Production and glycan binding characterization of human properdin and structural elucidation of c-Jun N-terminal kinase 3 inhibitors

Dissertation

der Mathematisch-Naturwissenschaftlichen Fakultät

der Eberhard Karls Universität Tübingen

zur Erlangung des Grades eines

Doktors der Naturwissenschaften

(Dr. rer. nat.)

vorgelegt von

Joana Tavares Macedo

aus Martingança, Portugal

Tübingen

2019

Gedruckt mit Genehmigung der Mathematisch-Naturwissenschaftlichen Fakultät der Eberhard Karls Universität Tübingen.

Tag der mündlichen Qualifikation: 29.07.2019
Dekan: Prof. Dr. Wolfgang Rosenstiel
1. Berichterstatter: Dr. Bärbel Blaum
2. Berichterstatter: Prof. Dr. Thilo Stehle

To the ones that have always been there:
a minha mãe, o meu pai e o meu irmão

To the one that will always be there:
my life partner, Tobi

'*Always look on the bright side of life*'
by Monty Python

Acknowledgements

Christina, Nils, Manu, Felix and Michael Buch, thank you for all the help inside and outside of the lab. Thank you for all the conversations, big laughs and little tears. I am so happy I have met you and I can call you friends. You made my life much better in the last years and I hope our friendship will continue, no matter where we are.

Francesco, you started as my flatmate, became my collaboration partner and now you are my friend. What a great path! Thank you for the delicious Italian food, the chance to contribute to your work and all the good moments. Munich is not so far away and I still need to visit Perugia, so we will definitely spend some more time together.

Queridas Aida, Paula and Joana, I cannot imagine how I would have managed the last years without friends like you. Aida, thank you for all your energy, motivation and conversations. Thank you for the beautiful speech last year, for showing me your home and for always being there for me. Paula, the beginning was not so easy, but I am glad I got the chance to get to know you better. Thank you for sharing my concerns but also for contributing to my hopes and dreams. You are such a funny chica! Joana, obrigada por seres o "meu Portugal" na Alemanha. Para além da sorte de ter alguém com quem falar português também tive a sorte de ganhar uma amiga para a vida.

Daniel, Rodrigo, Michael and Diego thank you for all the good and funny moments together. I am sure we will meet more times in the future.

DSP PD team and new colleagues from Celonic, thank you for your warm welcome and the easy integration into the team. Although the last months have been challenging, I am happy with my decision and I am really thankful for all your motivating words.

Dear Karin, Klaus and Florian, thank you for receiving me in your family and taking care of me just like my Portuguese family would do. Thank you for all the motivating words and for your understanding. I promise I will have more time for you in the near future.

Mãe, pai e mano, obrigada pelo vosso apoio incondicional. Obrigada por estarem sempre desse lado e acreditarem sempre em mim. Tenho tanta sorte em ter uma família como vocês. Estamos longe fisicamente, mas bem pertinho no coração!

My love Tobi, you deserve the biggest thank you of all. Thank you so much for believing in me and for never letting me down. Your constant support was essential from the beginning until now and none of this would have been possible without you. Amo-te!

Table of Contents

Abstract

Properdin, the only known positive regulator of the alternative complement pathway (AP), has been a focus of controversy since its discovery by Louis Pillemer in 1954 [1], [2]. In the last years evidence accumulated that this glycoprotein acts as a pattern recognition molecule and initiates the AP upon recognition of specific glycan markers, apart from its established role in the stabilization of AP convertases [3]. In this work, human full-length properdin was produced in mammalian cells and the purified protein was found to be active and to form the previously reported dimeric, trimeric and tetrameric cyclic structures. STD-NMR experiments performed with a set of glycans suggested that negative charge is required for binding to properdin and that glycosaminoglycans (GAGs) are potential pathogen-associated molecular patterns (PAMPs) for properdin and may mediate direct AP activation by properdin. The results also confirmed reports that both positive and negative regulators of the AP, namely properdin and complement Factor H, bind to different epitopes on identical glycans.

The structural complexity that contributes to the diversity of GAGs and glycans represents a significant challenge for their isolation for functional and structural studies. A common approach involves the use of GAG-depolymerizing enzymes, such as heparinase I from *Pedobacter heparinus*, followed by fractionation of the obtained oligosaccharides [4], [5]. However, since the most biological relevant oligosaccharides have a higher degree of polymerization than those usually obtained from heparinase, a structure-based engineering approach to rationally design this enzyme could alter the product distribution. Although recombinant heparinase I could be produced in *E. coli* and expression constructs with varying N- and C-terminal sequences were tested, an intrinsic low thermal and conformational stability was observed and no protein crystals were obtained for structure determination by X-ray crystallography and subsequent enzyme engineering.

Over the past decades, inhibition of the c-Jun N-terminal kinase 3 (JNK3), a mitogen-activated protein kinase (MAPK) involved in the regulation of cellular responses to extracellular stimuli has become a promising strategy for treatment of neurodegenerative disorders such as Alzheimer's and Parkinson's diseases. However, up to date, no inhibitors targeting JNK3 have been approved by the FDA [6], [7]. Altering the substitution pattern of a pyridinylimidazole scaffold from a dual p38α/JNK3 MAPK inhibitor proved to be effective in shifting the selectivity towards JNK3 [8]. A similar binding mode of the two most potent inhibitors with an IC_{50} value <1 µM in the ATP binding pocket was confirmed by X-ray crystallography. While selectivity was achieved by addressing the hydrophobic region I of JNK3 with a small methyl group, addition of a *S*-methyl group contributed to the stability of the G-rich loop of JNK3, thus increasing the inhibitory potency. Future strategies to increase the inhibitory potency while preserving selectivity were devised from the determined JNK3 crystal structures.

Zusammenfassung

Properdin, der einzige bekannte positive Regulator des alternativen Wegs des Komplementsystems (abgekürzt mit AP für alternative complement pathway), ist seit seiner Entdeckung durch Louis Pillemer im Jahr 1954 Gegenstand von Kontroversen [1] [2]. In den letzten Jahren wurden verschiedene Hinweise dafür beschrieben, dass dieses Glykoprotein neben seiner bekannten Rolle bei der Stabilisierung von AP-Konvertasen auch als Mustererkennungsmolekül fungieren und den AP direkt initiieren kann [3]. Properdin wurde in der vorliegenden Arbeit rekombinant in Säugetierzellen hergestellt und es wurde gezeigt, dass das gereinigte Protein aktiv ist und die zuvor beschriebenen zyklischen Strukturen (Dimere, Trimere und Tetramere) bildet. STD-NMR-Experimente, die mit einer Reihe von Glykanen durchgeführt wurden, legen nahe, dass eine negative Ladung für die Bindung an Properdin erforderlich ist und dass Glykosaminoglykane (GAGs) als potenzielle Pathogen-assoziierte molekulare Muster (abgekürzt PAMPs für pathogen-associated molecular patterns) agieren und zur AP-Aktivierung durch Properdin führen könnten. Die Ergebnisse bestätigten auch die bestehende Vermutung, dass sowohl positive als auch negative Regulatoren des AP (Properdin und Komplementfaktor H) an identische Glykane binden können, die Bindung jedoch über die Erkennung verschiedener Epitope erfolgt.

Die strukturelle Komplexität, die zur Vielfalt der GAGs und anderer Glykane beiträgt, stellt gleichzeitig eine große Herausforderung für deren Isolierung für funktionelle und strukturelle Studien dar. Ein verbreiteter Ansatz ist die Verwendung von GAG-depolymerisierenden Enzymen wie Heparinase I aus Pedobacter heparinus. Dabei werden die entstandenen Oligosaccharide mit unterschiedlichen Molekulargewichten mittels anschließender Fraktionierung erhalten [4], [5]. Die Entwicklung eines strukturbasierten Protein-Engineerings für dieses Enzyms könnte die Produktverteilung so verändern, dass die biologisch relevantesten Oligosaccharide mit einem höheren Polymerisationsgrad isoliert werden können. Heparinase I konnte erfolgreich in E. coli hergestellt werden und verschiedene Expressionskonstrukte mit variierenden N- und C-terminalen Sequenzen wurden getestet. Dabei wurde jedoch in allen Konstrukten eine intrinsische niedrige thermische und konformationelle Stabilität beobachtet und es wurden keine Proteinkristalle zur Strukturbestimmung mittels Röntgenstrukuranalyse erhalten.

