)(Br) complex (Phase 2). The amine coordinates to the (L)Pd(Ar)(Br) complex (Phase 3), and the addition of base facilitates deprotonation of the amine and elimination of the bromide to generate the amido complex



Figure 2.3. The catalytic cycle for the palladium-catalyzed aromatic aminations.

(L)Pd(Ar)[NR(R')] (Phase 4). This complex is decomposed through reductive elimination to yield the target compound (ArN(R)R') and the regenerated phosphine-palladium complex (Phase 5).

#### 2.2 Synthesis of the combinatorial library

The scaffolds **1a** and **1b** were immobilized on the *p*-nitrophenyl-carbonate-activated Wang resin using DIPEA (N,N-diisopropylethylamine) in DMF (dimethylformamide) (Scheme 2.2). The loading (81% for **6a** and 61% for **6b**) was rewarding, considering the poor nucleophilicity of the NI of piperazine-2-carboxylates. The practical loading was determined by cleavage of a weighed sample of resin followed by analysis on HPLC-UV calibrated with

Scheme 2.2. Solid-phase chemistry.<sup>a</sup>



<sup>a</sup> Reagents and conditions: (a) *p*-nitrophenyl carbonate Wang resin, DIPEA, DMF, 80 °C (b) SnCl<sub>2</sub>·H<sub>2</sub>O, DMF;
(c) Chemset 8 (R<sub>3</sub>XCl, X = SO<sub>2</sub> or CO), pyridine, DCM; (d) TMSOK, THF; (e) Chemset 11 (NHR<sub>4</sub>R<sub>5</sub>), PyBOP, NMM, DMF; (f) TFA, 30% in DCM.



Figure 2.5. The electrophiles in chemset 8.



Figure 2.6. The nucleophiles in chemset 11.



Figure 2.4. Diagrams showing a reservoir containing the resin (*left*), reservoirs fitted into a rack (*middle*), and the rack viewed from above (*right*).

known solutions of scaffolds. Longer reaction times did not improve the loading. Several methods have been reported for reduction of nitro compounds on solid support.<sup>9,10,11</sup> Addition of SnCl<sub>2</sub>·H<sub>2</sub>O in DMF to **6a** and **6b** resulted in a clean reduction to the anilines **7a** and **7b**. with no resin cleavage. Premature cleavage of the Wang linker has been reported using SnCl<sub>2</sub>·H<sub>2</sub>O due to formation of hydrogen chloride and/or Sn(IV) with Lewis acid capabilities.<sup>12</sup> The scaffold 7a was transferred to 80 reservoirs (A–H, 1–10) placed in a rack as shown in Figure 2.4, and 7b was transferred to a similar rack. The first diversification of the library was done by adding the electrophiles from chemset 8 in DCM/pyridine to the rows (A–H) in the two racks, which furnished  $8 \times 2 = 16$  different compounds (Figure 2.5). The resulting set of esters **9a** and **9b** were hydrolyzed to the corresponding carboxylic acids **10** using potassium trimethylsilanoate in THF (tetrahydrofuran). The second diversification of the library was achieved by coupling the amines in chemset 11 to the free acids 10a and 10b using pyBOP (benzotriazol-1-yl-oxytripyrrolidino-phosphonium hexafluorophosphate) and NMM (N-methyl morpholine) in DMF; this yielded the amides 12a and 12b (Figure 2.6). The amines in chemset 11 were added to the columns (1-10) in the two racks, which gave 80 x 2 = 160 different compounds. The immobilized target compounds were cleaved from the resin with 40% TFA (trifluoroacetic acid) in DCM to provide the desired library consisting of the TFA salts 13a and 13b.

#### 2.3 Purity and yields

Every library member was analyzed by HPLC/MS, and all contained a molecular ion consistent with the expected compound. To further enhance the quality of the library, 39 compounds were subjected to preparative HPLC purification with mass-triggered fraction collection. The purity was greater than 70% for 134 of 160 of the library members and was in no case less than 40%; the average purity was 82%. Notably, almost all of the impurities could be identified by their mass. The most prominent impurities were identified as molecules with a mass 106 or 212 g/mol greater than expected. A plausible explanation for this is that, during the cleavage of chemset **12**, some of the Wang linker was liberated as *p*-hydroxy benzyl cation, which may have been added to the members of **13** by electrophilic substitution. In accordance with that, chemset **13b** (20% impurities), which was more prone to electrophilic aromatic substitution than chemset **13a** (8% impurities), also contained a larger amount of this impurity.

The average overall yields from the immobilization of the scaffolds to the cleavage of the target compounds from the resin  $(1a/1b \rightarrow \rightarrow \rightarrow 13a/13b)$ , Scheme 2) were 61% and 50%, respectively. The yields were calculated based on NMR integrals calculations using hexamethyldisiloxane as internal standard.

#### 2.4 Conclusions

Piperazine-2-carboxylic acid was orthogonally protected and coupled with 2fluoronitrobenzene and 3-bromonitrobenzene by direct aromatic substitution and palladiumcoupled amination, respectively. The two nitrophenyl-piperazine scaffolds were immobilized on solid support and combinatorially derivatized with 8 electrophiles and 10 nucleophiles to yield a library of 160 compounds with an average purity of 82%. The synthesized screening library of **13a** and **13b** contains the 4-phenyl-2-carboxy-piperazine pharmacophore, which is known to possess antibacterial and oxytocin receptor-antagonistic activities. The methodology used is well adapted for creating larger libraries, and it may be possible to develop a lead from these compounds.



#### **3** Serine proteases

The second part of the thesis (*papers II–III*) describe the design and synthesis of inhibitors of the two serine proteases thrombin (*paper II*) and HCV NS3 (*paper III*). Thrombin regulates the clotting ability of the blood, and inhibiting this protease should decrease the risk of formation of unwanted blood clots that can block narrowed veins and cause a stroke or a heart attack. In patients infected with the hepatitis C virus (HCV) the NS3 protease of the virus is responsible for cleaving the viral polyprotein into smaller active proteins, thus inhibiting NS3 will prevent the assembly of new viruses.

Proteases regulate a wide array of biological responses, for instance they are vital in the replication phase of many viruses and parasites (*e.g.*, HIV and HCV, and the microorganisms that cause malaria), and they also play important roles in the human body by participating in activities such as immune responses, blood coagulation, digestion, and fertilization. Proteases catalyze the hydrolytic cleavage of proteins into smaller peptide fragments by promoting cleavage of certain amide bonds. These enzymes are divided into four major classes called the serine, aspartic, cysteine and metallo proteases, and they all cleave amide bonds to release a carboxylic acid (*N*-terminal) and an amine (*C*-terminal) part of the substrate (Figure 3.1). In a



Figure 3.1. The mechanism of action of serine proteases.

protease-substrate complex, the first substituent in the *N*-terminal direction from the scissile bond is called P1, the second is called P2, and so on; the first substituent in the *C*-terminal direction is called P1', the second P2', and so on (Figure 3.1).<sup>13</sup> Correspondingly, the subsites in the enzyme that these Pn or Pn' substituents interact with are designated S1 or S1', S2 or S2', and so on. Proteases in the same family can be divided into subfamilies, and the proteases in the subfamilies are often very similar, therefore is it extremely important to design inhibitors that are highly selective for one particular protease, unless the objective is to inhibit all the proteins in a subfamily. Thrombin and HCV NS3 are both chymotrypsin-like serine proteases. The serine proteases are made up of a catalytic triad (*i.e.*, aspartic acid, histidine, and serine) and an oxyanion hole, and in chymotrypsin-like proteases the latter

consists of the peptide backbone of the same serine as in the catalytic triad and the backbone of a glycine in proximity. The catalytic triad cleaves the amide bond, and the oxyanion hole stabilizes the high-energy tetrahedral intermediates. In Figure 3.2, the amino acids are numbered as in the actual chymotrypsin protease; of course the numbers differ for different proteases, and the specific amino acids in the oxyanion hole may also vary in different serine proteases. The hydroxyl group of the serine in the catalytic triad, *Ser195* is highly nucleophilic due to hydrogen bonding to the electron-rich *His57* (**A**). The imidazole ring in *His57* is rich in electrons as the result of hydrogen bonding to the deprotonated carboxylic acid group in *Asp102*. The substrate interacts with the protease, whereupon the scissile bond is moved into the active site (**B**) and the carbonyl group of the scissile bond becomes more electropositive, because it forms a hydrogen bond network with the amino acids in the oxyanion hole. The electropositive carbonyl is attacked by *Ser195*, which leads to formation of a high-energy oxyanion tetrahedral intermediate that is stabilized in the oxyanion hole (**C**). The tetrahedral intermediate collapses and the *C*-terminal cleavage product (the amine) is



Figure 3.2. The mechanism of serine proteases.

released, while whereas the carboxylic acid part remains covalently bound to the enzyme through the serine (**D**). The imidazole ring in *His57* forms a hydrogen bond with a water molecule and thereby makes it more nucleophilic (**E**). The water molecule subsequently makes a nucleophilic attack on the carbonyl, and, for the second time, an oxyanion tetrahedral intermediate is formed (**F**). The tetrahedral intermediate collapses, and this time the *N*-terminal cleavage product (the carboxylic acid) is released from the enzyme, and the protease is ready for another catalytic round (**A**).

### 4 Thrombin protease inhibitors (Paper II)

*Thromboembolism* has long been recognized as a major health problem in the Western world, and it entails the unwanted presence or formation of a blood clot (*thrombus*) (Figure 4.1). If such a clot is dislodged and moves through the blood stream, it is called an *Embolus*. A blood clot that become lodged in a narrowed blood vessel and block circulation, can lead to life-



Figure 4.1. Thromboembolism includes all clinical events in which a thrombus or embolus blocks a blood vessel.

threatening conditions as stroke (CT and CE) and heart attack (AT). Thrombi can also arises in the knees (DVT), where they can block veins and cause morbidity, or dislodge as emboli and induce morbidity in the lungs (PE). In the United States, one death due to stroke occurs every 3.3 minutes, and about 70–80% of such mortality is caused by CT or CE.<sup>14</sup> Around 600.000 people develop PE annually in that country, and 60,000 of those individuals die from the condition. Furthermore, about two million Americans are hit by DVT each year, and roughly five in 1,000 Caucasians have a heart attack (AT). Anticoagulant treatment is given to approximately 0.7% of the population in the western world today, even though the traditional anticoagulant drugs have several drawbacks.<sup>15</sup> At present, warfarin and other coumarins are the only such agents that can be administered orally. However, their anticoagulant effects can be erratic as a result of frequent drug-drug and food-drug interactions, which means that patients need to avoid concomitant medication and maintain a well-controlled diet. In addition, there is a narrow therapeutic window (small difference between safe and effective concentrations).<sup>16</sup> The low-molecular-weight heparins provoke a more predictable response, but some patients develop immune reactions to those substances.<sup>17</sup> Nonetheless, due to their substantial molecular size, heparins must be administered subcutaneously, which tends to

hinder long-term outpatient use. Clearly, there is a strong medical need for a novel, welltolerated oral anticoagulant with a wide therapeutic window. Many large pharmaceutical companies are now searching for such a drug, and the first to be approved will no doubt achieve blockbuster status (over 1 billion US dollars in revenue).

Coagulation per se is a vital process in that it prevents the harmful loss of blood from injured vessels and a blood clot, whether desired or not, is regulated by the coagulation



Figure 4.2. A simplified schematic representation of the blood coagulation cascade. Therapeutic anticoagulants are in bold type.

cascade. In short, this cascade is initiated when damage to epithelial surfaces or some form of trauma triggers a cascade of biological responses that eventually converge in the cleavage of the zymogen prothrombin through factor  $X_a$ , which generates the serine protease thrombin (Figure 4.2). Thrombin cleaves off the fibrinopeptides A and B from fibrinogen, which is thereby transformed into fibrin. The fibrin monomers are cross-linked (mediated by factor XIII<sub>a</sub>) and aggregate to produce a network that traps platelets and erythrocytes, resulting in formation of a blood clot.

The widely used anticoagulant warfarin inhibits the synthesis of vitamin K, which is necessary for the production of prothrombin. The heparins activate antithrombin, which is the natural inhibitor of thrombin. Since thrombin is one of the key enzymes in the coagulation cascade, it has long been looked upon as a viable drug target for preventing thromboembolism. A sort of proof that the concept of using thrombin inhibitors as drugs is actually feasible is provided by the fact that the saliva of medicinal leeches contains the thrombin inhibitor hirudin (a peptide comprising 65 amino acids). It is believed that the direct inhibition of thrombin will lower the level of fibrin in the blood stream, which will decrease the formation of thrombi and thereby reduce the risk of thromboembolism.

## 4.1 The development of thrombin inhibitors based on the D-Phe-Pro-Arg motif

In the mid 1950s, Bettelheim showed that thrombin, similar to many other enzymes, is feedback inhibited, and in this case the effect involves the cleavage product fibrinopeptide A, which is a competitive inhibitor of the reaction between thrombin and fibrinogen.<sup>18</sup> This finding was the starting point for development of the substrate-based thrombin inhibitors. In the late 1960s, Blombäck and coworkers at the Karolinska Institute proposed that fibrinopeptide A (16 amino acids,  $H_2N$ -Ala-Asp-Ser-Gly-Glu-Gly-Asp-**Phe**-Leu-Ala-Glu-Gly-Gly-Gly-**Val-Arg**-*COOH*) has an ordered ( $\alpha$ -helical) structure that position the P9-Phe in space very close to the P2-Val.<sup>19</sup> That suggestion was confirmed by NMR studies at the end of the 1980s, and soon thereafter the a X-ray structure of fibrionopeptide A in complex with  $\alpha$ -thrombin was reported.<sup>20</sup> Moreover, in the late 1970s, optimizations performed by Bajusz *et al.* resulted in the synthesis of the aldehyde tripeptide D-Phe-Pro-Arg-H.<sup>21</sup>

Our research group in Linköping has worked in collaboration with investigators at AstraZeneca to design and synthesize thrombin inhibitors based on the X-ray crystal structure of the covalently bound inhibitor PPACK (D-Phe-Pro-Arg-CH<sub>2</sub>Cl) in complex with  $\alpha$ -thrombin (Figure 4.3, *left*).<sup>22,23</sup> The irreversible protease inhibitor PPACK is itself not suitable as a drug, because the electrophilic  $\alpha$ -chloroketone moiety can bind covalently to countless nucleophiles en route to the intended target, which gives unwanted side effects. PPACK has been used as a tool to map essential structure-activity relationships (SARs) that can be incorporated into reversible inhibitors. The motif D-Phe-Pro-Arg has been used as a lead identification template for reversible inhibitors, which has resulted in the synthesis of several



**Figure 4.3.** Diagrams showing interactions between PPACK (IC<sub>50</sub> 0.017 nM) and thrombin (*left*), along with melagatran (IC<sub>50</sub> 4.7 nM), which is the active substance in the AstraZeneca investigational drug Exanta<sup>TM</sup>. (*right*).

thrombin inhibitors.<sup>24,25</sup> A proof of concept of this design strategy is the AstraZeneca drug Exanta<sup>TM</sup>, which contains the active ingredient melagatran and is now on the market in some countries (Figure 4.3, *right*).