In den letzten Jahrzehnten wurde die Inhibition der c-Jun-N-terminalen Kinase 3 (JNK3), einer mitogenaktivierten Proteinkinase (MAPK), die an der Regulation der Zellreaktionen auf extrazelluläre Stimuli beteiligt ist, zu einer vielversprechenden Strategie für die Behandlung von neurodegenerativen Erkrankungen wie Alzheimer und Parkinson beschrieben. Bisher wurden jedoch keine Inhibitoren gegen JNK3 von der FDA zugelassen [6], [7]. Durch die Veränderung des Substitutionsmusters bzw. der Seitenketten eines Pyridinylimidazolgerüsts

eines dualen p38α/JNK3-MAPK-Inhibitor konnte die Selektivität in Richtung JNK3 verschoben werden [8]. Mittels Röntgenstrukturanalyse konnte für die beiden potentesten Inhibitoren mit IC_{50}-Werten <1 µM sehr ähnliche Bindungsmodi in der ATP-Bindungstasche bestätigt werden. Dabei wurde die Selektivität durch die Verwendung einer Methylgruppe, die mit der hydrophoben Region I von JNK3 interagiert, erhöht. Die Verwendung einer S-Methylgruppe erhöhte die inhibitorische Wirksamkeit durch die Stabilisierung einer Glycin-reichen Loopregion von JNK3. Anhand der bestimmten JNK3-Kristallstrukturen wurden weitere Strategien zur Erhöhung der inhibitorischen Wirksamkeit unter Beibehaltung der Selektivität entwickelt.

Abbreviations

AMP-PCP	β,γ-methyleneadenosine 5′-triphosphate disodium salt
AMP-PNP	Adenylyl-imidodiphosphate
AP	Alternative pathway of the complement system
ATP	Adenosine triphosphate
BSA	Bovine serum albumin
B. thetaiotaomicron	Bacteroides thetaiotaomicron
CCP	Complement control protein
CD	Circular dichroism
CNS	Central nervous system
CP	Classical pathway of the complement system
DSF	Differential scanning fluorimetry
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
DTT	Dithiothreitol
FACS	Fluorescence-activated cell sorting
FB	Factor B
FD	Factor D
FDA	U.S. Food and Drug Administration
FH	Factor H
FI	Factor I
E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EGTA	Ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
EM	Electron microscopy
GAG	Glycosaminoglycan
GST tag	Glutathione *S*-transferase tag
HEK	Human embryonic kidney
HEPES	(4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid)
HS	Heparan sulfate
IC_{50}	Half maximal inhibitory concentration
IEX	Ion exchange chromatography
IPTG	Isopropyl-β-D-thiogalactopyranoside
JNK	c-Jun N-terminal kinase
LB	Luria broth
LMWHs	Low molecular weight heparins

LP	Lectin pathway of the complement system
MALDI-MS	Matrix-assisted laser desorption/ionization mass spectrometry
MAPK	Mitogen-activated protein kinase
MRE	Mean residue weight ellipticity
MS	Mass spectrometry
MW	Molecular weight
MWCO	Molecular weight cutoff
NGLs	Neoglycolipids
NHS	Normal human serum
NMR	Nuclear magnetic resonance
OD_{600}	Optical density at 600 nm
PAMP	Pathogen-associated molecular pattern
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PDB	Protein data bank
PEG	Polyethylene glycol
PEI	Polyethylenimine
P. heparinus	*Pedobacter heparinus*
pI	Isoelectric point
PL	Polysaccharide lyase
PTM	Post-translational modification
PVDF	Polyvinylidene fluoride
RBCs	Red blood cells/ Sheep erythrocytes
RMSD	Root mean square deviation
RT	Room temperature
SEC	Size exclusion chromatography
SDS-PAGE	Sodium dodecyl sulfate-polyacrylamide gel electrophoresis
SPR	Surface plasmon resonance
STD-NMR	Saturation transfer difference nuclear magnetic resonance
TBS-T	Tris-buffered saline solution with 0.5% (v/v) Tween-20
TCEP	Tris (2-carboxyethyl) phosphine
TEM	Transmission electron microscopy
TEV	Tobacco Etch Virus
TF antigen	Thomsen-Friedenreich antigen
T_m	Melting or unfolding temperature
TSR	Thrombospondin
UV	Ultraviolet
VWA	Von Willebrand Factor A-type

1 Introduction

1.1 The complement system

The human complement system, an essential component of the innate immune response, serves as a first line of defense against pathogenic intruders and altered host cells [9]. Studied for the first time in the late 19th century, complement was initially identified as a heat-labile lytic fraction of human plasma that 'complemented' antibodies in their ability to recognize and fight bacteria [10]. To prevent inappropriate attack and thus avoid host tissue and organ damage, which can lead to autoimmune and chronic inflammatory diseases, the complement system must be tightly regulated. Complement comprises more than 30 soluble blood proteins and cell-surface bound proteins that act in plasma, tissues or within cells and serve as substrates, enzymes or modulators [11], [12]. Since the largest part of complement proteins are serum proteins and cellular receptors, glycosylation is an important factor in the complement system. These glycoproteins are synthesized in the liver, in macrophages and in lymphoid tissue and thus consist of a mixture of glycosylation variants (glycoforms) [13]. Proteins at the heart of the complement system, the so-called convertases, circulate in the blood in inactive forms (so-called zymogens). Upon complement activation, large conformational changes occur and successive enzymatic reactions serve to opsonize pathogens and induce inflammatory responses, which help immune cells to fight infections [14].

Depending on the context, the complement system can be activated through three distinct pathways, known as the classical (CP), lectin (LP) and alternative (AP) pathways (Figure 1.1). Activation of the CP and LP occurs upon binding of pattern recognition molecules to molecular entities that are typical of pathogens [15]. The CP is triggered when complement protein C1 binds to foreign targets that have IgG or IgM antibody molecules attached via the hexameric pattern recognition molecule C1q [16]. The LP pathway is initiated in response to recognition of mannose residues on pathogen surfaces by mannose-binding lectins and ficolins [17]. Conversely, the AP is thought to be permanently active at low-level (so-called 'AP tickover') and, therefore, requires continuous regulation. This pathway can be amplified through a positive feedback loop in the presence of pathogens [18].

Zymogens are inactive precursors of complement proteases that require proteolytic cleavage to become active enzymes. Local activation of zymogens, such as C3 and C5, leads to their cleavage into smaller fragments, which subsequently associate with other proteins and

cleave and activate further zymogens in the complement pathway. The activation of a small number of complement proteins is therefore highly amplified by a cascade of enzymatic reactions, resulting in the generation of a large complement response [19]. In this way, the complement system can be ubiquitously present in an inactive form but rapidly become activated at sites of infection [14].

1.1.1 The central role of complement protein C3

All three pathways of the complement system culminate at the cleavage of the central zymogen C3. After activation of the complement system by antibody complexes, mannose-containing glycans or by spontaneous and induced hydrolysis of a labile thioester bond present in C3 on surfaces, C3 is cleaved into its active fragments [20]. Proteolytic activation of C3 by an enzymatic complex called C3 convertase, either the CP and LP C3 convertase C4bC2a or the AP C3 convertase C3bBb, releases the small fragment C3a (~9 kDa) and the larger fragment C3b (~180 kDa) [21]. C3a is an anaphylatoxin that triggers a range of chemotactic and pro-inflammatory responses by binding to the C3a receptor on various effector cells, including recruitment of phagocytes to the site of infection, activation of leukocytes and an increase in vascular permeability [22]. C3b, on the other hand, is an opsonin that, upon cleavage of C3, undergoes a large structural change leading to exposure of a highly reactive but short-lived thioester. The exposed thioester moiety can tag cells or proteins by covalently binding to carbohydrates as well as hydroxyl and amino groups found on cell surfaces in close proximity to its site of generation [23], [24]. Once C3b has covalently attached to a pathogenic surface, terminal complement pathway activation occurs in the same manner, regardless of the pathway that initiated complement activity [25].

1.1.2 Direct activation of the alternative complement pathway

Unlike the CP and LP that rely on specific recognition molecules, the AP activates on any surface that is not actively protected by complement regulatory proteins [24], [26]. This system is initiated in the fluid-phase by the tickover as well as by contact activation on biological surfaces (e.g. platelets) and artificial surfaces (e.g. gas bubbles, biomaterial surfaces) and converts C3 into a bioactive form called $C3(H_2O)$ [27]–[29]. The thioester domain of $C3(H_2O)$ or C3b produced by the CP or LP exposes a binding site for another member of the AP known as Factor B (FB), which binds in a Mg^{2+}-dependent manner and yields the pro-convertases $C3(H_2O)B$ or C3bB. This interaction occurs via the Von Willebrand Factor A-type (VWA) domain and three complement control protein (CCP) domains of FB. Upon activation, the serine protease Factor D (FD) binds to and cleaves the Factor B part of $C3(H_2O)B$ or C3bB which releases the N-terminal fragment Ba and generates the AP C3 convertase complex $C3(H_2O)Bb$ or C3bBb [30], [31]. During normal

physiological conditions, these C3 convertases constantly activate the complement system by interacting with additional C3 molecules and generating additional small amounts of C3b that expose the internal thioester bond and produce even more C3bBb [11]. However, C3 convertases are not very stable on pathogenic surfaces and have a short half-life of 90 seconds [32]. In this amplification loop, more C3b molecules are generated and deposited on cell surfaces, which trigger opsonization [33]. C3b can be further cleaved by Factor I (FI) into iC3b, C3dg and C3d and these fragments can be recognized by complement receptors and promote phagocytosis [34]. During the labeling of surfaces, C3b molecules bind to or near the C3 convertases that produced these C3b and generate a new enzymatic complex, called C5 convertase (C3bBbC3b). C5 convertases have a different substrate specificity and cleave the zymogen C5 into its active fragments C5a, a potent chemoattractant that acts as an inflammatory mediator by binding to its C5aR receptor, and the opsonin C5b [35], [36]. C5b can bind to C6, followed by a sequential binding of C7, C8 and C9 to form a multiprotein pore complex called membrane attack complex (MAC). This part of the complement system is termed 'terminal pathway' since all activation pathways converge at the level of C5.

Properdin, also called Factor P, can bind to and stabilize the AP convertases thereby increasing their half-life by 5- to 10-fold and leading to enhanced complement activity [37]. In fact, evidence for the AP as an individual pathway emerged for the first time in the early 1950's when Pillemer reported a novel protein termed properdin (derived from the Latin verb perdere, 'to destroy') and its apparent capacity to activate the complement system on several targets in the absence of antibodies, i.e. without involvement of the CP [1], [38].
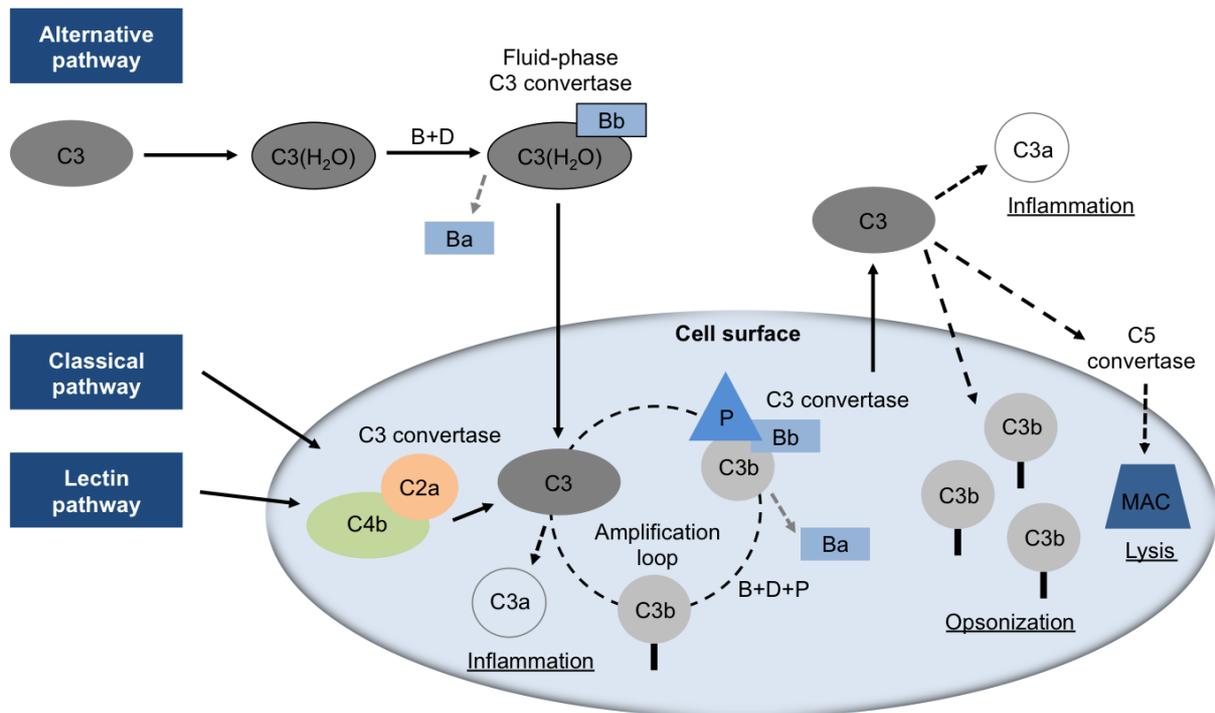
Figure 1.1 Overview of the complement system and the role of the alternative pathway. Complement activation is initiated by the CP, LP and AP. Pattern recognition molecules such as antibodies and mannose-binding lectins generate the C3 convertase C4bC2a. Unlike the CP and LP, the AP is activated by spontaneous hydrolysis of C3 originating the fluid-phase C3 convertase C3(H$_2$O)Bb. C3 convertases cleave plasma C3 molecules to generate C3a and C3b. C3b molecules can then covalently attach to cell surfaces and, together with FB and FD, lead to the formation of C3bBb convertases via an amplification loop. Properdin binds to C3bBb, extends its half-life and thus promotes efficient C3b deposition on cell surfaces. All pathways culminate in the formation of C3 and C5 convertases, which in turn generate the major effector molecules of the complement system: the potent pro-inflammatory anaphylatoxins (C3a and C5a) and the opsonins (C3b and C5b), which trigger phagocytosis and the membrane attack complex that can directly lyse targeted surfaces. All together these mechanisms contribute to pathogen and altered host cell elimination. Figure adapted from [11].

1.1.3 Properdin as an initiator of the alternative pathway

Glycans play essential roles in mammalian biology, for instance in cellular communication and in the modulation of immune and inflammatory responses. Although complement system regulation is achieved mainly by sets of protein-protein interactions, as outline above, protein-glycan interactions equally contribute to the maintenance of complement homeostasis [39], [40]. Among the complement regulators, interactions with glycosaminoglycans (GAGs) have been reported for properdin, the only known positive regulator of complement that may also act as a pattern-recognition molecule [41]. Despite all efforts that have been made to date, the understanding of the biological role of properdin as well as its mechanism of action is still fragmentary due to its oligomeric heterogeneity and the different sources of properdin used in biochemical studies [42]. The atomic three-dimensional structure of properdin has not yet been solved and controversies continue

to surround its role as a possible initiator of alternative complement pathway since its discovery by Pillemer and its coworkers in 1954 [1], [2].

1.1.4 Regulation of the complement system

Although the AP effectively activates and amplifies on most pathogen surfaces, host cells and tissues need to be protected from inadvertent complement activation. Soluble or membrane-bound complement regulatory proteins can provide protection against complement activation by down-regulating the central proteolytic activity of the amplification and opsonization steps [43], [44]. The membrane-bound regulators, such as complement receptor 1 (CR1/CD35), membrane-cofactor protein (MCP/CD46/CD59), MAC-inhibitory protein (CD59) and decay-accelerating factor (DAF/CD55) are common proteins in the regulation of all three pathways. The process to differentiate between host cells from targets by the AP requires participation of Factor H (FH), C3b and potentially properdin via its role as a pattern-recognition molecule [3], [45]. Defective complement regulation due to mutations in the complement genes or the presence of autoantibodies against complement regulators has been associated with several diseases, such as atypical hemolytic uremic syndrome, membranoproliferative glomerulonephrithis in the kidney and aged-related macular degeneration in the eye [46], [47]. Besides, certain bacteria and viruses have found ways to utilize these host protection mechanisms and escape complement response [15]. For example, orthopox and herpes viruses express soluble proteins that structurally and functionally mimic host regulators [48].

FH is an elongated 150 kDa plasma glycoprotein and the main soluble regulator of the AP. This protein composed of 20 CCP modules can recognize specific markers on host cells and control complement activation on self-surfaces in addition to its regulatory activity in the fluid-phase [49], [50]. FH engages most effectively with C3b molecules when these are attached to self-surfaces carrying specific polyanionic self-markers such as GAGs and sialic acid [40], [51]. FH regulates complement by accelerating convertase decay activity via irreversible dissociation of Bb from C3 and C5 convertase complexes, thus inhibiting the complement amplification loop. In addition, FH can also act as a cofactor for FI-mediated cleavage of C3b and $C3(H_2O)$ generating the inactive fragments iC3b and $iC3b(H_2O)$, which can no longer form AP convertases [52], [53].