Fibrinogen is the natural substrate of thrombin, and, like all human enzymes, it consists solely of L-amino acids. In the D-Phe-Pro-Arg motif, the phenylalanine moiety has D-stereochemistry, thus the phenyl moiety in D-Phe has a different *vector* then the P3 group of the substrate. The D-Phe is directed away from the true S3 pocket and is instead guided into a distal (D) pocket. As mentioned above, it is probably the P9-Phe of fibrinopeptide A that occupies that space in the fibrinopeptide A / $\alpha$ -thrombin complex. Nonetheless often the S3 terminology are used, but that is wrong. According to the nomenclature for thrombin inhibitors generated using the D-Phe-Pro-Arg motif, the P3 group fits in the D pocket, the P2 group interacts with the S2-pocket (also called the proximal (P) pocket), and the P1 group occupies the S1 pocket (also called the specificity pocket). Figure 4.3 shows a two-dimensional (2D) map of PPACK and its interaction with thrombin, and it can be seen that thrombin has an essentially hydrophobic D pocket that is formed by the amino acids Leu99,



**Figure 4.4.** (A) The D-Phe-Pro-Arg motif; (**B-E**) compounds based on that motif that has been synthesized in Linköping thrombin projects using 5- and 6-membered rings as proline bioisosteres.

Trp215, and Ile174. The S2 pocket is also hydrophobically flanked by the amino acids Tyr60A and Trp60D, and situated in the bottom of the deep S1 pocket is the acidic Asp189. Also of interest in the 2D structure are the hydrogen bonding interactions with the backbones of Gly216 and Ser214. The approach used most often in P1 design is to create an ionic bond between the inhibitor and Asp189, and, in both the D-Phe-Pro-Arg motif and PPACK, the guanidino group in Arg constitutes such a bridge. In the early 1990s, researchers at AstraZeneca used what was at that time a new method of computer-assisted design and discovered that the more constrained and rigid *p*-amidino benzylamine (Pab) group was an even better P1 moiety than the Arg.<sup>26</sup> The Pab group indeed enhanced the overall activity and was used in melagatran, which is the active substance in the prodrug Exanta<sup>TM</sup>. In melagatran, the P2-Pro and the P3-D-Phe have been replaced with azetidine and a D-cyclohexyl amino acid, respectively. Despite these replacements the structural similarities with the motif (D-Phe-Pro-Arg) are still pronounced. The collaboration between AstraZeneca and Linköping University has led to development of several different thrombin inhibitors that were

synthesizied using the D-Phe-Pro-Arg motif as template (Figure 4.4). In all these compounds, the P1-Arg has been substituted with the para-amidino benzylamine (Pab) group, the P2proline has been replaced with various proline isosteres, and the P3-D-Phe has been replaced with different amines. In a previous study, the activity of inhibitor  $14{4}$  (970 nM) was promising (Figure 4.5).<sup>27</sup> A comparison of the D-Phe-Pro-Arg motif and  $14{4}$  reveals the following structural similarities: the Pab group is an Arg bioisostere, the cyclopentene is a proline bioisostere, and the *N*-ethyl aniline group is a D-Phe bioisostere. (Figures 4.4 and 4.5).

Paper II describes an optimization study of the lead compound  $14\{4\}$ , and it also discusses attempts to compare and rationalize the activity of the different cyclopentene isomers in series 14, 31 and 32, and to compare and explain the cyclohexenes in series 17 and the cyclopentenes in series 14 (Figure 4.5).<sup>27</sup> A total of 35 target compounds were synthesized, and the most active of those,  $14\{22\}$ , had an IC<sub>50</sub> of 49 nM (Figure 4.5).



Figure 4.5. The lead compound  $14\{4\}$  (970 nM); the target compounds in the series 14, 17, 31 and 32; the optimized inhibitor  $14\{22\}$ .

#### 4.2 Synthesis of the scaffolds

The scaffold **15** was the key intermediate in the synthesis of the final products in series **14**, and the previously used synthetic pathway to the racemic proline isostere **15** involved five steps and gave an overall yield of 12% (Figure 4.6). In short, the target compounds in series **14** were synthesized from scaffold **15** by amide couplings. The carboxylic acid **15** was synthesized from the aldehyde **20** through methylester protection followed by oxidation. The aldehyde **20** was obtained from the hydrolysis of the bicyclic compound **19**, which was synthesized by adding dichloroacetyl chloride to cyclopentadiene (produced by retro-Diels-Alder reaction of the dicyclopentadiene **18**).



Figure 4.6. The previously reported retrosynthetic route to the target compounds in series 14.

We were intrigued by a novel and versatile synthetic pathway, which, in addition to yielding **14**, facilitated synthesis of the six-membered analog **17** (Figure 4.7). This new route took into account that these compounds are olefins. Over the past few years, olefin metathesis has emerged as a powerful novel carbon-carbon bonding technique (Section 4.3). We believed that the desired cycloalkenes **15** and **16** could be synthesizied from the diolefins **22** and **23** by applying ring-closing metathesis (RCM) chemistry. The carbon framework of **22** and **23** is identical to 4-methyl itaconic acid (**21**), which is commercially available.



Figure 4.7. The new retrosynthetic route to target compounds 14 and 17.

By creating the kinetic enolate of **21** (4-methylitaconate), we assumed that it would be simple to synthesize the racemic diolefine **22** by standard techniques (Scheme 4.1), but it turned out that the  $\beta\gamma$ -unsaturated **21** could not be alkylated.<sup>28</sup> Adding a co-solvent such as hexamethylphosphoramide (HMPA) can increase the reactivity of alkyl cations. However, HMPA is highly carcinogenic, and a special permit is required to use this agent in Sweden. Consequently, we deemed DMPU to be an acceptable replacement for HMPA, and applying that co-solvent made it possible to alkylate **21** with 4-bromobutene, which furnished **22** in 30% yield. Apparently the enolate of 4-methyl itaconate is highly aggregated, thus this compound must be in a strongly polar environment to be reactive. Aggregation means that the positive metal ion binds to the negatively charged enolate, which is thereby rendered

Scheme 4.1. Synthesis of the proline isosteres.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) LDA, LiI, THF:DMPU (1:1), -80 °C 4-bromobutene; (b) LDA, LiI, THF:DMPU (1:1), 5-bromopentene, -80 °C; (c) 3-8 mol % second generation Grubbs' catalyst ( $C_{46}H_{65}C_{12}N_2PRu$ ), DCM, 40 °C.

unreactive. Addition of a strong polar aprotic solvent will cause the metal ions to chelate with the solvent, which, in turn, will free the enolate and make it more inclined to react with the electrophile. The addition of 0.4 equivalents of lithium iodide raised the yield of 22 to 48%. The yields of this alkylation were low, because the stable isomerized product 2-methyl-but-2enedioic acid 4-methyl ester was readily formed (10-40%). When we used 5-bromopentene as electrophile instead of 4-bromobutene, 23 was obtained in 42% yield. Although the yield in these novel alkylations was not impressive, we tested asymmetric alkylations with auxiliaries developed by Evans and Oppolzer, but, unfortunately all attempts failed. After the successful alkylations that had yielded the racemic compounds 22 and 23, it was time for metathesis reactions that were intended to give the corresponding cyclized racemic compounds 15 and 16. Initially, we tried to cyclize the successfully alkylated compounds and derivatives thereof using the first-generation Grubbs catalyst 24 and the Schrocks catalyst 25, these metathesis trials are summarized in Section 4.3.1 (Figure 4.11). When the second-generation Grubbs catalyst 26 became commercially available from Strem Chemicals in August 2000, we subjected the dienes 22 and 23 to high-yield cyclization with 3-8 mol percent tricyclohexylphosphine [1,3-bis(2,4,6-trimethylphenyl)-4,5-dihydroimidazol-2-ylidene] [benzylidine]ruthenium (IV)) dichloride (26) in DCM at 40 °C,<sup>42</sup> which yielded the RCM products 15 and 16 (98% and 85%, respectively).

The proline isosteres **15** and **16** were synthesized along a two-step reaction route with an overall yield of 47% and 35%, respectively. As previously mentioned, the old synthetic pathway to **15** involved five steps and an overall yield of 12%.

#### 4.3 The olefin metathesis reaction

The synthesis of the final products is described in section 4.4. This section (4.3) concerns metathesis and some metathesis trials that were conducted before the above-mentioned reaction conditions were discovered. Furthermore a short history of the metathesis reaction is included here, along with a discussion of the catalysts and the mechanisms involved.

The metathesis reaction is actually very well known and has undoubtedly been performed by virtually every person with at least a background in junior high school chemistry, because ordinary acid/base neutralization is precisely such a reaction (Figure 4.8A). Metathesis is a double exchange reaction, which means that some molecules of the reactants exchange positions and create two new products. Olefin metathesis entails catalytically induced reactions in which olefins undergo bond reorganization and release



**Figure 4.8.** (A) The ordinary acid/base neutralization is a metathesis reaction; (B) the olefin ring closing metathesis reaction; (C) the olefin cross metathesis reaction.

volatile ethene, which constitutes the underlying driving force of the reaction (Figure 4.8B-C).<sup>29,30</sup> The most reliable and most widely applied of these reactions are the ring closing metathesis (RCM) reaction, whereas the type of intermolecular metathesis referred to as cross metathesis (CM) is somewhat more complicated, because two reactants A and B can form the dimers AA, AB, and BB. The geometry of a CM product can be both *cis* and *trans*, even if the *trans* geometry is usually preferred. RCM reactions are run in very diluted solutions (~ 0.01 M) and the CM reactions in fairly concentrated solutions (~ 0.2 M). An ordinary flash column chromatography usually removes the catalyst but if it is absolutely necessary to remove all traces of catalyst it is possible to trap it with the use of lead tetraacetate or phosphine reagents such as P(CH<sub>2</sub>OH)<sub>3</sub> and PPh<sub>3</sub>O).<sup>31</sup>

In the 1950s, several patent applications described a new type of reaction, which at that time was called "coördination polymerization," and the first articles describing such chemistry were published in the early 1960s.<sup>32,33</sup> However, it was not until 1968 that the reaction was given its right name (olefin metathesis) and the underlying mechanism was correctly described in 1971 by Chauvin, Y.<sup>34,35,36</sup> Olefin metathesis remained an obscure and

unreliable reaction through the 1970s and 1980s due to the lack of commercially available well-defined catalyst.<sup>37,38</sup> The catalyst revolution started in the early 1990s, research groups headed by R. R. Schrock and R. H. Grubbs published papers reporting that they had used molybdenum and ruthenium, respectively, to synthesize new and powerful, well-defined metathesis catalysts.<sup>39,40</sup> The early molybdenum (Schrock) catalysts were highly reactive and sensitive to air, they decomposed in protic environments but they could yield metathesis products that were both trisubstituted and sterically hindered (Figure 4.9).<sup>39</sup> The first generation of Grubbs catalysts were moderately reactive and slightly sensitive to air, but the commercially available 24 could tolerate almost any type of functionalities (e.g., esters, amides, ethers, alcohols, and carboxylic acids) and it could yield disubstituted and, in rare cases, also trisubstituted alkenes (Figure 4.9).<sup>40,41</sup> Due to the higher reactivity, the commercially available Schrock catalyst was still a good option, but that changed in the late 1990s, when Grubbs developed, among other ones, the second-generation Grubbs catalyst 25, which surpassed the reactivity of the Schrock catalyst while retaining the air stability and excellent functional group compatibility of the first generation catalyst 24 (Figure 4.9). These catalysts could easily yield alkenes through metathesis that was trisubstituted and in some cases even tetrasubstituted.<sup>42</sup> In 2000, Amir H. Hoveyda published a paper describing



Figure 4.9. Structures of the most commonly used metathesis catalysts, all of which are commercially available.

experiments that included the catalyst **26**, which, after the first metathesis cycle, has the same reactive species as **25** (Figure 4.9).<sup>43</sup> The second-generation Hoveyda/Grubbs catalyst **26** is, in most circumstances, equipotent to the second-generation Grubbs catalyst **25**, although this can vary from better to worse. Despite the development of functional first-generation metathesis catalysts in the 1990s, it was not until the 2000s that the use of metathesis reactions made real headway and this approach became the tool of the average organic chemist. The second-generation Grubbs catalysts have had a major impact on the field of organic chemistry.

I wrote this during the summer 2005: "In my opinion it is very likely that the Nobel Prize in chemistry is imminent for Robert H. Grubbs and Richard R. Schrock, and possibly also Amir H. Hoveyda, for their work in developing metathesis catalysts." Actually, my

prediction was almost true, as Robert H. Grubbs and Richard R. Schrock, and Chauvin, Y (for the correct mechanism) got the Nobel Prize 2005.

For readers who are unfamiliar with the metathesis reaction, it should be pointed out that the active catalysts of the first turnover are the original benzylidene species 24, 25, and 26, and the dissociation products are styrene derivatives. In the subsequent catalytic turnovers, the more reactive (and unstable) methylidene species E is the active catalyst, and the dissociation product is ethene (Figure 4.10). The RCM catalytic cycle with the Grubbs catalysts begins with the dissociation of tricyclohexylphosphine from 24 and 25, or, in the



Figure 4.10. General mechanism of the olefin ring-closing metathesis reaction.

case of 26, with the loss of coordination of the isopropoxide moiety. The loss of the ligand generates the four-coordinate intermediate **A**. Species of **A** derived from catalysts 25 and 26 (*N*-heterocyclic carbene ligands) will have  $k_2 > > k_{-1}$ , whereas species derived from 24 (triphenyl phosphine ligand) will have  $k_2 \sim k_{-1}$ , which explains why the second generation of catalysts are more reactive then the first. In a formal [2 + 2] addition, the benzylidene ligand in **A** attacks a diolefin, creating the metallacyclobutane **B**, which decomposes to release styrene and generate the key intermediate **C**. An intramolecular [2 + 2] cycloaddition creates the metallacyclobutane **D**, which decomposes to yield the very active methylidene intermediate catalyst **E** and the cyclized compound **F**. The methylidene compound **E** attacks another diolefin, producing the metallacyclobutane **G**, decomposition of which liberates

ethene and again creates the key intermediate C. Ethene quickly evaporates, rendering  $G \rightarrow C$  irreversible, which is the driving force of the metathesis reaction.

#### 4.3.1 Metathesis trials

As mentioned in the previous section, several metathesis trials were performed on various cycloalkenes before the second-generation Grubbs catalyst was released (Figure 4.11). Since the attempts to use the first-generation Grubbs catalyst (24) for direct RCM of 22 had failed, we tried to cyclize the *tert*-butyl ester derivative with both 24 and the Schrock catalyst, but with no success. Inasmuch as it is known that electron-withdrawing groups reduce the reactivity in RCM, the next approach was to convert the carboxylic acid to the alcohol.<sup>44</sup> Trace amounts (3%) of the cyclized alcohol product were obtained using the first-generation Grubbs catalyst. Encouraged by that result, we continued with the benzoylated derivative and obtained the cyclized product benzoic acid 5-methoxycarbonyl-cyclopent-1-enyl ester in 35% yield.<sup>42b</sup> This ester could be hydrolyzed and oxidized to yield the proline isostere **15**, but that



Figure 4.11. Metathesis trials aimed at synthesizing 5- and 6-membered cycloalkenes.

meant that this new metathesis strategy was hardly better than the old cyclopentadiene synthetic pathway (Figure 4.6). Luckily, the emergence of the second-generation Grubbs catalysts made it feasible to synthesize **15** and **16** directly from the corresponding diolefins. The second-generation Grubbs catalyst (25) also improved the yield of all other derivatives, except the *tert*-butyl ester derivative, which remained impossible to cyclize, probably due to steric hindrance. After successful synthesis of the scaffolds 15 and 16, we also tested synthesizing 4- and 7-membered cycloalkenes, as well as a tetrasubstituted cyclopentene (Figure 4.12). Attempts at synthesizing the tetrasubstituted alkene gave an isomerisized anhydride (Figure 4.12). Trials aimed at making a 4-membered ring vielded the Cope rearranged product, and efforts focused on generating the 7-membered ring gave only degradation products. In all likelihood, the catalyst was trapped by the neighbouring carboxylic acid functionality (or ester carbonyl). Another possibility is that intramolecular hydrogen bonding between the carboxylic acid and the ester puts the long alkene tail off distance to the disubstituted alkene, thus lowering the population of the reactive conformer. This effect could also explain the lower yield of the cyclohexene 16 compared to the cyclopentene 15. The addition of titanium isopropoxide has been reported to increase the yield in similar cases *i.e.* both unproductive chelates and problem with unreactive conformer.<sup>45,46</sup> Unfortunately the addition of 10 mol % had no effect in this case.



Figure 4.12. Metathesis trials aimed at synthesizing a tetrasubstituted cyclopentene and 4- and 7- membered cycloalkenes.

#### 4.4 Synthesis of the final products

The synthesized scaffolds 15 and 16 were subjected to two amide-coupling derivatization steps to yield the final products in series 14 and 17 (Scheme 4.2), which were subsequently used to optimize the activity of the lead compound  $14\{4\}$ . The first derivatizations of scaffolds 15 and 16 were achieved by performing coupling reactions at r.t. with various amines taken from the P3 building blocks 27 (Figure 4.13), mainly using BOP-Cl (bis-(2-

Scheme 4.2. Synthesis of the final products used for activity optimizations.<sup>a</sup>



<sup>*a*</sup>**Reagents and conditions:** (a) BOP-Cl, DIPEA, **27** (the P3-building blocks. HNR<sup>1</sup>R<sup>2</sup>), DMF; (b) LiOH, dioxane-H<sub>2</sub>O (1:1); (c) DCC:HOBt, Pab(Z), DIPEA, THF (50 °C); (d) TFMSA, anisole, DCM.



Figure 4.13. The P3 building blocks 27 (HNR<sup>1</sup>R<sup>2</sup>),  ${}^{a}R = H \text{ in } 14\{7\} \text{ and } 31\{7\}, R = {}^{t}Bu \text{ in } 27\{7\}, 28\{7\}, and 30\{7\}.$ 

Oxo-3-oxazolidinyl)-phosphinic-chloride) as coupling reagent, which yielded 28 and 29. Surprisingly, trials with HATU (O-(7-azabenzotriazol-1-yl)-N,N,N,N'-tetramethyluronium hexafluorophosphate), one of the most versatile coupling reagents known, resulted in sluggish reactions and low yields. Some of the sterically hindered anilines in 27 (the P3 building blocks) could be coupled only at elevated temperatures using BOP-Cl. Before the optimized coupling conditions were found we used  $30\{7\}$  as a model compound and tested several other conditions (PyBOP, DCC:HOBt, EDC:HOBt, HATU in THF, or DMF at r.t. or under heating). The methylesters in the series 28 and 29 were hydrolyzed with lithium hydroxide in dioxane-water (1:1). This resulted in lithium carboxylates, which were dried by coevaporation with methanol-toluene (1:1) and then coupled with Pab(Z) x 2 HCl using DCC:HOBt (dicyclohexyl carbodiimide:1-hydroxy benzotriazol) and DIPEA in THF (50 °C).26 Use of other coupling reagents (i.e. BOP-Cl, HATU, and PyBOP) in DMF did not improve the yields. The Z-protecting group and the *tert*-butyl-protecting group in members derivatized with  $27{7}$  were simultaneously hydrolyzed using TFMSA in DCM at 0 °C, which yielded the target compounds in series 14 and 17. All chemical reactions gave better vields with the cyclopentene derivatives than with the cyclohexene derivatives.

Subsequently, we also synthesized the target compounds in series **31** and **32** (Scheme 4.3), that together with series **14**, were used in a directional vector study (Table 4.1, Section 4.5). The scaffold **33** was easily prepared (90% yield) in a straightforward manner from commercially available 1-cyclopentene-1,2-dicarboxylic anhydride in methanol in the presence of zinc chloride. The target compounds in series **31** were synthesized in the same way as the members of series **14** and **17** (Scheme 4.3). The acid **15** was coupled with Pab(Z) x

Scheme 4.3. Synthesis of target compounds needed for the vector study.<sup>a</sup>



"Reagents and conditions: (a) methanol,  $ZnCl_2$ ; (b) BOP-Cl, DIPEA, 27 (the P3 building blocks  $HNR_1R_2$ ), DMF; (c) LiOH, dioxane-H<sub>2</sub>O (1:1); (d) DCC:HOBt, Pab(Z), DIPEA, THF (50 °C); (e) TFMSA, anisole, DCM.

2 HCl using DCC:HOBt and DIPEA, which generated the benzamidine(Z)-methylester **34** (90% yield; Scheme 4.3). Saponification with lithium hydroxide in dioxane-water gave the lithium salt **35** (85% yield), which was dried through co-evaporation with toluene/DCM (1:1) to afford water-free crystals. These crystals were dissolved in DMF (50 °C) and coupled with several P3 building blocks **27**, using BOP-Cl as coupling reagent. The drying procedure after the saponification was necessary due to the polar and hygroscopic nature of the molecule. The Z-group was cleaved using TFMSA and anisole in DCM to yield the target compounds in series **32**. All these compounds were racemic, but we purified **14**{*1*} on a chiral column, and the enantiomers (*S*)- and (*R*)-**14**{*1*} could be separated and individually tested for activity.

The *N*-ethyl- and *N*-cyclohexyl-substituted anilines in the P3 building blocks **27** (Scheme 4.4) were synthesized from acetaldehyde or cyclohexanone and various anilines to form the corresponding imines, which were reduced using borane-pyridine and molecular sieves (4 Å) in methanol.<sup>47</sup> Compound **27**{7} was synthesized using *o*-toluidine, *tert*-butyl bromoacetate, and potassium carbonate in DMF at 80 °C.

In paper II, a total of 84 new molecules were synthesized and characterized by <sup>1</sup>H-NMR, <sup>13</sup>C-NMR, and high resolution mass spectroscopy.

**Scheme 4.4.** Synthesis of the commercially unavailable amines in the P3 building blocks **27**.<sup>a</sup>



<sup>a</sup>**Reagents and conditions**: (a) acetaldehyde, molecular sieves (4 Å), borane-pyridine, methanol; (b) cyclohexanone, molecular sieves (4 Å), borane-pyridine, methanol; (c) *tert*-butyl bromoacetate,  $K_2CO_3$ , DMF (80 °C).
## 4.5 Structure-activity relationships

To fully elucidate the interactions between thrombin and the ligands in series 14, we cocrystallized  $\alpha$ -thrombin with the racemate of 14{14} (Figure 4.14-4.16), which was the most potent compound at that time the X-ray analysis was performed. Figure 4.14 illustrates the overlay of the X-ray structure of the  $\alpha$ -thrombin-14{14} complex and the  $\alpha$ -thrombin-PPACK complex. Only the (*R*)-enantiomer of 14{4} has strong enough affinity for  $\alpha$ thrombin to be crystallized with it. The  $\alpha$ -thrombin-(*R*)-14{14} and  $\alpha$ -thrombin-PPACK complexes exhibit similar binding modes, as can be seen in Figure 4.14. A direct comparison with PPACK with respect to affinity for the protease may seem irrelevant, since PPACK, via its chloromethyl and ketone group, bind covalently to His57 and Ser195 (Figure 4.3). However, it has been reported that non-covalent inhibitors such as melagatran exhibit substantial affinity for thrombin and interact with the enzyme in the same way as PPACK



**Figure 4.14.** Connolly surface map of the X-ray structure of the  $\alpha$ -thrombin-**14**{*14*} complex (yellow) at 1.9 Å resolution superimposed on the X-ray structure of the  $\alpha$ -thrombin-PPACK complex (magenta).



**Figure 4.15**. Electron density map of **14**{*14*} showing the well-defined region of the P1-benzamidino in the S1 cavity.



Figure 4.16. Electron density of the P3 group and the fitted structure of  $14\{14\}$  show the slightly disrupted electron density in the D-pocket.



**Figure 4.17.** Possible orientation of the P3 phenyl group that can lower the energy of the ligand binding conformation.

does, and they form an extensive hydrogen bond network with Ser214 and Gly216.<sup>25</sup> The benzamidine group of (*R*)-14{14} fits nicely into the S1 pocket, and there are very strong electrostatic interactions between the amidine group and the carboxylate of Asp189, with N-O distances of 2.45 and 2.71 Å (Figure 4.15). Furthermore, in a manner similar to PPACK, the cyclopentene ring of 14{14} fits well in the hydrophobic proximal pocket (S2) of thrombin. In the X-ray structure of the 14{14}-thrombin complex, the electron density of the P3 5-chloro-2-methylphenyl is slightly disrupted in relation to the D pocket, which indicates that the phenyl moiety adopts more than one conformation (Figure 4.16 and 4.17). As can be clearly seen in Figures 4.14 and 4.17, the inhibitor does not occupy the whole D pocket, and, in light of that observation, we synthesized inhibitors 14{19-22}.



Figure 4.18. The activity of some compounds showing the synergic effects of 2,5-phenyl substitution in the 14 series and the lower affinity of the six-membered derivative  $17\{4\}$  compared to the five-membered analog  $14\{4\}$ .

Further optimizations of cyclopent-2-ene-1, 2-dicarboxylic acid derivatives based on Xray analysis data led to compounds that exhibit 20 times greater activity against thrombin *in vitro* compared to the lead compound (*R*)-**14**{4} (Figure 4.18).<sup>27</sup> The 2,5-substitutions on the phenyl moiety in the P3 position constitute the major improvement. Due to the disrupted electron density in the D pocket, either one of the two substituents probably interacts with the D pocket (Figure 4.16). An alternative explanation for the increased activity is that adoption of the conformation shown in Figure 4.17 would lower the energy of ligand bonding conformation. The synergic effect of 2,5-substitution is illustrated by comparing the 2,5methoxy-substituted compound **14**{20}(79 nM) with the *ortho*-substituted **14**{*1*} (240 nM) and *meta*-substituted **14**{6} (250 nM) (Figure 4.18). Moreover, larger groups on the nitrogen in the P2-P3 amide linker further improved the activity ( $14\{22\}$ , 49 nM), because they allowed the P3 phenyl moiety to adopt the thermodynamically favorable *cis* configuration around the rotamerically hindered P2-P3 amide linker. Moreover, the increased flexibility of the cyclohexene ring in series 17, did unfortunately not, improve the ligand-enzyme interactions compared to the cyclopentene-containing analogs in series 14, indicated by the comparison of  $17\{4\}$  and  $14\{4\}$  (Figure 4.18).

O RaRin	$\bigtriangledown$	O RaRaN	$\langle \rangle$	O RaRaN	$\square$
0	Pab	1121111	O Pab	1121111	O Pab
Cnd	IC <sub>50</sub>	Cnd	IC <sub>50</sub>	Cnd	IC <sub>50</sub>
Сра	( µM)	Сра	μM)	Сра	(µM)
<b>14</b> { <i>1</i> }	0.24	<b>31</b> { <i>l</i> }	24	<b>32</b> { <i>l</i> }	7.8
<b>14</b> {2}	4.6	<b>31</b> {2}	> 130	<b>32</b> {2}	22
<b>14</b> { <i>3</i> }	0.68	<b>31</b> { <i>3</i> }	33	<b>32</b> { <i>3</i> }	19
14{4}	0.97 <sup><i>a</i></sup>	<b>31</b> {4}	28	<b>32</b> {4}	16

**Table 4.1.** Activities of compounds  $14\{1-4\}$ ,  $31\{1-4\}$ , and  $32\{1-4\}$ .

<sup>*a*</sup> The lead compound, the (R)-enantiomer.

Continuing this work, we performed a directional vector study of the isomeric compounds  $14\{1-4\}$ ,  $31\{1-4\}$ , and  $31\{1-4\}$  (Table 4.1). The spatial dissimilarity of target compounds in series 14, 31, and 32 is due to different bond hybridization of the P1- and P3directing atoms of the central ring scaffold. We assumed that the best activity would be achieved by using the same hybridization configuration as in the D-Phe-Pro-Arg motif and PPACK. Target compounds in series 14 and Pro have sp<sup>3</sup> hybridization of the P1-directing atom (tetrahedral  $\rightarrow$  theoretical bond angle 109.5°) and sp<sup>2</sup> hybridization of the P3-directing atom (planar  $\rightarrow$  theoretical bond angle 120°). The bond hybridization of the directing atoms is reversed in target compounds in series 32, whereas target compounds in series 31 have only sp<sup>2</sup>-directing atoms. The theoretical dihedral angle between the P1- and P3-carbonyl atoms in the P2 scaffold is 0° in series **31** but is  $< 0^{\circ}$  in Pro and in the series **14** and  $> 0^{\circ}$  in the series **32**. Examination of the potency of the target compounds  $14\{1-4\}$  and  $32\{1-4\}$  given in Table 4.1 indicates that reversing the hybridization of the directing atoms compared to the D-Phe-Pro-Arg motif lowered the activity towards thrombin. On average, members of series 32 show a 20-fold decrease in activity compared to target compounds in series 14 that have the same P3 moiety. This also applies to the target compounds  $31\{1-4\}$ , in which the shift from sp<sup>3</sup> to sp<sup>2</sup> resulted in, on average, a 50-fold decrease in activity compared to the analogs of target compounds in series 14. Thus, we may safely conclude that within the series of P2cyclopentene dicarboxylic derivatives, the motif with P1-sp<sup>3</sup> and P3-sp<sup>2</sup> directional vectors

provides the best interactions with thrombin, which is consistent with the directional vectors in the D-Phe-Pro-Arg motif.

## 4.6 Conclusions

By using RCM chemistry, we were able to develop a new and improved synthetic pathway to the previously reported thrombin inhibitors in series 14, which are based on the D-Phe-Pro-Arg motif. A total of 84 compounds were synthesized in the series 14, 17, 31, and 32. All these compounds were characterized, including the 34 final products, which were also tested against thrombin for inhibitory properties. Initial optimizations of the lead compound  $14{4}$ (970 nM) yielded thrombin inhibitors with activity in the nanomolar range. X-ray crystallographic and computer analyses of  $14{14}$  (220 nM) led to synthesis of compound  $1{22}$ , which is even more active (49 nM). A directional vector study illustrated the importance of using P1- and P3-directional vectors that are consistent with the D-Phe-Pro-Arg motif. The results of the SAR analysis indicate that the compounds in series 14 can be further optimized to provide more potent inhibitors.