Whereas human FH functions to down-regulate AP activation, properdin up-regulates this pathway by stabilizing C3 convertases, which generate additional C3b molecules and lead to opsonization and formation of the lytic/terminal pathway [54]. These two key regulatory proteins often have opposing roles and can also directly compete with each other. It has

been reported that properdin could possibly limit FH cofactor activity by sterically interfering with the interaction between FH and FI [55]. Additionally, structural data obtained by electron microscopy (EM) have also indicated that properdin could affect FH decay accelerating activity [56].

1.1.5 Functional and structural features of properdin

Properdin is a 53 kDa plasma glycoprotein synthetized as a single-chain molecule of 469 amino acids, including a 27-amino acid leader sequence and 442 amino acids in the mature protein [57]. In contrast to other complement proteins, which are mainly produced in the liver, properdin is synthesized and secreted by various cell types, including neutrophils, monocytes, primary T cells and endothelial cells resulting in plasma levels ranging from 4-25 µg/mL [58]. With an isoelectric point (pI) greater than 9.5, this highly positively charged protein at neutral pH is composed of several identical subunits that bind to each other in a head-to-tail manner under physiological conditions to form cyclic dimers (P_2), trimers (P_3), and tetramers (P_4) in a 26:54:20 ratio [37], [59]. Each properdin monomer adopts a flexible rod-like structure and consists of seven non-identical tandem repeats of about 60 amino acids named thrombospondin type I repeats (TSR) (Figure 1.2A) [60]. The TSR motif was originally found in the adhesive glycoprotein thrombospondin but is also known to be present in the C6, C7, C8α, C8β and C9 terminal components of complement [61]. TSR domains of properdin have been numbered from the truncated N-terminal domain TSR0 followed by TSR1-6 [62]. Based on a sequence alignment between thrombospondin and properdin together with structural studies, highly conserved residues such as six cysteines, three tryptophans and two arginines have been identified within the core of each TSR domain, which comprises three antiparallel β-strands held together by three disulfide bonds (Figure 1.2B) [41], [63]. Although the three-dimensional structure of properdin has not been determined to date, crystallization conditions have been recently reported and structures of other TSRs, such as TSR2 and TSR3 of human thrombospondin are available [63], [64]. The properdin monomer contains a single *N*-glycosylation site and 14 *C*-mannosylation sites at tryptophan residues that are part of a WxxW motif [65]. The *N*-linked oligosaccharide is attached to Asn 401 in TSR6 [66]. Unlike *N*-glycosylation, which is a rather common post-translational modification (PTM), *C*-mannosylation is an unusual PTM of tryptophan residues and its biological function is not yet clear. Recent studies highlighted the extensive heterogeneity of *C*-mannosylation in the different TSR domains of properdin and indicated a possible role for this modification in protein folding and stability [67], [68].

A

B
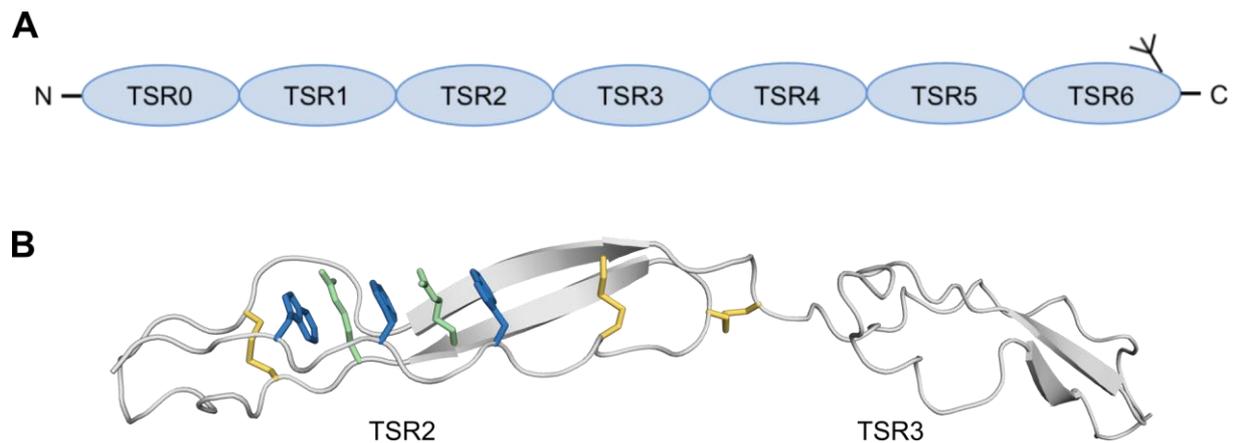
Figure 1.2 Properdin monomer and three-dimensional structure of a TSR domain. A) Schematic representation of the properdin monomer and arrangement of TSR domains. The properdin monomer is composed of seven thrombospondin repeats labeled TSR0-6. The single *N*-linked glycosylation site located in TSR6 is indicated. B) Crystal structure of the homolog TSR2 and TSR3 domains from human thrombospondin (PBD 1LSL) [63]. The protein backbone is depicted in gray and side chains of the proposed key tryptophan and arginine residues that are part of the WxxW motif are shown in blue and green for TSR2, respectively. Disulfide bonds from TSR2 are shown in yellow. Figure adapted from [3], [102].

Different approaches have been used to understand the biological role of properdin domains and oligomerization [3]. Higgins *et al.* investigated properdin deletion mutants, lacking single TSR domains and reported that some of these TSRs mediate properdin function and oligomerization. Whereas deletion of TSR3 did not affect properdin function, properdin lacking TSR4 was able to bind C3b but failed to stabilize C3bBb. Instead, properdin lacking TSR6 was unable to form oligomers, while properdin lacking either TSR5 or TSR6 was unable to bind C3b. Proteolytic nicking within TSR5 also interfered with C3b binding and C3bBb stabilization [66]. Key roles implicated for domains 5 and 6 were further supported by subsequent studies in which antibodies raised against human TSR5 and mouse TSR5 and TSR6 domains effectively inhibited properdin function *in vivo* and *in vitro*, respectively [69], [70]. According to structural studies performed by Pedersen *et al.* and Alcorlo *et al.*, TSR5 is needed for the interaction with C3b and this binding occurs on top of the C345C domain of C3b. Besides, properdin was also suggested to contact the VWA domain of Factor B and TSR4 may be responsible for this interaction [56], [71]. The structural studies by Alcorlo *et al.* as well as by Sun *et al.* produced models for properdin oligomerization (Figure 1.3). These studies suggested roles for the TSR0-1 and TSR5-6 domains in mediating contacts at the vertexes of properdin oligomers and also indicated variable combinations of the vertexes. These vertexes were assigned to be the structural unit of recognition and stabilization C3bBb convertases [56], [62]. In addition, previous studies reported the influence of properdin oligomers in C3 convertase stabilization and showed that the tetramer had a greater convertase stabilizing activity than both the trimer and the dimer

[59]. Properdin tetramers were also the most efficient at inducing complement-mediated platelet-granulocyte aggregate formation [72]. This increased activity of tetramers is likely due to an increased avidity of properdin for the C3 convertases caused by the presence of multiple convertase binding sites [56]. By characterizing a patient-derived monomeric properdin, which contained a single point mutation (E244K) in TSR3, Pedersen *et al.* demonstrated that properdin oligomerization is essential for *in vivo* function [71].