# 5 Inhibitors of the HCV NS3 protease (Paper III)

Based on the therapeutic success of targeting the HIV protease, the HCV NS3 protease represents a major target in the search for antiviral agents against the hepatitis C virus (HCV).<sup>48</sup> Recently, a phase II study demonstrated the antiviral efficacy of a NS3 protease inhibitor in patients suffering from hepatitis C.<sup>85h</sup>

WIIO region	Total pop.	Infected pop.
who region	(Millions)	(Millions)
Africa	602	31.9
Americas	785	13.1
Eastern Mediterranean	466	21.3
Europe	858	8.9
Southeast Asia	1500	32.3
Western Pacific	1600	62.2
Total	5811	169.7

**Table 5.1.** Number of people infected with the hepatitis C virus in regions designated by the WHO.

HCV is a single-stranded positive RNA virus of the family Flaviviridae, and it has a narrow host range in that only humans and chimpanzees are susceptible to infection.<sup>49</sup> Chronic infection with HCV is a major health problem that can cause cirrhosis, hepatocelluar carcinoma, and liver failure.<sup>50</sup> The World Health Organization (WHO) has estimated that approximately 170 million people are infected with HCV worldwide (Table 5.1).<sup>51</sup> The virus is spread through blood contact and through exchange of bodily fluids (*e.g.*, via blood transfusions, recycling of infected needles, and sexual intercourse). Around 20% of those who have been infected for 10 to 20 years develop cirrhosis, and about 5% of such prolonged cases progress to liver cancer.<sup>52</sup> HCV is one of the major indications for liver transplantation, and at present there is no vaccine.<sup>53</sup> The only therapy available is expensive, and comprises of a high dose of PEGylated  $\alpha$ -interferons in combination with ribavirin, and it results in only an approximately 50% sustained virological response in infected patients patients of genotype 1, the dominant genotype in the western world.<sup>54</sup> The limited efficacy of this treatment clearly highlights the immense need for new and more efficacious therapeutic options.

The life cycle of HCV is illustrated in Figure 5.1, and the letters in boldface type in this discussion refer to the stages shown in the figure. The cycle begins when the virus enters the blood and comes in contact with a liver cell (hepatocyte) (A).<sup>54b</sup> The glycoproteins E1 and E2 bind to the receptors CD81 and LDLR on the hepatocyte, and the virus particle is delivered to the cytoplasm of the cell through endocytosis (B). In the cytoplasm, the virus is uncoated (C), and the RNA is released and translated to a polyprotein by the host cell ribosomes (D). The HCV polyprotein is cleaved into at least 10 active viral proteins by host enzymes and the viral

enzymes NS2 and NS3 (E). The viral capsule, which is assembled in the ER of the hepatocyte (F), consists of at least three viral proteins: the core of the nucleocapsid (C) and two glycoproteins (E1 and E2). The polymerase NS5B is responsible for replication of the RNA genome to a double RNA helix (G). The double-stranded RNA is unwound by the NS3 helicase (H), and the positive single-stranded part is packaged into the assembled nucleocapsid (I) to form a mature virion. The constructed virions are subsequently released from the cell through vesicle fusion and are ready to infect other hepatocytes (J).

As enlightenment to fully understand the completeness of the virus cycle, I would like to point out that only some single-stranded RNA are packaged into a nucleocapsid and the



**Figure 5.1.** The life cycle of HCV. The cycle comprises the following stages: (**A**) attachment; (**B**) endocytosis; (**C**) uncoating of the virus; (**D**) translation of RNA to the polyprotein; (**E**) proteolytic processing of the polyprotein; (**F**) formation of the nucleocapsid; (**G**) replication; (**H**) unwinding of the double-stranded RNA; (**I**) viral assembly and vesicle transport; (**J**) vesicle fusion and virion release.

other ones are used as templates in the translation process to produce more polyproteins (**D**). Moreover, the negative single stranded RNA, that together with the positive single-stranded RNA forms the RNA-double helix, can after unwounding (**H**) be used as template in the replication to produce more positive single-stranded RNA (**G**).

The HCV genome and its encoded gene products have recently been characterized in detail.<sup>55</sup> The genome consists of a 5' non-translated region (NTR), which includes the internal ribosome entry site (IRES), followed by the open reading frame (ORF) that is translated to a



Figure 5.2. The single-stranded positive RNA genome of the hepatitis C virus.

single polyprotein (Figure 5.2). There is also another NTR in the 3' region. The HCV polyprotein is made up of approximately 3000 amino acids, which are proteolytically processed into four structural proteins (C, E1, E2, and P7) and six nonstructural proteins (NS2, NS3, NS4A, NS4B, NS5A, and NS5B) (Figure 5.3).<sup>56</sup> These cleavages are mediated by host enzymes and the viral proteases NS2 and NS3. The host enzymes cleave the structural proteins, and the NS2 enzyme is responsible for cleavage of the NS2-NS3 junction through autocatalysis, which frees the *N*-terminal part of the NS3 protein. The NS3 protein is a bifunctional enzyme: its *N*-terminal domain exhibits protease activity, and its *C*-terminal domain displays RNA helicase/ATPase activity. The NS3 protease is responsible for the autocatalytic *cis* cleavage of the NS3-NS4A junction and the *trans* cleavage of the NS4A-NS4B junction. The NS3-NS4A complex mediates the cleavage at the NS4B-NS5A and NS5A-NS5B junctions.



Figure 5.3. The polyprotein that is expressed by the ORF of the HCV genome.

# 5.1 Development of NS5A/5B substrate-based inhibitors of the HCV NS3 protease

In the late 1990s, it was discovered that the HCV NS3-NS4A protease complex is feedback inhibited by several of its cleavage products.<sup>57,85a</sup> Compound **36** corresponds to the *N*-terminal cleavage product of a substrate derived from the NS5A/NS5B cleavage site, and it inhibits the NS3 protease with an IC<sub>50</sub> of 7100 nM (Figure 5.4).<sup>85a</sup> The iterative structure based design process started with the substrate-based hexapeptide **36**, which was *N*-terminally capped with an acetyl group, and the oxidatively sensitive cysteine was replaced with norvaline. This gave



Figure 5.4. The truncated NS5A/5B cleavage product 36 has been used as template in the structure based design process.

less polar inhibitors that had greater stability, but the activity remained low. Real progress was made in when a benzyl substituted hydroxyproline moiety was introduced instead of the proline moiety, this gave the modestly potent hexapeptide **37**.<sup>85b</sup> Hexapeptides are not viable as drugs, because they have a large molecular mass, and they are usually too polar to pass through cell membranes, so efforts were made to achieve truncation of the P5 and P6 groups. Unfortunately, this resulted in loss of inhibitory activity, but replacing the benzyl ether by a naphthyl ether and performing a ring fusion of the propyl moiety of norvaline to a cyclopropyl group, as well as another ring fusion of isolecine to cyclohexyl glycine, resulted in the tetrapeptide inhibitor **38**.<sup>85c</sup> Although the tetrapeptide inhibitors at this stage were still less potent then their hexapeptide counterparts, they were smaller and less polar due to the

removal of the two Asp amino acids, thus they were much more appropriate as lead candidates according to the Lipinski rule of five.<sup>5</sup> Extensive optimization of the P2 substituent led to compound **39**, which had a respectable  $IC_{50}$  value of 800 nM against the HCV NS3 protease.<sup>85d</sup> Optimizations of the P1 group led to the vinylcyclopropyl derivative **40**, which further improved the potency to a promising  $IC_{50}$  of 170 nM (Figure 5.5).<sup>85e</sup> Making the P3 building block a little more bulky and extension of the P2 substituent with a phenyl moiety



Figure 5.5. Some examples of potent tetrapeptides used to deveolp the clinical candidate BILN 2061 (43).

gave compound **41** an impressive  $IC_{50}$  of 1 nM.<sup>85f</sup> However, **41** and similar compounds with excellent  $IC_{50}$  values had poor membrane permeability and were thus not suitable as drugs. It was obvious that some major structural changes had to be made. Earlier NMR analysis of compound **37** had shown that the P1 and P3 groups in the active conformation are located close to each other and undergo extensive rigidification upon binding.<sup>85b</sup> A macrocyclization between P1 and P3 would give rise to less peptidic structure, and the rigidification would probably lead to a more desirable pharmacokinetic profile than the corresponding linear peptides. Truncation of the P4 moiety, capping the P3 amine with a cyclopentyl carbamate, and using cyclopentyl glycine in the P3 position yielded compound **42**, which had an  $IC_{50}$  of 400 nM.<sup>85g</sup> This was a promising result, because the activity was still moderately high despite

the truncation of the P4 moiety and introduction of the new P3-cyclopentyl glycine moiety. Further optimizations of the P2 substituent and a macrocyclization between the P1 and the P3 position of **42** led to the excellent inhibitor **43** (**BILN 2061**), which has very good membrane permeability and an EC<sub>50</sub> of 3 nM against replicon cells.<sup>85h</sup> The clinical phase II study of **BILN 2061** showed that it is an effective protease inhibitor with an antiviral impact in humans infected with HCV. Treatment with **BILN 2061** (200 mg) twice a day for two days decreased the viral load in patients infected with HCV genotype 1 to nearly undetectable levels within 24–28 h.<sup>85h</sup> However, the antiviral efficacy of BILN 2061 is less pronounced in patients with HCV genotype 2 or  $3.^{58}$  Further clinical trials are currently on hold due to cardiotoxicity observed in recent toxicological studies in monkey receiving high doses of **BILN 2061** for four weeks.



Figure 5.6. A P1- and P3-substituted P2 carbonyl-L-proline moiety (A) compared with some P1- and P3-substituted P2-proline bioisosteres (B-C).

Previous investigations of thrombin inhibitors performed in our laboratory have shown that proline can be successfully replaced with five- and six-membered carbocyclic isosteres (Figure 5.6).<sup>59,60</sup> We were interested to explore whether the structural motif of the cyclopentene moiety, *i.e.* a *trans*-4-hydroxy-cyclopentene dicarbonyl derivative, could provide advantages over the P2-hydroxyproline moiety found in compound **36-43**. To generate the initial SAR from this novel series we selected 7-methoxy-2-phenylquinolin-4-ol as P2 substituent and a variety of P1 and P3 amino acid derivatives (Figure 5.7). Although **BILN 2061** has entered clinical studies, we believed that the less complicated compound **41** would be a better structural motif to mimick



Figure 5.7. The generic structure of the target compounds 44a and compound 44a $\{3,4\}$ , which is the hydroxy-cyclopentene analog of 41.

Among other compounds, we synthesized  $44a\{3,4\}$ , which is the hydroxycyclopentene analog of the L-hydroxyproline compound 41 (Figure 5.7). An activity comparison of the two compounds shows that they are equipotent, which proves that in this case the hydroxy-cyclopentenes are good L-hydroxy proline bioisosteres.

#### 5.2 Synthesis of the scaffold

A retrosynthetic analysis of the compounds in series **44a** confirmed that they could be synthesized from the chiral scaffold **A** by two ordinary peptide couplings and one Mitsunobu



Figure 5.8. Retrosynthetic route to target compounds 44a.

reaction (Figure 5.8).<sup>61,86</sup> We reasoned that it should be possible to synthesize the scaffold **A** from the diester **B** by a regioselective monohydrolysis. Synthesis of the alcohol **B** should be feasible through stereoselective reduction of the ketone **C**. Also, the  $\alpha\beta$ -unsaturated ketone **C** should be available through olefination of **D**. The enantiomeric synthesis of **D** is described in the literature.<sup>62</sup> Regrettably, the olefination of **D** led to racemization, so we instead chose to use the more easily synthesized racemic *syn* derivative of **D**, namely compound **48** (Scheme 5.1).

Scheme 5.1. Synthesis of the previously reported intermediates.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) *p*-toluene sulfonic acid, methanol, reflux; (b)KMnO<sub>4</sub> (aq),
0 °C; (c) *i* acetic anhydride, 130 °C; *ii* NaOAc, 130 °C; (d) CuBr<sub>2</sub>, refluxing THF;
(e) CaCO<sub>3</sub>, DMF;100 °C.

The ketone **48** was synthesized as previously reported from commercially available and inexpensive (*syn*)-tetrahydrophtalic anhydride **45** (Scheme 5.1).<sup>62,63</sup> The anhydride **45** was esterfied and oxidatively cleaved to the dicarboxylic acid **46**, which was cyclized and

Scheme 5.2. Synthesis of the hydroxyproline bioisostere scaffold.<sup>a</sup>



<sup>a</sup>**Reagents and conditions:** (a) NaBH<sub>4</sub>, methanol, -30 °C; (b) LiOH, 0 °C, water/dioxane (1:1).

decarboxylated to the *syn*-cyclopentanone **47**. The ketone was  $\alpha$ -brominated with cupric bromide and  $\beta$ -eliminated to the olefin **48** using calcium carbonate. As mentioned in the discussion concerning the retrosynthesis, we tried to produce chiral **C** from chiral **D** by using CuBr<sub>2</sub>/CaCO<sub>3</sub> or PhSeBr/LDA followed by oxidation, but this strategy always resulted in the racemate **48**.<sup>64</sup> The racemization of **48** was confirmed by polarimetry and NMR analysis performed using the chiral shift reagent europium-tris[3-(heptafluoropropylhydroxymethylene)-(+)camphorate].<sup>65,68</sup>

The ketone 48 was reduced to the alcohol 49 in a 6:1 diastereomeric ratio (92%) using sodiumborohydride in cold (-30 °C) methanol to reduce the possibility of unwanted 1,4addition (Scheme 5.2). When the reduction was tried at room temperature, the yield was poor and the *syn/anti* ratio was 1:1. It has been reported that adding CeCl<sub>3</sub> will decrease the amount of 1,4-addition, but in our study that approach did not improve the yield or the diastereoselectivity, and instead it actually diminished both.<sup>66</sup> We assumed that *svn*-49 was the major compound and that *anti*-49 was the minor compound, since the  $\beta\gamma$ -unsaturated ester probably acted as a steric hindrance and guided the hydride donation to the opposite face. This hypothesis was confirmed by experiments using nuclear Overhauser effect spectroscopy (NOESY; Section 5.2.1). The next step was the monohydrolysis of 49 to 50. We believed that it would be possible to achieve selective monohydrolysis of the  $\beta\gamma$ -unsaturated ester, leaving the  $\alpha\beta$ -unsaturated ester intact. This should be feasible for two reasons. First, hydrogen bonding between the 4-hydroxyl and the carbonyl oxygen of the  $\beta\gamma$ -unsaturated ester is plausible, and this neighboring group participation (NGP) effect makes the  $\beta\gamma$ -unsaturated carbonylcarbon more electropositive and thus more inclined to be hydrolyzed.<sup>67</sup> Second, the carbonyl carbon in an  $\alpha\beta$ -unsaturated ester has greater electron density than a carbonyl carbon of a  $\beta\gamma$ -unsaturated ester, hence it is less electropositive. The syn/anti (6:1) diastereomeric mixture **49** was monohydrolyzed using 1.0 equivalent of LiOH in dioxane/water, and the hydrolysis mixture could be further purified by crystallization (hexane/diethylether) to pure monomethylester **50**. The yield was not impressive (27%), but it was taken after several crystallizations without regeneration of any material. The regiochemistry of **50** was determined by a heteronuclear multiple bond correlation (HMBC) spectrum see section 5.2.1 for further information.

#### 5.2.1 Structure determination of the scaffold using 2D-NMR

In a NOESY spectrum, hydrogens that are in proximity to each other exhibit nuclear Overhauser effects (NOEs), thus it is possible to distinguish *syn*-**49** from *anti*-**49** (Figure 5.9). The reduction of **48** described in the previous section yielded a *syn/anti* (6:1) diastereometric

Scheme 5.3. Inversion of the stereochemistry using the Mitsunobu reaction.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) BnCOOH, PPh<sub>3</sub>, DIAD, dry THF; (b) NaOMe, MeOH.

mixture of **49**, which was purified by flash column chromatography. Unfortunately, the diastereomer of **49** that was formed in a smaller amount (presumably *anti*-**49**) was always strongly contaminated with the predominant diastereomer (presumably *syn*-**49**). The presumed *anti*-**49** was synthesized from the presumed *syn*-**49** by use of the Mitsunobu inversion reaction (Scheme 5.3). In the NOESY spectrum of the presumed *syn*-**49**,  $H_1$  and  $H_4$  interacted with the same methylene hydrogen ( $H_{5\beta}$ ), which confirmed that the compound did



Figure 5.9. Significant NOE effects in *syn*-49 (illustrated by the 1*S*,4*R* enantiomer) and *anti*-49 (illustrated by the 1*S*,4*S* enantiomer), and the structure-determining HMBC couplings in 50.

indeed have the *syn* configuration (Figure 5.9).<sup>68</sup> Also, in the NOESY spectrum of the presumed *anti*-49,  $H_4$  interacted with  $H_{5\alpha}$ , and  $H_1$  interacted with  $H_{5\beta}$ , which corroborated the *anti* configuration (Figure 5.9).

To confirm the regiochemistry of **50**, we performed an HMBC experiment, because the spectra that are produced can reveal  ${}^{13}C \leftrightarrow {}^{1}H$  couplings that are separated by two or three bonds, even those that span heteroatoms. The HMBC spectrum we obtained showed the expected interactions between the methylester hydrogens and the carbonyl carbon that also participated in interactions with the alkene hydrogen (Figure 5.9).<sup>68</sup> These interactions confirmed that the synthesized compound **50** was the  $\beta\gamma$ -unsaturated acid. The *syn* stereochemistry of **50** was verified by a NOESY spectrum that showed the same relevant NOEs as seen for *syn*-**49**.<sup>68</sup>

## 5.3 Synthesis of the final products

The carboxylic acid **50** was coupled with the P1 building blocks in series **51** (H<sub>2</sub>NCHR<sup>1</sup>COOtBu) to yield the amides **52** (Scheme 5.4, and Figure 5.10). The secondary alcohols **52** were coupled with 7-methoxy-2-phenyl-quinolin-4-ol (R<sup>2</sup>OH) using the Mitsunobu reaction, which gave the ethers **53**. <sup>69,85f,</sup> The methylesters **53** were hydrolyzed to the carboxylic acids **54** and then coupled with the P3/P4 building blocks **55** (H<sub>2</sub>NR<sup>3</sup>) to provide the amides **56** (Figure 5.10). These amides were purified by HPLC, but, of all twelve

Scheme 5.4. Synthesis of the target compounds.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) DIPEA, HATU, **51** (P1 building blocks), 0 °C, DMF; (b) R<sup>2</sup>OH, (methoxy-2-phenyl-quinolin-4-ol), PPh<sub>3</sub>, DIAD, THF; (c) LiOH, dioxane/water (1:1); (d) **55** (P3/P4 building blocks), DIPEA, coupling reagent (HATU and DMF or DCC/HOBt and THF); (e) triethylsilane, TFA, DCM.



Figure 5.10. The P3/P4 building blocks in series 55 (HNR<sup>3</sup>R<sup>4</sup>) (*left*) and the P1 building blocks in series 51 (H<sub>2</sub>NCHR<sup>1</sup>COOtBu) (*right*).

Scheme 5.5. Synthesis of the commercially unavailable P1 building blocks.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) Boc<sub>2</sub>O, DMAP, *tert*-butanol; (b) Pd-C, 95% ethanol, H<sub>2</sub>; (c) LiOH, dioxane/water (1:1); (d) Fmoc-ONsu, NaHCO<sub>3</sub>; (e) TBTA, 50 °C, DCM; (f) DMF, piperidine.

of these diastereomeric pairs, only one pair could be separated into two compounds, namely  $56a\{2,2\}$  and  $56b\{2,2\}$ . The *tert*-butyl esters in 56 were hydrolyzed to target compounds 44, using TFA and triethylsilane as a cation scavenger.<sup>70</sup> After HPLC chromatography, another four of the target compounds in series 44 could be separated into the chiral final products 44a and 44b. The chiral compounds (44a and 44b), the diastereomeric mixtures (44), and two *tert*-butyl esters (56) were tested for NS3-protease-inhibiting properties. The activity data and complete structures of the tested compounds are presented in Section 5.5. (Table 5.2). The stereochemistry of the diastereomers was confirmed by ROESY and TOCSY 2D-NMR spectra in combination with computer calculations (Section 5.4).

The P1 and P3/P4 building blocks in series **51** and **55** were all synthesized from commercially available amino acids, with the exception of the amines **51**{*1*} and **55**{*6*}, which could be purchased. The P1 building block **51**{*2*} was easy to produce from (*S*)-2-benzyloxycarbonyl-amino-pentanoic acid (**57**) by *tert*-butylester protection using Boc<sub>2</sub>O followed by Z-deprotection using catalytic hydrogenation (Scheme 5.5). The cyclopropane derivative **51**{*3*} was synthesized from (*1R*,*2S*)-1-amino-2-vinyl-cyclopropanecarboxylic acid ethyl ester (**58**).<sup>69,71,72,73</sup> The ethylester was hydrolyzed and the amine was Fmoc protected (**59**).<sup>74</sup> The carboxylic acid was *tert*-butyl ester protected using TBTA, and the Fmoc was removed using piperidine in DMF.<sup>75</sup> The target compound **51**{*3*} could be purified by flash column chromatography, but it was not preferred as the free amine **51**{*3*} decomposed on the column; instead, coupling of the crude **51**{*3*} to the scaffold **50** as quickly as possible was favored. The free amine **61** was synthesized from **60** (Boc-(*R*)-cyclohexylglycine) through methylester protection using silver(II)oxide and methyliodide, and the Boc group was deprotected using TFA (Scheme 5.6). The amine **61** was coupled with Z-(*R*)-Val-OH, and the

Scheme 5.6. Synthesis of the P3/P4 building blocks 55{1} and 55{2}.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) Ag<sub>2</sub>O, methyliodide, acetone; (b) TFA, triethyl silane, DCM; (c) Z-(*R*)-Val-OH, HATU, DIPEA, DMF; (d) Pd-C, 95% ethanol, H<sub>2</sub>; (e) Z-(*S*)-Val-OH, HATU, DIPEA, DMF.

Z-group was deprotected using catalytic hydrogenation, which yielded the P3 building block  $55\{1\}$ . The P3 building block  $55\{2\}$  was synthesized by the same protocol but using the (*S*)-amino acids instead of the (*R*)-amino acids (Scheme 5.6). Synthesis of  $55\{3\}$  was begun by using Boc<sub>2</sub>O and TEA in dioxane/water (1:1) to achieve Boc protection of *tert*-leucine (64) (Scheme 5.7). The acid was coupled with 63 ((*S*)-amino-cyclohexyl-acetic acid methyl ester), an intermediate formed during synthesis of  $55\{2\}$ , which yielded the amide 65. The Boc group was removed using TFA and TES in DCM, and that gave the P3/P4 building block  $55\{3\}$ . The methylamide derivative  $55\{4\}$  was synthesized from 65, an intermediate in the

Scheme 5.7. Synthesis of the P3/P4 building blocks 55{3} and 55{4}.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) Boc<sub>2</sub>O, TEA, dioxane/water (1:1); (b) 63, HATU, DIPEA, DMF;
(c) TFA, TES, DCM; (d) LiOH, dioxane/water (1:1); (e) methylamine, DIPEA, HATU, DMF.

Scheme 5.8. Synthesis of the P3/P4 building block 55{5}.<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) paraformaldehyde, p-TsOH, toluene; (b) TFA, triethyl silane, DCM; (c) HATU, DIPEA, methylamine, DMF; (d) piperidine, DMF; (e) Boc-*tert*-leucine, DIPEA, HATU, DMF; (f) TFA, triethylsilane, DCM.

synthesis of **55**{*3*}. Thereafter, the methylester **65** was hydrolyzed, the acid was coupled with methylamine, and, lastly, the Boc group was removed using TFA (Scheme 5.7). The *N*-methylated species **55**{*5*} was synthesized from the commercially available **66** (Fmoc-(*S*)-cyclohexylglycine-OH) (Scheme 5.8). First the amide nitrogen in **66** condensed with paraformaldehyde to give an imine, which was attacked intramolecularly by the carboxylic acid to form an oxazolidinone intermediate.<sup>76</sup> The oxazolidinone was hydrolyzed with TFA to yield the *N*-methylated compound, which was coupled with methylamine to give the amide **67**. The Fmoc-group was removed with piperidine in DMF, and the resulting amine was coupled with Boc-*tert*-leucine, and finally the Boc group was removed using TFA to yield the P3/P4 building block **55**{*5*}.

The P2-substituent 7-methoxy-2-phenyl-quinolin-4-ol ( $R^2OH$ ) was synthesized as described in the literature from ethyl benzoylacetate and *m*-anisidine through imine formation and a high temperature condensation reaction (ethanol as leaving group) (Scheme 5.9).<sup>81f</sup>

To summarize the syntheses, it can be said that a total of 63 novel compounds were produced, and they were characterized by NMR. Also, 18 final products were synthesized and tested for activity against the HCV NS3 protease (Section 5.5).

Scheme 5.9. Synthesis of the P2-substituent (R<sup>2</sup>OH).<sup>a</sup>



<sup>a</sup>Reagents and conditions: (a) 4 M HCl, reflux in toluene; (b) diphenyl ether, 280 °C.

#### 5.3.1 Chemical stability

A successful route to the final products in series **44a** was described in the previous section. However, as will be shown here, these compounds can be unstable, so it was necessary to implement careful control of reaction conditions in the synthetic route. It is a known fact that 1,2-unsaturated esters are susceptible to 1,4-nucleophilic attack, which initially presented a problem in the synthesis of the ethers in series **53**. When we used excesses (> 2 equiv.) of



Figure 5.11. The problems associated with Michael's addition, along with the attempted synthesis of  $53{2}$ .

triphenylphosphin and DIAD, we obtained not only the desired ether, but also the 1,4-DIAD adduct and this byproduct was confirmed by NMR and Maldi-TOF analyses (for spectra see supporting info to *paper III*) (Figure 5.11). To circumvent this undesired reactivity in the continued synthesis of ethers **53** less excess of reagents was generally employed. At this point, it is also important to realize the base liability of the target compounds in series **44**. Our



Figure 5.12. The base-induced  $\beta$ -elimination of quinolines.

earlier synthetic pathways had used the ethylester variant of the P1 building block **58**, but we found that, during hydrolysis, the P2 substituent underwent  $\beta$ -elimination more readily than the ethylester was hydrolyzed. (Figure 5.12).<sup>68</sup> NMR and LC/MS data are provided as supplementary information in *paper III*. We tried several different bases and conditions to help diminish or eliminate the  $\beta$ -elimination problem, but all failed. These attempts included the hydrolysis reagents LiOH (1–8 equiv.), KOH (1–8 equiv.), HCl (6 and 12 M), pig liver

esterase and LiOH (1 equiv.), KOSiM<sub>3</sub>, TFMSA, and BBr<sub>3</sub>. Microwave assisted organic chemistry LiOH (1–8 equiv.) or HCl (6 M). Transesterification reagents *tert*-butylacetate and potassium *tert*-butoxide (0.05–3 equiv.).<sup>77</sup>

In summary, these compounds show some susceptibility of 1,4 Michael's addition reactivity and the potential of quinoline  $\beta$ -elimination.

## 5.4 Structure determination of the target compounds using 2D-NMR

Due to the lack of success in synthesizing a chiral scaffold, it was necessary to perform preparative LC/MS to be able to isolate some members of series 44 as chiral compounds. Based on activity measurements, we assumed that the most active diastereomer would have the same stereochemistry as the lead compound 41, in other words that it would have 1R,4R stereochemistry in the cyclopentene ring. To confirm the stereochemistry we subjected 44a $\{2,3\}$  and 44b $\{2,3\}$  to TOCSY and NOE (ROESY and NOESY) analysis.<sup>68</sup> In a TOCSY spectrum, it is possible to see the couplings of entire spin systems (Figure 5.13, *left*). That is



**Figure 5.13**. (*Left*) The total correlation systems that include the PnNH hydrogens in **11**{2,3}. As can be seen, all PnNH hydrogens are part of different spin systems, hence a TOCSY spectrum reveals the chemical shifts of the PnNH hydrogens. (*Right*) Overlay of energy-minimized **44a**{2,3} (red) and **44b**{2,3} (blue) showing theoretical NOE interactions between P3NH and P2 $\alpha$ H in **44a** {2,3} (M = medium) and **44b**{2,3} (S = strong).

to say, because all the Pn $\alpha$ H (P1 $\alpha$ H, P2 $\alpha$ H, P3 $\alpha$ H, and P4 $\alpha$ H) are part of different spincorrelation systems, a TOCSY spectrum will reveal the chemical shift of both the Pn $\alpha$ H and the corresponding PnNH. This information is needed to interpret NOE spectra, which disclose spatial relationships between hydrogens and therefore make it possible to determine which hydrogens are in proximity to each other. According to computer models, there should be a relevant difference in the distance between the P3NH and P2 $\alpha$ H of the diastereomers, making it possible to distinguish between them (Figure 5.13, *right*).<sup>78</sup> In Figures 5.14 and 5.15 (showing TOCSY spectra of **44a**{2,3} and of **44b**{2,3}), there are a total of four correlation lines, which respectively correspond to the P1, P2, P3, and P4 spin systems.



Figure 5.14. The designations of 44a {2,3} (top) and its TOCSY spectrum (below).



The first line (lowest chemical shift) of  $44a\{2,3\}$  is the P4 $\alpha$ H line, followed by the P2 $\alpha$ H, P1 $\alpha$ H, and P3 $\alpha$ H lines. There is a slightly different order for 44b{2,3}, namely first P4 $\alpha$ H, then P1 $\alpha$ H and P2 $\alpha$ H, and last the P3 $\alpha$ H line. The ROESY and NOESY spectra were identical with regard to the NOE interactions; we chose the ROESY spectra for further studies, and the important parts of the spectra can be seen in Figure 5.16. Figure 5.17 presents plausible configurations (energy minimized) of the diastereomers, which are used to explain the measured NOE interactions.<sup>78,79</sup> The closer the distance between the molecules, the stronger the NOE interaction, and a useful guide is as follows: strong cross peaks, 2.0–2.5 Å; medium cross peaks, 2.0-3.0 Å; weak cross peaks, 2.0-5.0 Å.<sup>80</sup> The P3NH and P4NH show the expected NOE interactions between their corresponding  $\alpha$ -hydrogens. The NOE effects between P1NH and P1 $\alpha$ H (16 $\leftrightarrow$ 17) is hard to see due to proximity to the stronger P1NH $\leftrightarrow$ P2 $\alpha$ H (16 $\leftrightarrow$ 15) NOE interaction. It is also a possibility that such interaction between P1NH and P1 $\alpha$ H (16 $\leftrightarrow$ 17) is really weak (2.95 Å in model) and actually undetectable. Moreover, the NOE interactions between P4NH and P3 $\alpha$ H ( $8 \leftrightarrow 9$ ) and between P1NH and  $P2\alpha H$  (16 $\leftrightarrow$ 15) are clearly visible. As illustrated in Figure 5.16, the intensity patterns of these two diastereomers are similar, with the exception of the NOE interaction for P1NH and P2 $\alpha$ H  $(11 \leftrightarrow 15)$ . Based on these results, it can be concluded that the structure-determining interaction between P1NH and P2 $\alpha$ H (11 $\leftrightarrow$ 15) is stronger in 44b{2,3} than in 44a{2,3}.

The weaker NOE interaction of  $11 \leftrightarrow \underline{15}$  in  $44a\{2,3\}$  compared to  $44b\{2,3\}$  indicates that  $44a\{2,3\}$  has the 1R,4R stereochemistry in the cyclopentene ring.