Properdin deficiency is a rare X-chromosomal linked disorder characterized by an increased susceptibility to infections with *Neisseria meningitidis* and a significantly higher risk of developing meningococcal infections compared to individuals without this deficiency [73], [74]. Three different types of properdin deficiencies have been described: type I which refers to complete absence of properdin, type II represented by low levels of properdin in plasma (1-10% of normal concentration), and type III with a normal concentration of dysfunctional properdin [75].
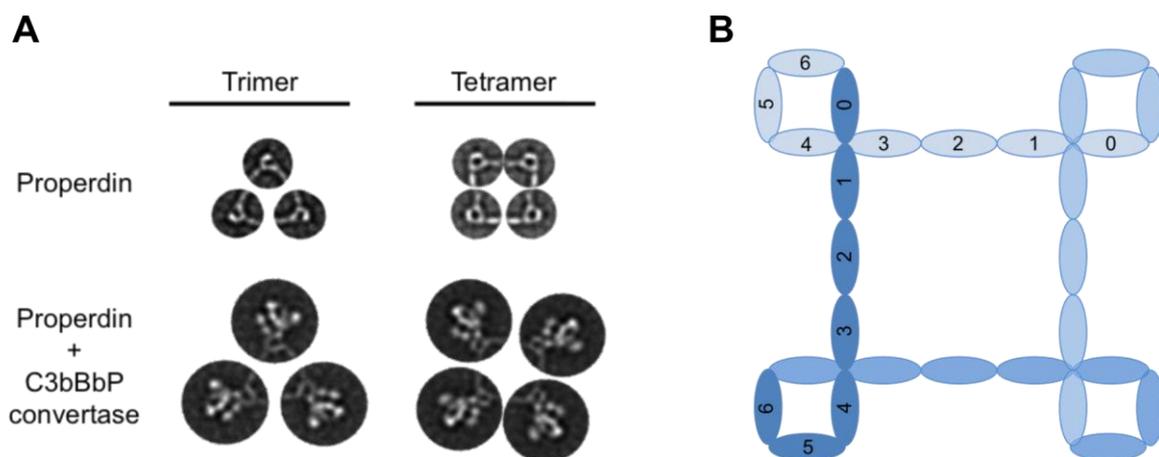
Figure 1.3 Model of properdin oligomerization and complement activation proposed by Alcorlo *et al.* [56]. A) 2D models of properdin and properdin-C3bBb convertase complexes. Models containing properdin trimers and tetramers were generated by combining the averages of single molecule images from negative-stain EM. B) Cartoon representation of properdin tetramer model. Molecular modeling suggested that the vertexes of properdin oligomers are composed by four TSR domains from two different monomers organized in a head-to-tail manner. Monomers are shown in different tones of blue and TSR domains are numbered from 0 to 6.

1.1.6 Properdin as an initiator of AP activity

When properdin was first discovered, it was thought to act as an initiator of AP activity [1]. However, this theory was later discredited and replaced by the now widely accepted role of properdin as a positive regulator of preexisting AP activity [37]. Within the past decade, several studies have provided new insights into the function of properdin and rediscovered its ability to act as a pattern recognition molecule and initiate AP activity [58]. Using surface

plasmon resonance (SPR), Hourcade *et al.* demonstrated that properdin has the ability to initiate the AP by providing a platform for *de novo* assembly of C3bBbP. Properdin immobilized on a biosensor surface and treated with purified complement proteins in the presence of Mg^{2+} associated with C3b or $C3(H_2O)$. The formed C3bP complex was able to bind FB and C3bBP was subsequently cleaved by FD, releasing Ba and generating surface-bound C3bBbP [76]. Whereas recognition of surfaces by C3b is relatively non-specific and thus requires a tight regulation, it is possible that properdin can selectively direct complement activation by recognizing and binding to specific target surfaces [77]. According to Spitzer *et al.*, once properdin has bound to a target surface, it initiates and propagates the complement response by recruiting C3b to the recognized surfaces and allowing *de novo* C3 convertase assembly as well as by providing C3 convertase stabilization (Figure 1.4) [78].
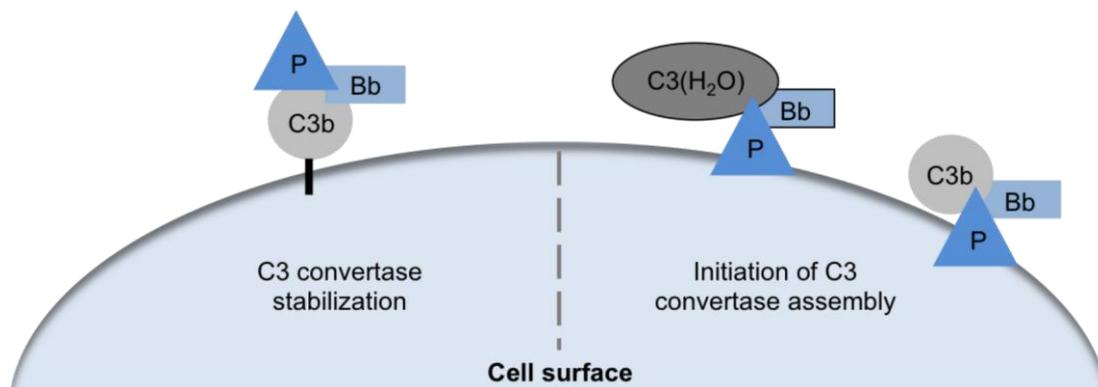
Figure 1.4 Proposed roles for properdin within the alternative complement pathway. Two models have been suggested for the initiation and assembly of AP C3 convertases on target surfaces. Mode highlighted on the left side: assembly of AP C3 convertase is initiated via C3b, generated by the non-specific fluid-phase activation of C3, which covalently binds to non-specific surfaces located near its site of generation. The C3 convertase is then stabilized by properdin and complement activation can be propagated via the amplification loop. Model highlighted on the right side: properdin-direct initiation of AP results from the selective binding of properdin to target surfaces, which recruits $C3(H_2O)$ or C3b and provides a platform for *de novo* C3bBb assembly. Figure adapted from [3], [78].

Several studies have supported the role of properdin as a pattern recognition molecule that can bind directly to microbial surfaces such as *Chlamydia pneumoniae*, late apoptotic and necrotic cells as well as biological substrates such as bacterial lipopolysaccharide (LPS), heparan sulfate (HP), and zymosan [78]–[83]. However, caution needs to be taken when interpreting results of properdin binding studies since large aggregates of properdin (P_n) are formed during purification of this protein or upon repeated freeze-thaw cycles. These non-physiological aggregates of properdin, described by Farries *et al.* as large amorphous aggregates, and also designated as "activated" properdin, have distinct properties from physiological oligomers given their highly positive charge [84]. First, while retaining the ability to stabilize AP convertases, these aggregates can also spontaneously induce AP activation

in solution leading to complement consumption [59]. Second, they may account for non-specific ionic interactions with target surfaces, such as *Neisseria meningitidis*, *Neisseria gonorrhoeae*, live Raji and Jurkat cells and non-activated platelets [85]–[87]. Separation of physiological forms of properdin (P_2, P_3 and P_4) from P_n forms was thus recommended using ion exchange or size exclusion chromatography [88].

With the exception of C3 convertase proteins, all cell surface molecules reported to directly interact with properdin to date are negatively charged [58]. Sulfated GAGs including heparin, heparan sulfate (HS), chondroitin sulfate (CS), dextran sulfate and fucoidan have been shown to bind to properdin [80], [89]–[91]. Besides, it has also been demonstrated that properdin can recognize apoptotic T cells via HS and CS and contribute to their clearance by enhancing complement-mediated opsonization and phagocytosis [80]. Furthermore, properdin was shown to bind to heparin and HS on tubular epithelial cells leading to complement activation [90]. Other ligands reported for properdin include surface DNA on late apoptotic and necrotic cells [81]. However, since several of these studies utilized purified, unfractionated preparations of properdin, it is possible that these samples contained aggregates, which may bind to surfaces that native properdin present in plasma would not bind [33]. The methodologies used to study the binding of properdin to target surfaces have also contributed to the controversy around this protein. Most of the experiments performed to analyze the role of properdin as a pattern recognition molecule did not allow discrimination between initial binding of properdin and binding of properdin secondary to C3b deposition under physiological conditions [92]. Previous work by Harboe *et al.,* where C3 cleavage was inhibited by peptide inhibitor compstatin Cp40, demonstrated that properdin binding to yeast zymosan and *Escherichia coli* (*E. coli*) is dependent on initial C3 deposition [93], [94]. Nevertheless, since the tickover mechanism is present under typical conditions *in vivo* and it is required to provide C3b molecules, the question of whether properdin can bind to target surfaces independently of C3 might not be of high relevance [95].

Further studies are thus necessary to determine specific interactions between properdin and surfaces and to identify new ligands for direct properdin binding to surfaces, such as glycans. The differences between the ability of purified physiological and non-physiological forms of properdin, as well as plasma- and neutrophil-derived properdin to bind surfaces should be carefully analyzed in order to determine if properdin can truly act as initiator of complement activation [58]. A better understanding of properdin function and regulation could potentially lead to the development of novel properdin inhibitors that block properdin-initiated activities without impairing its role in convertase stabilization [42].