**Figure 5.16.** The designations of  $44a\{2,3\}$  (*top left*) and  $44b\{2,3\}$  (*top right*), and the ROESY spectra showing the interactions between the Pn-NH with Pn- $\alpha$ H hydrogens in  $44a\{2,3\}$  (*lower left*) and  $44b\{2,3\}$  (*lower right*).



**Figure 5.17.** Models of **44a**{2,3} (*left*) and **44b**{2,3} (*right*) showing the theoretical distances (Å) between Pn-NH $\leftrightarrow$ Pn- $\alpha$ H.

## 5.5 Structure-activity relationships of the final products

The final products in series **44** are very active in enzyme assays, so the SAR data on these compounds can be used as a tool in future projects focused on scaffold hopping (mimicking). All inhibitors discussed below contain the 7-methoxy-2-phenyl-quinolin-4-ol group as P2 substituent, because this group has had ample precedence in previous work, and it apparently stabilizes the catalytic machinery in the correct geometry by shielding the catalytic triad of the protease from exposure to the solvent. <sup>81a,82,83</sup> In addition, this P2 substituent probably interacts with the helicase domain of the NS3 protein, which enhances the overall activity.<sup>84</sup> The final products in series **44a** are more potent then the corresponding diastereomers in series **44b** (Table 5.2). The P2 hydroxycyclopentene moiety in series **44a** has most likely the same stereochemistry as the L-hydroxyproline moiety that has been incorporated in the inhibitors described in section 5.1 (Figures 5.4 and 5.5).<sup>48,81,85</sup>

In continuation, comparison of inhibitors incorporating the smallest P1 amino acid  $8{1}$ with inhibitors incorporating the homoanalog  $8_{\{2\}}$  indicates that the extension of one carbon atom in the P1 position increases the activity roughly four times; this is exemplified by, for example, compounds  $44\{1,4\}$  and  $44\{2,4\}$  (280 and 67 nM, respectively). Inhibitors containing the vinylcyclopropyl amine  $8{3}$  in the P1 position exhibit dramatically improved activity, which is demonstrated by comparison of, for instance,  $44a\{2,3\}$  and  $44a\{3,3\}$  (30) and 1.1 nM, respectively).<sup>86</sup> In the P3/P4 position, the S,S stereochemistry is preferred for all reported inhibitors using the P2 hydroxyproline building block.<sup>85</sup> Although this is true for those inhibitors, the novel P2 cyclopentene scaffold presented in our studies introduces a reversal of direction of the peptide chain at the P2 position, which results in inhibitors with two carboxy termini and also introduces a one-carbon extension along the P3/P4 amino acid chain. It might be expected that addition of a retro proline mimetic at P2 would reverse the S,S stereochemical preference at P3/P4. However, the one-atom extension should again reverse the preference back to S.S. By comparing the 2D-structures of 1 and  $44a_{3,4}$ , it seemed likely that the S.S stereochemistry would be more favorable (Figure 1B and 1D), and this hypothesis was supported by a 3D computer overlay analysis.<sup>78</sup> Further verification was provided by comparison of the activity of the P3/P4 R,R-derivative with that of the corresponding S.S-derivative, which made it evident that the S.S-derivative 44{1,2} ( $K_i$  523 nM) indeed was more potent then the R,R-derivative  $44a\{1,1\}$  (K<sub>i</sub> >2000 nM). The S3 pocket is considered to be relatively small, but, despite that, the activity was enhanced when a valine was replaced with a *tert*-butyl glycine moiety in P3 (*cf.*,  $44a\{2,2\}$  and  $44a\{2,3\}$ , 102 and 30 nM, respectively). It is possible that even bulkier P3 groups would further improve the activity. The interaction with the S4 pocket is very important for activity, as exemplified by the P1–P3 derivative  $44\{1,6\}$  (> 2000 nM) and the P1–P4 derivative  $44\{1,2\}$  (523 nM). The

Compd	Structure	<i>K</i> <sub>i</sub> (nM)	Compd	Structure	$K_i$ (nM)
<b>44a</b> { <i>1,1</i> }		>2000	44a <i>{3,3}</i>	$ \begin{array}{c} \begin{array}{c} \begin{array}{c} \begin{array}{c} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \\ \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \\ \end{array} \\ \end{array} \\ \end{array} \\ \end{array} \\ \begin{array}{c} \\ \\ \end{array} $	1.1
<b>44</b> { <i>1,2</i> } ( <i>anti</i> )		523	44a <i>{3,4}</i>		1.3
<b>44</b> { <i>1,3</i> } ( <i>anti</i> )		503	<b>44</b> {3,5} ( <i>anti</i> )	$\begin{array}{c} & & & \\ & & & \\ & & & \\ & & & \\ H \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & &$	>5000
<b>44</b> {1,4} (anti)		280	<b>44b</b> {2,2}		2560
<b>44</b> { <i>1,6</i> } ( <i>anti</i> )	OR2 NH OH	>2000	44b{2,3}		>2000
44a{2,2}		102	44b <i>{3,3}</i>	$ \begin{array}{c} & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & $	108
44a{2,3}		30	<b>44b</b> { <i>3,4</i> }		23
<b>44</b> {2,4} ( <i>anti</i> )	HN C NH O HN C OH	67	<b>56</b> {3,3} (anti)		>2000
<b>44</b> {3,2} ( <i>anti</i> )	$ \begin{array}{c} & & & \\ & & & & \\ & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & & & & \\ & $	24	<b>56</b> {3,4} (anti)		>2000

**Table 5.2.** Activity data on target compounds.

free carboxylic acid in the P1 position is essential for HCV NS3 protease inhibitory properties, which can be compared with, for example, the acid  $44a\{3,4\}$  (1.3 nM) and the *tert*-butylester  $13\{3,4\}$ , the latter of which is devoid of all activity. As previously reported, the carboxylic acid of the inhibitor probably forms hydrogen bonds with Ser139 in the catalytic triad of the protease.<sup>86</sup> The methyl ester derivative  $44a\{3,3\}$  (1.1 nM) and the methyl amide derivative  $44a\{3,4\}$  (1.3 nM) are virtually equipotent, which suggests that a hydrogendonating bond is not required at this position. In contrast, *N*-methylation of the P4NH derivative  $44a\{3,4\}$  (1.3 nM) to give the P4NMe derivative  $44a\{3,5\}$  (> 5000 nM) resulted in a dramatical loss of inhibitory activity, suggesting that the P4NH is involved in a hydrogen bond network with the NS3 enzyme.

## 5.6 Conclusions

In the present work, a novel synthetic route was developed to the hydroxyproline bioisostere **50**, and the regio and stereochemistry of this scaffold was confirmed by NOESY and HMBC spectra. The scaffold **50** was derivatized with various amines and the quionoline 7-methoxy-2-phenyl-quinolin-4-ol, which yielded the final products in series **44**. The stereochemistry of products **44a**{2,3} and **44b**{2,3} was verified by recording TOCSY and ROESY NMR spectra in combination with computer analysis. A total of eighteen inhibitors of the HCV NS3 protease were synthesized in the series **56**, **44a**, and **44b**, and the most potent of these has an impressive Ki of 1.1 nM (Table 5.2). The SAR data obtained should provide a useful tool in future projects focused on scaffold hopping.





# **6** Acknowledgments

I would like to thank the following people:

Professor **Ingemar Kvarnström** (University of Linköping) for being an excellent supervisor and for stopping me when I got carried away in matters concerning synthesis and writing. Professor **Bertil Samuelsson** (Medivir AB) for never-ending ideas about how to improve

Professor **Bertil Samuelsson** (Medivir AB) for never-ending ideas about how to improve syntheses and manuscript preparation.

Dr. **Ingemar Nilsson** (AstraZeneca R&D, Mölndal) for providing professional assistance with medicinal chemistry, both in the writing of papers and at meetings.

Dr. Åsa Rosenquist (Medivir AB) for improvement of manuscripts and practical advice.

**AstraZeneca R&D, Mölndal**, for financial support, and the following people who work at that facility:

Dr. Djordje Musil for spending his nonexistent sparetime on the X-ray structure of  $14\{14\}$ , Dr. Karl-Erik Karlsson for time spent at the HRMS machine on my account, Dr. Shalini Andersson for letting Hanna Nelander be my personal chiral HPLC consultant, Hanna Nelander for performing the chiral separation of  $44\{1\}$ , Marléne Fredenwall and Peter Abrahamsson for practical help with synthesizing the combinatorial library reported in *paper I*, and Erika Gyzander and the Biochemistry Group for performing the biological tests reported in *paper II*.

Medivir AB for financial support, as well as the following people who work at that company:
 Taitana Maltseva for recording excellent NMR spectra, Lotta Vrang and Elizabeth
 Hamelink and the biochemistry group for biological testings reported in *papers III*, and Hans Wallberg and Xiao-Xiong Zhou for good ideas at meetings.

Past and present colleagues in the Kvarnström group:

**Per Ola Johansson** for being an excellent wingman during our years of PhD studies together; **Anders Dahlgren** for practical help and all discussions about global issues and weird things; **Jonas Brånalt** for sending chemistry articles to me and allowing me stay at "Hotel Balltorp"; **Yantao Chen** for sharing vital knowledge of Chinese culture (Kung Fu movies); **Jonas Nilsson** for being a friend and valuable sounding board; **David Pyring** and **Nelly Bibak** for sharing with me an office, vital knowledge, and nonsense; **Marcus Bäck** for helpfulness and for being the 747 man; **Veronica Sandgren** for being kind and cheerful; **Fredrik Wångsell** the heir to my office and my chair,—thanks for being a good roommate and chemist for the past two years.

My Master's thesis students who have assisted me in my projects:

Patrick Johansson, Elisabet Palmqvist, Cecelia Andresen, Linda Karis Fredrik Wångsell, Maria Wikberger, Maja Wiklund, Camilla Romanov, Veronica Sandgren, Per-Olof Bergh, and Malek Mirzayan.

Good ol' days, the early years (1999–2002):

Professor Uno Carlsson, Karin Carlsson, Magdalena Svensson, Malin Person, Per Hammarström, Dick Andersson, and Martin Karlsson mainly for being there at lunch time during my first years at the Department, but also for some enlightenment regarding biochemistry and beer. Johan Ekerot, Jan Lundqvist, Jonas "Baltzer" Nilsson, Anders Dahlgren, Daniel Nöteberg, Jonas Nilsson, Per-Ola Johansson, Jonas Brånalt, Yantao Chen and Joakim Svensson for all the corridor conversations and after-work activities.

The new deal, the present day (2003–2005):

Markus Hederos, Andreas Åslund, Per-Ola Johansson, Fredrik Wångsell, Marcus Bäck, Timmy Fyrner, Patrik Nygren, Johan Viljanen, Tysken (Janosch) and the quisling Johan Olson who went to Stockholm. The Vallfarten lunch troika over the last two years—I want to thank them for all the coffee and other beverages we have drunk together, all the rock concerts we went to, and for making the organic chemistry corridor a cozy place to be.

**Patricia Ödman** for linguistic help with this thesis. **Anders Hallberg** for help with manuscripts.

Mina föräldrar, **Inger och Sören Thorstensson**, för stöd, omtanke och visad förståelse och min bror, **Frank Thorstensson**, för att han aldrig ger upp och för att han är min bror.

Anngelica Jarl—Tack för att du finns och står ut med mig.

# 7 References

- B. Jönsson, Intervju med Dr. Jan Lundberg (head of discovery research, AstraZeneca). Kemist söker kemister, 2002, *Kemivärlden*, 31-32.
- (2) Teague, S. J.; Davis, A. M.; Leeson, P. D.; Oprea, T. I. The design of leadlike combinatorial libraries *Angew. Chem. Ind. Ed.* **1999**, *38*, 3743-3748; German version: *Angew. Chem.* **1999**, *111*, 3962-3967.
- (3) Morita, S.; Ootsubo, K.; Uchida, M. 1993 JP 05078326.
- (4) Bell, I. M.; Freidinger, R. M.; Guare , J. P.; Sparks, M. A.; Williams, P. D. 1999 GB 02326639; US 5968938.
- (5) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings. *Adv. Drug Deliv. Rev.* **1997**, *23*, 3-25.
- Bigge, C. F.; Hays, S. J.; Novak, P. M.; Drummond, J. T.; Johnson, G.; Bobovski, T. P. New preparations of the N-methyl-D-aspartate receptor antagonist, 4-(3-phosphonopropyl)-2-piperazinecarboxylic acid (CPP). *Tetrahedron Lett.* 1989, *30*, 5193-5196.
- (7) Hamann, B., C.; Hartwig, J. F. Systematic variation of bidentate ligands used in aryl halide amination. Unexpected effects of steric, electronic, and geometric perturbations. *J. Am. Chem. Soc.* **1998**, *120*, 3694-3703.
- (8) John P. Wolfe and Stephen L. Buchwald, Scope and limitations of the Pd/BINAPcatalyzed amination of aryl bromides. John P. Wolfe and Stephen L. Buchwald, J. Org. Chem. 2000, 65, 1144-1157.
- (9) Arumugam, V.; Routledge, A.; Abell, C.; Balasubramanian, S. Synthesis of 2-oxindole derivatives via the intramolecular Heck reaction on solid support. *Tetrahedron Lett.* 1997, *38*, 6473-6476.
- (10) Hari, A.; Miller, B. L. A new method for the mild and selective reduction of aryl nitro groups on solid support. *Tetrahedron Lett.* **1999**, *40*, 245-248.
- (11) Phillips, G. B.; Wei, G. P. Solid phase synthesis of benzimidazoles. *Tetrahedron Lett.* 1996, 37, 4887-4890.
- (12) Pei, Y.; Houghten, R. A.; Kiely, J. S. Synthesis of 4-amino-3,4-dihydro-2(1H)quinolinonesvia β-lactam intermediates on the solid-phase. *Tetrahedron Lett.* **1998**, *38*, 3349-3352.
- (13) Schechter, I., Berger, A. On the size of the active site in proteases. I. Papain. *Biochem. Biophys. Res. Commun.* 1967, 27, 157-162.