1.2 Role of glycans and glycosaminoglycans

Glycans, one of the four molecular building blocks of life (nucleic acids, amino acids, lipids, and glycans), have a wide structural and functional diversity and participate in multiple biological processes in cells and tissues, including those that are fundamental for the modulation of the immune system [39], [96]. Initial recognition and interaction with cell surfaces is often mediated by glycans due to their structural diversity as well as their quantity and high avidity [97]. For example, the AP of complement is regulated on cell surfaces via glycans such as sialic acid or GAGs, which can dictate whether a positive or negative immune response occurs, including complement activation, as a result of the different sugar composition of a tissue's glycomatrix [40], [98]. Sialic acids, formed by a nine-carbon backbone, are found attached to the end of glycan chains on human glycoproteins and can influence complement activation through Factor H binding [99]. The predominant form of sialic acid in humans is *N*-acetylneuraminic acid (Neu5Ac), which can be enzymatically removed with neuraminidase and convert non-activators of complement into activators [100], [101]. In the case of properdin, several glycans, including GAGs such as heparin and HS, have been identified as potential pathogen-associated molecular patterns (PAMPs) [9], [102]. Zaferani *et al.* have, for example, shown that besides recognizing distinct HS epitopes on renal tubular epithelial cells, properdin and FH require different sulfation patterns for interaction [103]. These findings demonstrate therefore the high impact of GAGs in the mediation of immune responses by recruiting both regulators of complement through the recognition of different sulfation motifs [54].

1.2.1 GAGs – Heparin and heparan sulfate

GAGs are polyanionic, polydisperse, linear polysaccharides of varying molecular weights that reside within the extracellular matrix and on the surface of almost all mammalian cells [104], [105]. These molecules are usually attached to proteins via a serine residue, one xylose moiety and two galactose moieties, forming the so-called proteoglycans [106]. Based on the composition of each monosaccharide as well as the position and configuration of the glycosidic linkage formed between repeating disaccharide units, GAGs can be classified in four groups: heparin/HS, chondroitin sulfate/dermatan sulfate, keratan sulfate and hyaluronic acid (Figure 1.5) [4]. In contrast to all other GAGs, which undergo post-polymerization modifications to exert their biological functions, such as epimerization, deacetylation and sulfation, hyaluronic acid exists solely as protein-free and is non-sulfated [107], [108].
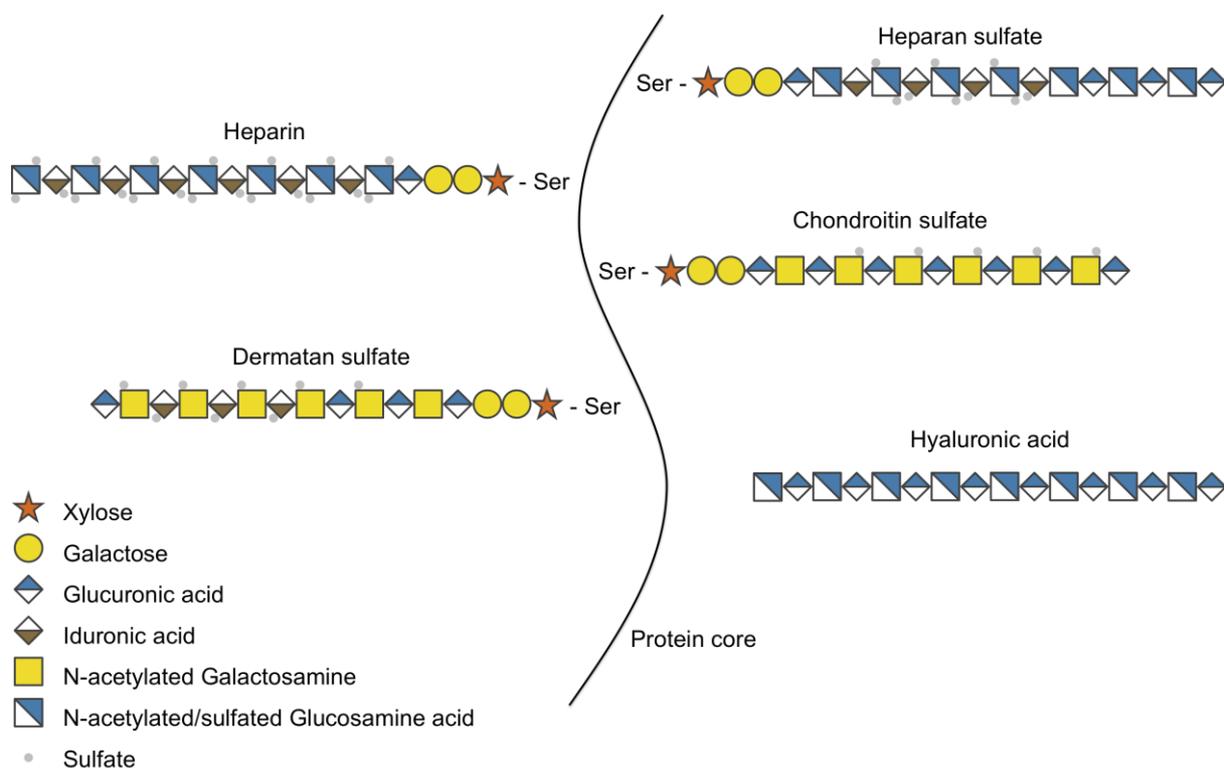
Figure 1.5 Schematic representation and composition of GAGs. Except hyaluronic acid, which exists as protein-free, GAGs such as heparin/HS, chondroitin sulfate/dermatan sulfate and keratan sulfate are attached to proteins via serine residues, followed by one xylose and two galactose moieties. GAGs composition varies greatly, depending on the type of GAG as well as on their sulfation and acetylation state. S and A labels in HS denote sulfated and non-sulfated domains, respectively. Figure adapted from [105], [109].

Over several years, heparin and HS, the structurally most complex members of the GAG family of polysaccharides, have been shown to interact with a large number of proteins and thereby regulate a wide range of biological processes, including cell growth and proliferation, inhibition of angiogenesis, inflammation and viral infectivity [110]. Although these two GAGs share similar biosynthetic routes, heparin is found primarily in secretory granules of mast cells and is structurally simpler than HS, which is expressed and secreted by a variety of mammalian cells and is predominantly located on cell surfaces [111].

Heparin and HS consist of the same disaccharide building blocks, an uronic acid residue, either α-L-iduronic acid (IdoA) or β-D-glucuronic acid (GlcA), 1,4-glycosidically linked to D-glucosamine (GlcN) [111], [112]. These building blocks can be subsequently modified by partial *N*-deacetylation/*N*-sulfation of GlcN units, C-5 epimerization of GlcA to IdoA residues and incorporation of *O*-sulfate groups at various positions [113] (Figure 1.5). As a result of such modifications that occur during chain extension in the biosynthesis of heparin and HS, the end products are highly heterogeneous with regard to chain length, charge, and substitution pattern [114]. Therefore, the combination of the different structural units and the arrangement of the disaccharides along a chain makes these molecules extremely complex,

17

with 32 (2^5) potential modifications for each disaccharide unit [115]. Heparin and HS are structurally closely related due to the same building blocks and identical glycosidic linkages, however they differ in their overall degree of sulfation [105]. While heparin exhibits a relatively uniform high level of sulfation with 2.5 to 3 sulfate groups per disaccharide unit, HS chain contains usually 0.5 to 1.5 sulfate groups per disaccharide and shows a more varied substitution pattern with alternating *N*-sulfated and *N*-acetylated domains (NS- and NA-domains, respectively) and a smaller amount of *N*-unsubstituted GlcN units [116], [117]. Also, whereas IdoA predominates in heparin, GlcA corresponds to the majority of the uronic acid present in HS [118], [119]. Heparin possesses the highest negative charge density of any know biological macromolecule and although the structural heterogeneity observed for HS is usually greater than for heparin, due to a higher occurrence of minor sequence variations, heparin can often mimic HS in *in vitro* studies. HS chains are generally longer than heparin and have an average molecular weight of 30 kDa as compared to 15 kDa for heparin [110], [113], [120].

As a result of their polyanionic character as well as their distinct composition, flexibility, length and specific sulfation patterns, heparin and HS have great potential for interaction with positively charged moieties such as plasma proteins thereby mediating their biological functions [108], [119]. To date, the most studied interaction occurs between a pentasaccharide sequence of heparin and the serine protease inhibitor antithrombin III (AT) of the coagulation cascade. This specific interaction leads to a conformational change in AT resulting in heparin's anticoagulant activity [121], [122]. Nowadays, heparin is widely used as an anticoagulant drug and pharmacological active pentasaccharides, such as Fondaparinux (also known as Arixtra), have been synthesized [123], [124]. However, since generation of GAGs with defined sequences is still difficult, the most commonly used heparins for pharmaceutical preparations comprise mixtures of high and low molecular weight heparins (LMWHs) [5], [125].

1.2.2 Challenges to investigate glycans and glycan-binding

The structural complexity that contributes to the diversity of GAGs and glycans in general also poses challenges for their isolation and molecular characterization [5]. As a result of the biosynthetic pathways, GAGs and glycans are highly heterogeneous and obtaining single molecular entities for functional and structural studies is still hardly possible [68]. Additionally, many glycan-protein interactions have low affinity with K_d-values typically within the milli- to micromolar range. [126]. In most studies using GAGs, enzymatic or chemical digestion prior to analysis is required to obtain defined oligosaccharides.

A common approach involves the use of GAG-depolymerizing enzymes, such as bacterial heparinases followed by fractionation of the smaller oligosaccharides according to length, chemical composition and charge [127]–[129]. Glycan-protein interactions can be subsequently studied by various methods such as nuclear magnetic resonance (NMR), SPR, X-ray crystallography and mass spectrometry (MS) [130], [131]. More recently, glycan arrays have also become a common tool to identify interactions involving glycans in a high-throughput manner. To address the challenge of glycan heterogeneity in nature, this method uses glycans or GAGs from natural sources or obtained via chemical and enzymatic synthesis. These molecules are then fluorescently tagged and attached to the surface of the array via covalent or non-covalent binding and glycan-protein interactions are determined by antibody-mediated fluorescence measurements and MS [97], [132].

1.3 GAG depolymerization by heparinases

Glycosaminoglycans, namely heparin and HS, play key roles in diverse biological activities through their interaction with functional proteins. However, identifying GAG sequences that can target proteins with a high level of specificity remains a challenge since it is still difficult to obtain chemically defined GAGs in high purity and quantity [108]. Most methods of GAG analysis involve a depolymerization step utilizing enzymatic digestion of long and heterogeneous GAG polymer chains to smaller fragments by polysaccharide lyases (PLs) [4]. The availability of these enzymes, especially heparinases from *Pedobacter heparinus* (*P. heparinus*), has greatly facilitated the analysis of heparin and HS since they cleave GAG chains with unique specificity [133], [134]. However, besides the high cost required for production of GAGs oligosaccharides with heparinases, the major product obtained is a disaccharide. According to previously published studies, defined GAGs with a higher degree of polymerization (dp) can target proteins with a higher level of specificity [135]–[137].

1.3.1 *Pedobacter heparinus* and the polysaccharide lyases

P. heparinus is a gram-negative, non-pathogenic soil organism that is capable of using either heparin or HS as a sole carbon, nitrogen and sulfur source. This bacterium, first described in 1956 by Paza and Korn as *Flavobacterium heparinum* (*F. heparinum*), has been extensively investigated due to its production of various enzymes involved in the degradation of polysaccharides, such as heparinases and chondroitinases [138]–[140]. Although, it is still unknown why a soil bacterium expresses PLs, it is likely that *P. heparinus* could use these enzymes to degrade GAGs from carcasses [141]. Based on their primary structure, these PLs have been classified into 23 families within the Carbohydrate Active enZymes (CAZymes) database [142].

Most heparinase-producing bacteria were found in soil, however heparinases isolated from human intestine bacteria such as *Bacteroides stercoris* HJ-15 and *Bacteroides thetaiotaomicron* (*B. thetaiotaomicron*) have been reported [141], [143], [144]. Heparinases from *P. heparinus* are the most actively studied and widely used heparin/HS degrading enzymes [104]. This soil bacterium is capable of producing three heparin lyases differing in primary sequence, size, charge properties and substrate specificities: the 42.8 kDa heparinase I (family PL13), the 84.1 kDa heparinase II (family PL21) and the 70.8 kDa heparinase III (family PL12) [145], [146]. These periplasmic enzymes are positively charged at neutral pH (pI 8.5-10) and have been isolated using either sonication or osmotic shock. In 1992, Lohse and Linhardt reported the first single purification method to purify all three

heparinases to homogeneity [145], [147]. Sequence alignment of heparinases I, II and III revealed a low sequence homology (15% residue identity) at both DNA and amino acids levels [148].

Heparinase I (EC 4.2.2.7) is specific for heparin and primarily cleaves next to IdoA (Figure 1.6A), whereas heparinase III (EC 4.2.2.8) is specific for HS, cleaving next to GlcA, and heparinase II (no EC number assigned) has a wider specificity and can depolymerize both heparin and HS [149], [150]. In contrast to eukaryotic GAG depolymerizing enzymes that predominantly utilize a hydrolytic mechanism to cleave the glycosidic linkage, heparinases from bacterial sources catalyze the cleavage of heparin and HS chains through a β-elimination mechanism [104], [151]. The detailed β-elimination mechanism was proposed by Gacesa in 1987 and involves neutralization of the negatively charged C5 carboxylate group of the substrate followed by a base-catalyzed proton abstraction at C5 and subsequent β-elimination of the 4-*O*-glycosidic bond [152], [153]. This degradation mechanism leads to cleavage of glycosidic linkages formed between the amino sugar and the uronic acid and results in the formation of two chemically distinct ends: the sugar on the newly formed non-reducing end containing an unsaturated ring and the sugar on the new reducing end, which remains saturated [146]. According to commonly used nomenclature, sugars of the substrate are numbered in both directions from the cleaved bond. Sugars from the cleavage site toward the reducing end are numbered +1, +2, etc., whereas sugars toward the non-reducing end are labeled -1, -2, etc. with cleavage taking place between the -1 and +1 subsites [154].

The double bond formed between the C4 and C5 carbons of the newly generated oligosaccharides at the non-reducing end moiety is conjugated with a carboxy moiety and serves as a chromophore that can be detected by UV light absorption at 232 nm (Figure 1.6B) [141]. End products of these enzymatic reactions are usually described by their degree of polymerization, which corresponds to the total number of monomer units in that molecule [129]. Following enzymatic depolymerization, resulting mixtures of oligosaccharide fragments are physically separated based on their different sizes and charges. First, size exclusion chromatography is used to separate mixtures into size-uniform fractions such as disaccharides, tetrasaccharides, hexasaccharides and larger oligosaccharides, which can then be further purified to individual oligosaccharides using strong anion exchange or reverse-phase ion-pair high-performance liquid chromatography [155]–[157].

A

B
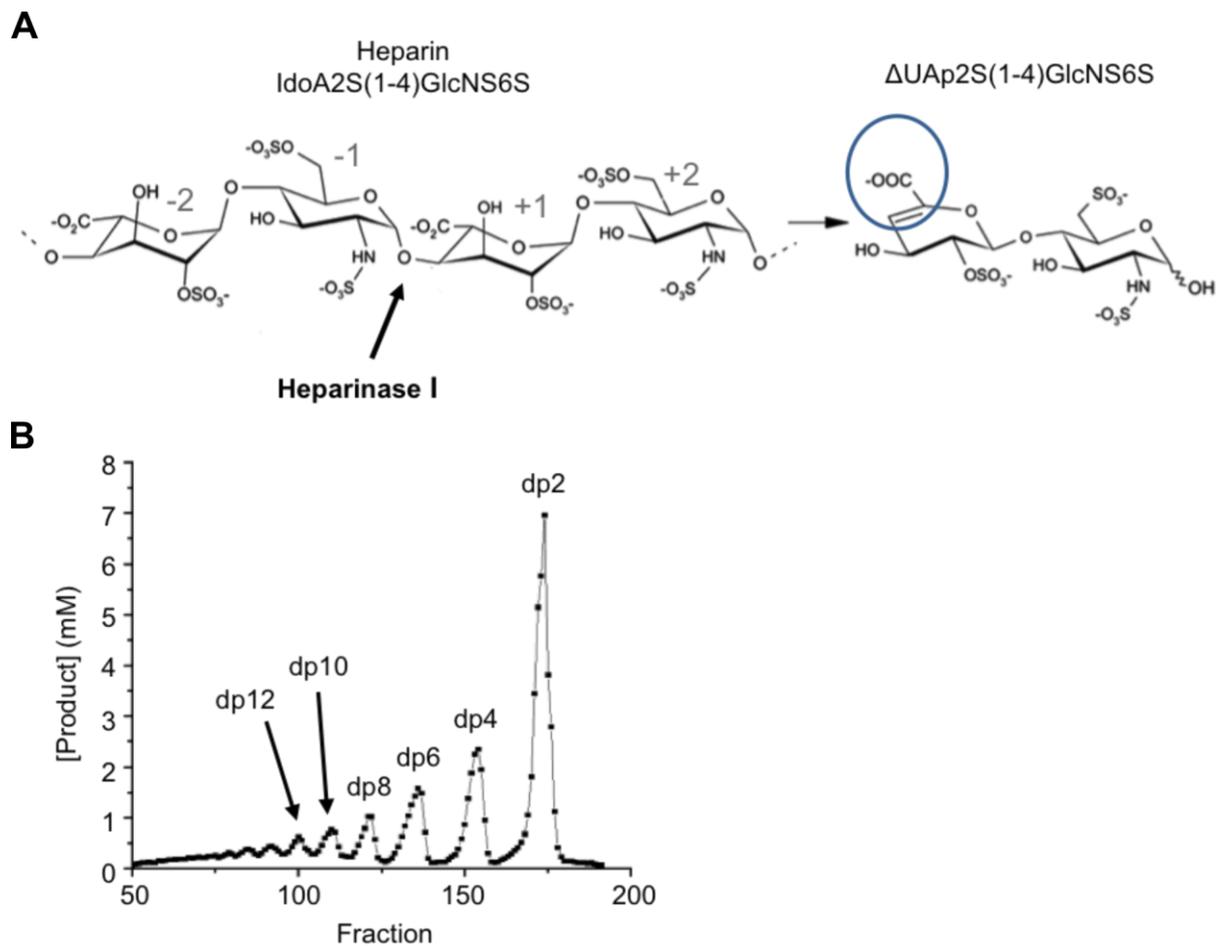
Figure 1.6 Heparin depolymerization by heparinase I. A) Degradation of heparin by heparinase I. Heparin cleavage takes place between subsites -1 and +1 (indicated in gray) and yields oligosaccharides containing a 4,5-unsaturated bond (highlighted in blue). B) Separation of heparin-derived oligosaccharides by size exclusion chromatography. The unsaturared bond can be detected by UV light absorption at 232 nm and thus be used to monitor the enzymatic reaction. The labels represent the number of saccharide units in the depolymerization products with dp2 (disaccharides) corresponding to the major product of the reaction. Product concentrations can be determined using the measured aborbance and the molar extinction coefficient for the unsatured disaccaride. Figure adapted from [155], [174].