- (14) Fang, J.; Alderman, M. H. Dissociation of hospitalization and mortality trends for myocardial infarction in the United States from 1988 to 1997. *Am. J. Med.* 2002, *113*, 208–214.
- (15) Gustafsson, D.; Bylund, R.; Antonsson, T.; Nilsson, I.; Nyström, J-E.; Eriksson, U.; Bredberg, U.; Teger-Nilsson, A-C. A new oral anticoagulant: the 50-year challenge. *Nature Revivews Drug Discovery*, **2004**, *3*, 649-659.
- (16) Hirsh, J.; Dalen J.; Anderson D. R.; Poller L.; Bussey H.; Ansell J.; Deykin D. Oral anticoagulants: mechanism of action, clinical effectiveness, and optimal therapeutic range. *Chest*, **2001**, *119* (Suppl. 1), 8–21.
- (17) Hirsh J.; Warkentin T. E.; Shaughnessy S. G.; Anand S. S.; Halperin J. L.; Raschke R.; Granger C.; Ohman E. M.; Dalen J. E. Heparin and low-molecular-weight heparin: mechanisms of action, pharmacokinetics, dosing,monitoring, efficacy, and safety. *Chest*, **2001**, *119* (Suppl. 1), 64–94.
- (18) Bettelheim, F. R. Clotting of fibrinogen. II. Fractionation of peptide material liberated. *Biochem. Biophys. Acta*, **1956**, *19*, 121–130.
- (19) Blombäck, B.; Blombäck, M.; Olsson, P.; Svendsen, L.; Åberg, G. Synthetic peptides with anticoagulant and vasodilating activity. *Scand. J. Clin. Lab. Invest.* **1969**, *24*, (Suppl. 107), 59–61.
- (20) (a) Ni, F.; Meinwald, Y. C.; Vasquez, M.; Scheraga, H. A. High-resolution NMR studies of fibrinogen-like peptides in solution: structure of a thrombin-bound peptide corresponding to residues 7-16 of the A alpha chain of human fibrinogen. Biochemistry 1989, 28, 3094–3105. (b) Stubbs M. T.; Oschkinat H.; Mayr I.; Huber R.; Angliker H.; Stone S. R.; Bode W. The interaction of thrombin with fibrinogen. A structural basis for its specificity. *Eur. J. Biochem.* 1992, 206, 187–195.
- (21) <u>This article is important in the development of the D-Pro-Arg based inhibitors:</u> Bajusz, S; Barabas, E.; Tolnay, P.; Szell, E; Bagdy, D. Inhibition of thrombin and trypsin by tripeptide aldehydes. *Int. J. Pept. Protein Res.* **1978**, *12*, 217-221.
- (22) <u>This article describes the X-ray of PPACK with alpha-thrombin, important starting point for the D-Phe-Pro-Arg inhibitors developed by AstraZeneca:</u> Bode, W.; Mayr, I.; Baumann, U.; Uber, R.; Stone, S. R.; Hofsteenge, J. The refined 1.9 A crystal structure of human alpha-thrombin: interaction with D-Phe-Pro-Arg chloromethylketone and significance of the Tyr-Pro-Pro-Trp insertion segment. *Embo. J.* **1989**, *8*, 3467-3475.
- (23) (a) Kettner, C.; Shaw, E. D-Phe-Pro-ArgCH2C1-A selective affinity label for thrombin. *Thromb. Res.* 1979, *14*, 969-973. (b) Stubbs. M. T.; Bode, W. Crystal structures of thrombin and thrombin complexes as a framework for antithrombotic drug design. *Perspectivies in drug discovery and design* 1994, *1*, 431-440.

- (24) (a) Bajusz, S. The story of D-Phe-Pro-Arg-H, the likely antigcoagulant of the future. *Biokémia* 1990, *14* 127-133 (b) Sanderson, P. E. J.; Naylor-Olsen, A. M. Thrombin inhibitor design *Curr. Med. Chem.* 1998, 289-304. (c) Sanderson, P. J. Small, noncovalent serine protease inhibitors. *Med. Res. Rev.* 1999, *19*, 179-197. (d) Adang, A. E. P.; Rewinkel, J. B. M. A new generation of orally active antithrombotics: CPIIb/IIIa, thrombin, and factor Xa inhibitors. *Drugs of the Future* 2000, *25*, 369-383. (e) Ripka, W. C.; Vlasuk, G. P. Antithrombotics/serine proteases. *Annual Reports in Medicinal Chemistry* 1997, *32*, 71-89 (f) Wiley, M. R.; Fisher, M. J. Small-molecule direct thrombin inhibitors. *Exp. Opin. Ther. Patents* 1997, *7*, 1265-1282.
- (25) (a) Babine, R. E.; Bender, S. L. Molecular recognition of protein-ligand complexes: applications to drug design. *Chem. Rev.* 1997, 97, 1359-1472. (b) Leung, D.; Abbenante, G.; Fairlie, D. P. Protease inhibitors: current status and future prospects. *J. Med. Chem.* 2000, 43, 305-320 (c) Ripka, A. S.; Rich, D. H. Peptidomimetic design. *Current Opinion in Chemical Biology*, 1998, 2, 441-452. (d) Antonsson T.; Bylund R.; Gustafsson D.; Nilsson I.; PCT Int Appl WO 94/29339; 1994; Method of thrombin inhibition. US-5,939,392; 1999.
- (26) <u>The patented synthesis of Pab:</u> Antonsson, T.; Bylund, R.; Gustafsson, N.; Nilsson, N. Trypsin-like protease-inhibiting peptide derivatives, their synthesis and therapeutic use. PCT Int Appl WO 94/29336, 1994.
- (27) Nöteberg, D.; Brånalt, J.; Kvarnström, I.; Linschoten, M.; Musil, D.; Nyström, J. E.; Zuccarello G.; Samuelsson, B. New Proline mimetics: synthesis of thrombin inhibitors incorporating cyclopentane- and cyclopentenedicarboxylic acid templates in the P2 position. binding conformation investigated by X-Ray crystallography. *J. Med. Chem.* **2000**, *43*, 1705-1713.
- (28) Dobarro A.; Velasco D. Synthesis of β-γ-Unsaturated N-Acyl-2-oxazolidinones. *Tetrahedron*, **1996**, *52*, 13525-13530.
- (29) (a) Buchmeiser M. R.; Homogenous metathesis polymerization by well-defined group VI and group VIII transition-metal alkylidenes: fundamentals and applications in the preparation of advanced materials. *Chem. Rev.* 2000, *100*, 1565-1604. (b) Nicolaou, K.C.; Bilger, P. G.; Sarlah, D. Metathesis reactions in total synthesis. *Angew.* 2005, *44*, 4490-4527.
- (30) (a) Stephen J. Connon, Siegfried Blechert. Recent Developments in Olefin Cross-Metathesis. *Angew. Chem. Int. ed.* 2003, *42*, 1900-1923. (b) Amblard, F.; Nolan, S. P.; Agrofoglio L.A. Metathesis strategy in nucleoside chemistry. *Tetrahedron*, 2005, *61*, ,7067-7080.
- (31) (a) Paquette, L. A.; Schloss, J. D.; Efremov, I.; Fabris, F.; Gallou, F.; Méndez-Andino, J.; Yang. J. A convenient method for removing all highly-colored byproducts generated during olefin metathesis reaction. *Org. Lett.* 2000, *2*, 1259-1261 (b)

Maynard H. D.; Grubbs, R. H. Purification technique for removal of ruthenium from olefin metathesis reaction products. *Tetrahedron Lett.* **1999**, *40*, 4137-4140. (c) Ahn, Y. M.; Yang, K.; Georg, G. I. A convenient method for the efficient removal of ruthenium byproducts generated during the olefin metathesis reactions. *Org.Lett.* **2001**, *3*, 1411-1413.

- (32) <u>Selected patents from the 1950s:</u> (a) E.I. du Pont de Nemours & Co.; British Patent 682,420 (November 12, 1952). (b) A.W. Anderson, et al.; U. S. Patent 2,905,645 (September 22, 1959) (Application August 16, 1954). (c) A. W. Anderson, et al.; French Patent 1,134,740 (December 3, 1956).
- (33) <u>The first scientific article describing a successful olefin metathesis reaction:</u> Truett W. L.; Johnson D. R.; Robinson I. M.; Montague B. A. Polynorbonene by Coördination Polymerization. J. Am. Chem. Soc. **1960**, 82, 2337-2340.
- (34) <u>The first article that correctly namegives the olefin metathesis reaction and suggests a</u> <u>mechanism:</u> Calderon N.; Ofesead E. A.; Ward J. P.; Judy W. A.; Scott K. W. Olefin Metathesis. I. Acyclic vinylenic hydrocarbons. J. Am. Chem. Soc. **1968**, 90, 4133-4140.
- (35) <u>Two examples on the first postulated correct mechanism:</u> (a) Dall'Asta, G.; Motroni, G. Site of ring cleavage in the ring-opening polymerization of low strained cycloolefins. *Eur. Polym. J.* **1971**, *7*, 707-710. (c) Hérisson, J.-L; Chauvin, Y. Catalysis of olefin transformations by tungsten complexes. II. Telomerization of cyclic olefins in the presence of acyclic olefins. *Makromol. Chem.* **1971**, *141*, 161-176.
- (36) <u>Proof of the postulated mechanism:</u> (a) Casey, C. P.; Burkhardt, T. J. Reactions of (diphenylcarbene)pentacarbonyltungsten(0) with alkenes. Role of metal-carbene complexes in cyclopropanation and olefin metathesis reactions. 1974, 96, 7808–7809.
  (b) Kress J.; Osborn J. A.; Greene R. M. E.; Ivin K. J.; Rooney J. J. First direct observation of the simultaneous presence and of the interconversion of chain-propagating metal-carbene and metallacyclobutane complexes in a catalytic olefin metathesis reaction: the ring-opening polymerization of norbornene. J. Am. Chem. Soc. 1987, 109, 899-901.
- (a) Examples on metathesis chemistry prior to well-defined catalyst: Kroll W. R.; (37) Doyle G.; Carbene complexes of the group VI metals as olefin disproportionation catalysts. J. Chem. Soc. 1971, 839. (b) Casey, C. P.; Burkhardt, T. J. (Diphenylcarbene)pentacarbonyltungsten(0). J. Am. Chem. Soc. 1973, 95, 5833-5834 Т (c) Casey, С. P.; Burkhardt, J. Reactions of (diphenylcarbene)pentacarbonyltungsten(0) with alkenes. Role of metal-carbene complexes in cyclopropanation and olefin metathesis reactions. J. Am. Chem. Soc. 1974, 96, 7808-7809. (d) Katz, T.J.; McGinnis, J.; Altus, C. Metathesis of a cyclic

trisubstituted alkene. Preparation of polyisoprene from 1-methylcyclobutene. J. Am. Chem. Soc. 1976, 98, 606-608.

- (38) <u>An example of Schrock's articles prior the discovery of Schrock catalyst:</u> Schrock R.
   R. On the trail of metathesis catalysts. J. Organomet. Chem. **1986**, 300, 249-262.
- (39) <u>The first reported synthesis of the Schrock catalyst:</u> Schrock R. R.; Mordzek John S.; Bazan G. C.; Robbins J.; Dimare M.; O' Regan M. Synthesis of molybdenum imido alkylidene complexes and some reactions involving acyclic olefins. J. Am. Chem. Soc. 1990, 112, 3875-3886.
- (40) <u>The first articles on the first generation Grubbs catalysts:</u> (a) Nguyen S. T.; Grubbs R. H. Syntheses and activities of new single-component, ruthenium-based olefin metathesis catalysts. J. Am. Chem. Soc. 1993, 115, 9858-9859. (b) Fu G. C.; Nguyen T.; Grubbs R. H.; Catalytic ring-closing metathesis of functionalized dienes by a ruthenium carbene complex. J. Am. Chem. Soc. 1993, 115, 9856-9857.
- (41) <u>The first time the optimized benzylidene catalyst 24 is described:</u> Schwab, P.; France, M. B.; Ziller, J. W.; Grubbs, R. H. A series of well-defined metathesis catalysts synthesis of [RuCl<sub>2</sub>(:CHR')(PR<sub>3</sub>)<sub>2</sub>] and their reactions. *Angew. Chem. Int. Ed. Engl.* 1995, *34*, 2039-2041.
- (42) <u>The first articles on the second generation Grubbs catalyst:</u> (a) Scholl M.; Trnka T. M.; Morgan J. P.; Grubbs R. H. Increased ring closing metathesis activity of ruthenium-based olefin metathesis catalysts coordinated with imidazolin-2-ylidene ligands *Tetrahedron Lett.* **1999**, *40*, 2247-2250 (b) Chatterjee A. K.; Grubbs R. H. Synthesis of trisubstituted alkenes via olefin metathesis. *Org. Lett.* **1999**, *1*, 1751-1753.
- (43) <u>The first article on the Grubbs/Hoveyda catalyst:</u> Garber, S. B.; Kingsbury, J. S.; Gray, B. L; Hoveyda, A. H. Efficient and recyclable monomeric and dendritic Rubased metathesis catalysts. J. Am. Chem. Soc. 2000, 122, 8168 8179.
- (44) Kirkland, T. A.; Grubbs, R.H. Effects of olefin substitution on the ring-closing metathesis of dienes. *J. Org. Chem.* **1997**, *62*, 7310-7318.
- (45) Fürstner, A.; Langermann, K. Total synthesis of (+)-ricinelaidic acid lactone and of (-)-gloesporane based on transition-metal-catalyzed C-C bond formations. J. Am. Chem. Soc. 1997, 119, 9130-9136.
- (46) Agami, C.; Couty, F.; Rabasso, N. Synthesis of (-)-β conhydrine and analogues using N-Boc-2-acyl oxazolidine methodology and ring closing metathesis. *Tetrahedron* 2000, 41, 4113-4116.
- (47) Bomann , M. D.; Guch , I. C.; DiMare , M. A mild, pyridine-borane-based reductive amination protocol. *J. Org. Chem.* **1995**, *60*, 5995-5996.
- (48) (a) De Francesco, R.; Tomei, L.; Altamura, S.; Summa, V.; Migliaccio, G. Approaching a new era for hepatitis C virus therapy: inhibitors of the NS3-4A serine

protease and the NS5B RNA-dependent RNA. *Antiviral Res.* **2003**, *58*, 1-16. (b) Griffith, R. C.; Lou, L.; Roberts, C. D.; Schm, U. HCV Anti-viral agents. *Annu. Rep. Med. Chem.* **2005**, *39*, 223-237. (c) Ni, Z-J.; Wagman, A. S. Progress and development of small molecule HCV antivirals. *Curr. Opin. Drug Discov. Devel.* **2004**, *7*, 446-459. (d) Goudreau, N.; Llinàs-Brunet, M. The therapeutic potenial of NS3 protease inhibitors in HCV infection. *Expert Opin. Investig. Drugs* **2005**, *14*, 1129-1144.

- (49) (a) Choo, Q. L.; Kuo, G.; Weiner, A. J.; Overby, L. R.; Bradley, D. W.; Houghton, M. Isolation of a cDNA clone derived from a blood-borne non-A, non-B viral hepatitis genome. *Science* 1989, *244*, 359-362. (b) Kuo, G.; Choo, Q. L.; Alter, H. J.; Gitnick, G. L.; Redeker, A. G.; Purcell, R. H.; Miyamura, T.; Dienstag, J. L.; Alter, M. J.; Stevens, C. E.; An assay for circulating antibodies to a major etiologic virus of human non-A, non-B hepatitis. *Science* 1989, *244*, 362-364. (c). Di Bisceglie, A. M. Hepatitis C. *Lancet* 1998, *351*, 351–355.
- (50) Griffith, R. C.; Lou, L.; Roberts, C. D.; Schm, U. HCV Anti-Viral Agents. Annual Reports in Medicinal Chemistry 2005, 39, 223-237.
- (51) WHO fact sheet N°164. Revised October 2000. http://www.who.int/mediacentre/factsheets/fs164/en/index.html
- (52) Freeman, A. J.; Dore, G. J.; Law, M. G.; Thorpe, M.; Overbeck, J. O.; Lloyd, A. R.; Marinos, G.; Kaldor, J. M. Estimating progression to cirrhosis in chronic hepatitis C virus infection *Hepatology* **2001**, *34*, 809-816.
- (53) Willems, M.; Metselaar, H. J.; Tilanus, H. W.; Schalm, S. W.; de Man, R. A. Liver transplantation and hepatitis C. *Transpl. Int.* 2002, *15*, 61-72.
- (54) (a) Cheng, S. J.; Bonis, P.; Lau, J.; Pham, N. Q.; Wong, J. B. Interferon and ribavirin for patients with chronic hepatitis C who did not respond to previous interferon therapy: a meta-analysis of controlled and uncontrolled trials. *Hepatology* 2001, *33*, 231-240. (b) Tan, S-L.; Pause, A.; Shi, Y.; Sonenberg, N. Hepatitis C therapeutics: current status and emerging strategies. *Nat. Rev. Drug Discov.* 2002, *1*, 867–881. (c). Di Bisceglie, A. M.; McHutchison, J.; Rice, C. M. New therapeutic strategies for hepatitis C. *Hepatology* 2002, *35*, 224–231. (d) Herrmann, E.; Lee, J. H.; Marinos, G.; Modi, M.; Zeuzem, S. Effect of ribavirin on hepatitis C viral kinetics in patients treated with pegylated interferon. *Hepatology* 2003, *37*, 1351-1358.
- (55) Rosenberg S. Recent advances in the molecular biology of hepatitis C. J. Mol. Biol. 2001, 313, 451-464.
- (56) (a) Reed, K. E.; Rice, C. M. Overview of hepatitis C virus genome structure, polyprotein processing, and protein properties. *Curr. Top. Microbiol. Immunol.* 2000, 242, 55–84. (b) Kolykhalov, A. A.; Mihalik, K.; Feinstone, S. M.; Rice, C. M.; Hepatitis C virus-encoded enzymatic activities and conserved RNA elements in the 3'

nontranslated region are essential for virus replication in vivo. J. Virol. 2000, 74, 2046–2051. (c) Gordon, C. P.; Keller, P. A. Control of hepatitis C: a medicinal chemistry perspective. J. Med. Chem. 2005, 48, 1-20.

- (57) Steinkühler, C.; Biasiol, G.; Brunetti, M.; Urbani, A., Koch, U.; Cortese, R.; Pessi, A.; De Francesco, R. Product inhibition of the hepatitis C virus NS3 protease. *Biochemistry* 1998, 37, 8899-8905.
- (58) Reiser, M.; Hinrichsen, H.; Benhamou, Y.; Reesink, H. W.; Wedemeyer, H.; Avendano, C.; Riba, N.; Yong, C-L.; Nehmiz, G.; Steinmann, G. G. Antiviral efficacy of NS3-serine protease inhibitor BILN-2061 in patients with chronic genotype 2 and 3 hepatitis C. *Hepatology* **2005**, *41*, 832-835.
- (59) Nöteberg, D.; Brånalt, J.; Kvarnström, I.; Linschoten, M.; Musil, D.; Nyström, J. E.; Zuccarello, G.; Samuelsson, B. New proline mimetics: synthesis of thrombin inhibitors incorporating cyclopentane- and cyclopentenedicarboxylic acid templates in the P2 position. Binding conformation investigated by X-ray crystallography. *J. Med. Chem.* 2000, *43*, 1705-1713.
- (60) Thorstensson, F.; Kvarnström, I.; Musil, D.; Nilsson, I.; Samuelsson, B. Synthesis of novel thrombin inhibitors. Use of ring-closing metathesis reactions for synthesis of the P2 cyclopentene- and cyclohexenedicarbarboxylic acid derivatives. *J. Med. Chem.* 2003, 46, 1165-1179.
- (61) Mitsunobu, O. Synthesis, 1981, 1
- (62) (a) Molander, G. A.; Quirmbach, M. S.; Silva, L. F.; Spencer, K.; Balsells, J. Toward the total synthesis of variecolin. *Org. Lett.* 2001, *3*, 2257-2260. (b) Rosenquist, Å.; Kvarnström, I.; Svensson, S. C. T.; Classon, B.; Samuelsson. B. Synthesis of enantiomerically pure *trans*-3,4-substituted cyclopentanols by enzymatic resolution. *Acta Chem. Scand.* 1992, *46*, 1127.
- (63) Dolby, L. J.; Esfandiari, S.; Elliger, C. A.; Marshall, K. S. Studies of the synthesis of the B, C and D rings of gibberlic Acid. J. Org. Chem. 1971, 36, 1277-1285.
- (64) <u>Elimination conditions using PhSeBr</u>: Wachtmeister, J.; Claesson, B.; Samuelsson, B.; Kvarnström, I. Synthesis of 2',3'-dideoxycyclo-2'-pentenyl-3'-C-hydroxymethyl carbocyclic nucleoside analogues as potential anti-viral agents. *Tetrahedron* 1995, *51*, 2029-2038.
- (65) McIldowie, M. J.; Mocerino, M.; Skelton, B. W.; White, A. H. Facile lewis acid catalyzed synthesis of C4 symmetric resorcinarenes. *Org. Lett.* 2000, *2*, 3869-3871.
- (66) (a) Luche, J-L. Lanthanides in organic chemistry. 1. Selective 1,2 reductions of conjugated ketones. J. Am. Chem. Soc. 1978, 100, 2226-2227. (b) Gemal, A. L.; Luche, J-L. Lanthanoids in organic synthesis. 6. Reduction of .alpha.-enones by sodium borohydride in the presence of lanthanoid chlorides: synthetic and mechanistic aspects. J. Am. Chem. Soc. 1981, 103, 5454-5459.

- (67) Honda, M.; Hirata, K.; Sueoka, H.; Katsuki, T.; Yamaguchi, M. A synthesis of (±)brefeldin A. *Tetrahedron Lett.* 1981, 22, 2679-2681.
- (68) Scanned spectra are included in supporting information for paper III.
- (69) WO 200364456; **2003**, WO 200208198; **2003**, WO 2001074768; **2002**.
- (70) Mehta, A.; Jaouhari, R.; Benson, T. J.; Douglas, K. T. Improved efficiency and selectivity in peptide synthesis: use of triethylsilane as a carbocation scavenger in deprotection of *tert*-butyl esters and *tert*-butoxycarbonyl-protected sites. *Tetrahedron Lett.* **1992**, *33*, 5441-5444.
- (71) <u>The synthesis of chiral amino-2-vinylcyclopropanecarboxylic acid vinyl.</u> Beaulieu, P. L.; Gillard, J.; Bailey, M. D.; Boucher, C.; Duceppe, J-S.; Simoneau, B.; Wang, X-J.; Zhang, L.; Grozinger, K.; Houpis, I.; Farina, V.; Heimroth, H.; Krueger, T.; Schnaubelt, J. Synthesis of (*IR*,2S)-1-amino-2-vinylcyclopropanecarboxylic acid vinyl-ACCA) derivatives: key intermediates for the preparation of inhibitors of the hepatitis C virus NS3 protease. J. Org. Chem. 2005, 70, 5869-5879.
- (a) <u>The first documented synthesis of amino-2-vinylcyclopropanecarboxylic acid vinyl.</u> Park, K-H.; Olmstead, M. M.; Kurth, M. J. Diasteroselective solid-phase synthesis of novel hydantoin- and isoxazoline-containing heterocycles. J. Org. Chem. 1998, 63, 6579-6585.
- (73) The (1*R*,2*S*)-1-amino-2-vinyl-cyclopropanecarboxylic acid ethyl ester was bought from Syngene-Custom Research, Bangalore, 20th KM Hosur Road, Electronics City, Bangalore 560100, India.
- (74) Carpino, L. A.; Han, G. Y. The 9-fluroenylmethoxycarbonyl amino-protecting group. J. Org. Chem. 1972, 22, 3404-3409.
- (75) (a) Pearson, D. A.; Lister-James, J.; McBride, W. J.; Wilson, D. M.; Martel, L. J.; Civitello, E. R.; Dean, R. T. Thrombus imaging using technetium-99m-labeled high-potency GPIIb/IIIa receptor antagonists. Chemistry and initial biological studies. J. Med. Chem. 1996, 39, 1372-1382. (b) Thierry, J.; Yue, C.; Potier, P. 2-Phenyl isopropyl and t-butyl trichloroacetimidates: useful reagents for ester preparation of N-protected amino acids under neutral conditions. Tetrahedron Lett. 1998, 39, 1557-1560.
- (76) Freidinger, R. M.; Hinkle, J. S.; Perlow, D. S.; Arison, B. H. Synthesis of 9fluorenylmethyloxycarbonyl-protected *N*-alkyl amino acids by reduction of oxazolidinones. *J. Org. Chem.* **1983**, *48*, 77-81.
- (77) Stanton, M. G.; Gagné. A mild protocol for the conversion of simple esters to *tert*butyl esters. *J.Org.Chem.*, **1997**, *62*, 8240-8242.
- (78) Computer analysis done in DS viewer Pro Suite 2002 and Chemdraw 3D 2004.
- (79) Picture done in DS viewer Pro Suite 2002.
- (80) Williams, D. H.; Fleming I. Spectroscopic methods in organic chemistr. *The McGraw-Hill Companies*. 1995 (Fifth edition.), ISBN 0-07-709147-7.
- (81) (a) Llinàs-Brunet, M.; Bailey, M. D.; Cameron, D.; Faucher, A-M.; Ghiro, E.; Goudreau, N.; Halmos, T.; Poupart, M-A.; Rancourt, J.; Tsantrizos, Y. S.; Wernic, D. M.; Simoneau, B.; Preparation of hepatitis C inhibitory tripeptides. WO 00/09543, 24 February 2000. (b) Tsantrizos, Y. S.; Cameron, D. R.; Faucher, A-M.; Ghiro, E.; Goudreau, N.; Halmos, T.; Llinàs-Brunet, M. Preparation of macrocyclic peptides active against the hepatitis C virus. WO 00/59929, 12 October 2000.
- (82) (a) Goudreau, N.; Dale, R.; Bonneau, C. R.; Plouffe, V. G. C.; Poirier, M.; Lamarre, D.; Llinàs-Brunet, M. NMR structural characterization of peptide inhibitors bound to the hepatitis C virus NS3 protease: design of a new P2 substituent. *J. Med. Chem.* 2004, 47, 123-132. (b) Orvieto, F.; Koch, U.; Matassa, V. G.; Muraglia, E. Novel, potent phenethylamide inhibitors of the hepatitis C virus (HCV) NS3 protease: probing the role of P2 aryloxyprolines with hybrid structures. *Bioorg. Med. Chem. Lett.* 2003, *13*, 2745-2748.
- (83) Barbato, G.; Cicero, D. O.; Cordier, F.; Narjes, F.; Gerlach, B.; Sambucini, S.; Grzesiek, S.; Matassa, V. G.; De Franscesco, R.; Bazzo, R. Inhibitor binding induces active site stabilization of the HCV NS3 protein serine protease domain. *Embo. J.* 2000, *6*, 1195-1206.
- (a) Yao, N.; Reichert, P.; Taremi, S. S.; Prosise, W. W.; Weber, P. C. Molecular views of viral polyprotein processing revealed by the crystal structure of the hepatitis C virus bifunctional protease-helicase. *Structure* 1999, 7, 1353-1363. (b) Johansson, A.; Hubatsch, I.; Åkerblom, E.; Lindeberg, G.; Winiwarter, S.; Danielsson, U. H.; Hallberg, A. Inhibition of hepatitis C virus NS3 protease activity by product-based peptides is dependent on helicase domain. *Bioorg. Med. Chem. Lett.* 2001, *11*, 203-206. (c) Johansson, A.; Poliakov, A.; Åkerblom, E.; Lindeberg, G.; Winiwarter, S.; Samuelsson, B.; Danielsson, U. H.; Hallberg, A. Tetrapeptides as potent protease inhibitors of hepatitis C virus full-length NS3 (protease-helicase/NTPase). *Bioorg. Med. Chem.* 2002, *12*, 3915-3922.
- (85) (a) <u>Synthesis of inhibitor 36</u>: Llinàs-Brunet, M.; Bailey, M.; Fazal, G.; Goulet, S.; Halmos, T.; Laplante, S.; Maurice, R.; Poirier, M.; Poupart, M-A.; Thibeault, D.; Wernic, D.; Lamarre, D. Peptide-based inhibitors of the hepatitis C virus serine protease. *Bioorg. Med. Chem. Lett.* 1998, *8*, 1713-1718. (b) <u>Synthesis of inhibitor 37</u>: LaPlante, S. R.; Cameron, D. R.; Aubry, N.; Lefebvre, S.; Kukolj, G.; Maurice, R.; Thibeault, D.; Lamarre, D.; Llinàs-Brunet, M. Solution structure of substrate-based ligands when bound to hepatitis C virus NS3 protease domain. *J. Biol. Chem.* 1999, *274*, 18618-18624. (c) <u>Synthesis of inhibitor 38</u>: Llinàs-Brunet, M.; Bailey, M.; Fazal, G.; Ghiro, E.; Gorys, V.; Goulet, S.; Halmos, T.; Maurice, R.; Poirier, M.; Poupart, M.

A.; Rancourt, J.; Thibeault, D.; Wernic, D.; Lamarre, D. Highly potent and selective peptide-based inhibitors of the hepatitis C virus serine protease: towards smaller inhibitors. Bioorg. Med. Chem. Lett. 2000, 10, 2267-2270. (d) Synthesis of inhibitor **39**:Beaulieu, P. L.; Llinàs-Brunet, M.; Therapies for hepatitis C infection: targeting the non-structural proteins of HCV. Curr. Med. Chem. 2002, 1, 163-176. (e) Synthesis of inhibitor 40: Llinàs-Brunet, M.; Bailey, M. D.; Bolger, G.; Brochu, C.; Faucher, A-M.; Ferland, J. M.; Garneau, M.; Ghiro, E.; Gorys, V.; Grand-Maître, C.; Halmos, T.; Lapevre-Paquette, N.; Liard, F.; Poirier, M.; Rhéaume, M.; Tsantrizos, Y. S.; Lamarre, D. Structure-activity study on a novel series of macrocyclic inhibitors of the hepatitis C virus NS3 protease leading to the discovery of BILN 2061. J. Med. Chem. 2004, 47, 1605-1608. (f) Synthesis of inhibitor 41: Llinàs-Brunet, M.; Bailey, M. D.; Ghiro, E.; Gorys, V.; Halmos, T.; Poirier, M.; Rancourt, J.; Goudreau, N. A systematic approach to the optimization of substrate-based inhibitors of the hepatitis C virus NS3 protease: discovery of potent and specific tripeptide inhibitors. J. Med. Chem. 2004, 47, 6584-6594. (g) Synthesis of inhibitor 42: Tsantrizos, Y. S.; Bolger, G.; Bonneau, P.; Cameron, D. R.; Goudreau, N.; Kukolj, G.; LaPlante, S. R.; Llinàs-Brunet, M.; Nar, H.; Lamarre, D. Macrocyclic inhibitors of the NS3 protease as potential therapeutic agents of hepatitis C virus infection. Angew. Chem. Int. Ed. Engl. 2003, 42, 1356-1360. (h) Synthesis of inhibitor 43: Lamarre, D.; Anderson, P. C.; Bailey, M.; Beaulieu, P.; Bolger, G.; Bonneau, P.; Bös, M.; Cameron, D. C.; Cartier, M.; Cordingley, M. G.; Faucher, A-M.; Goudreau, N.; Kawai, S. H.; Kukolj, G.; Lagacé, L.; LaPlante, S. R.; Narjes, H.; Poupart, M-A.; Rancourt, J.; Sentjens, R. E.; St George, R.; Simoneau, B.; Steinmann, G.; Thibeault, D.; Tsantrizos, Y. S.; Weldon, S. M.; Yong, C-L.; Llinàs-Brunet, M. An NS3 protease inhibitor with antiviral effects in humans infected with hepatitis C virus. Nature 2003, 426, 186-189.

(86) Rancourt, J.; Cameron, D. R.; Gorys, V.; Lamarre, D.; Poirier, M.; Thibeault, D.; Llinàs-Brunet, M. Peptide-based inhibitors of the hepatitis C virus NS3 protease: structure-activity relationship at the C-terminal position. J. Med. Chem. 2004, 47, 2511-2522.