Even though the substrate specificities of heparinases I-III are well known and the β-elimination mechanism is highly conserved across different PLs, the sequence of events that occur during enzymatic depolymerization of heparin is still controversial [146], [158]. Elucidation of the residues critical for catalysis as well as the substrate binding domains has been useful to understand the mechanism of action of these enzymes and for their development as tools for heparin and HS analysis. Although heparinases from *P. heparinus* are the best characterized, only structures of heparinases II and III are available up to date and no three-dimensional structure of heparinase I has been reported [159]–[161].

1.3.2 Structure and mechanism of heparinase I

Heparinase I from *P. heparinus* was isolated and purified to homogeneity for the first time by Yang *et al.* in 1985 [162]. Since then the properties of this enzyme have been well characterized through a combination of site-directed mutagenesis, chemical and proteolytic digestions and other biochemical methods [163]. Heparinase I is a monomeric protein with an approximate molecular weight of 43 kDa and contains a large number of lysine residues (10%) consistent with its cationic nature (pI of 8.5). However, this enzyme is very sensitive to thermal denaturation and is most stable at temperatures below 30°C [151], [162]. As previously mentioned, this enzyme is found in the periplasm of *P. heparinus* and its primary structure comprises a typical prokaryotic native leader sequence that consists of 21 amino acids. Cloning and expression of this enzyme lacking the signal peptide in *E. coli* has been successfully accomplished and showed that these first 21 amino acids are not essential for enzymatic activity [164].

Multiple biochemical studies have been carried out to understand the structure-function relationship of heparinase I and its mechanism to depolymerize heparin. Based on these studies, a model of heparinase I binding site was proposed with the substrate binding pocket containing the catalytic domain comprising amino acids critical for its activity: C135, H203 and K199 [165]–[167]. Besides, it was also hypothesized that the thiol group of C135 is negatively charged and due to is positively charged surrounding, it could act as a nucleophile and thereby initiate the β-elimination reaction in the active site [165]. H203 has also been implicated as a second acid catalyst serving to protonate the leaving hexosamine sugar. The highly basic heparin-binding site of heparinase I contains two Cardin-Weintraub heparin-binding consensus sequences (CB-1 and CB-2) motifs that seem to be involved in Ca^{2+} binding and enzymatic activity with CB-1 and CB-2 as high and low affinity sites, respectively. These binding sites might bridge heparin to heparinase I through Ca^{2+} in a ternary complex and orient the substrate in the active site region during catalysis [168], [169]. Although heparin binding to heparinase I is independent of Ca^{2+}, this ion is important for its catalytic activity [169], [170]. The role of the positive charges within the heparin binding domain has also been investigated and results indicated that lysines 132, 198 and 199 influence the enzymatic activity of heparinase I, with K199 showing a dominant effect, which suggested that it has a direct role in catalysis. [167]. All these findings are therefore consistent with the hypothesis that besides the consensus heparin binding motifs, the basic nature of the heparin binding domain present in heparinase I provides the necessary charge complementarity for specific heparin binding and creates the basic environment required for the catalytic activity [165], [169].

Using matrix-assisted laser desorption ionization/mass spectrometry (MALDI-MS) and capillary electrophoresis, it was discovered that the mode of action of heparinase I is a two-step mechanism, with depolymerization occurring from the non-reducing end of the polysaccharide chain and leading to cleavage of the entire heparin molecule. In an initial step, heparin cleavage is preferentially exolytic at the non-reducing end of the substrate, although around 5% of endolytic cleavage also takes place at internal linkages. In a second step, the reaction is highly processive since this enzyme has a strong preference for cleaving the same substrate molecule sequentially towards the reducing end of the heparin chain before releasing it. After cleavage, the polysaccharide slides along the active site of heparinase I until the next cleavable bond is correctly oriented for catalysis [171], [172]. The trisulfated heparin disaccharide ΔUA2S-GlcNS6S is the major end product of this reaction [173].

Although no structure of heparinase I from *P. heparinus* has been solved yet, crystal structures of heparinase I from *B. thetaiotaomicron* without substrate and in complex with varying substrate lengths, including the H151A inactive mutant in complex with a dodecasaccharide heparin (Figure 1.7), were determined by Han *et al.* [174].
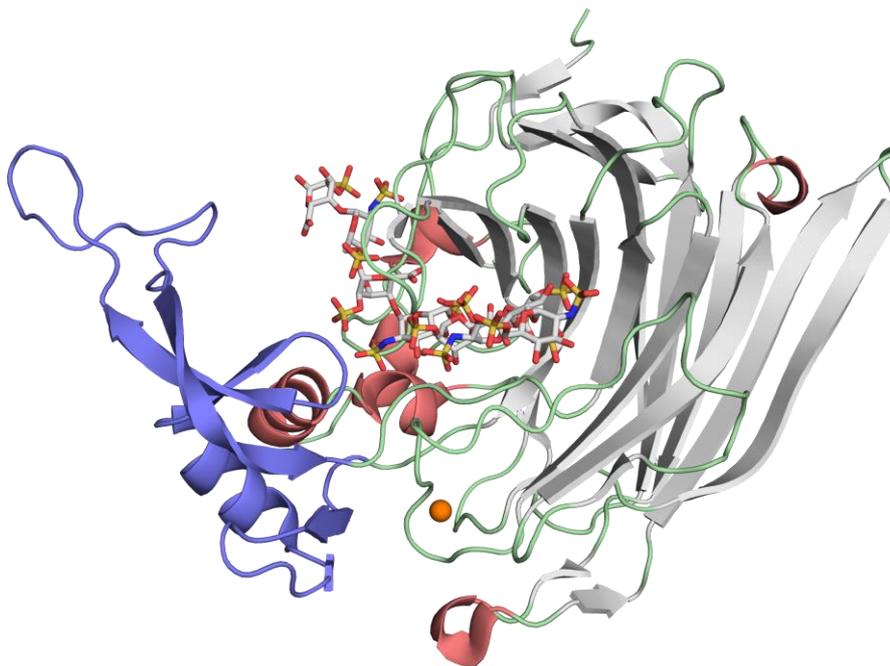
Figure 1.7 Overall structure of H151A-inactive mutant of heparinase I from *B. thetaiotaomicron* in complex with a dodecasaccharide heparin (PDB 3INA). Protein backbone is highlighted according to different secondary structure elements, except the thumb domain, which is colored in purple. The active site, containing the heparin chain depicted in stick representation, transverses the heparinase I and is located next to the β-jellyroll fold. The single Ca^{2+} ion is shown as an orange sphere. Figure adapted from [174].

This enzyme shows the classical β-jellyroll fold made of two antiparallel eight stranded β-sheets bent in the middle. In addition, a thumb-like insertion domain (thumb domain) was observed and it appears to participate in substrate binding. The resulting curvature in the inner-sheet of the β-jellyroll creates a long, extended canyon where the heparin substrate binds. This highly positively charged cleft also comprises the active site and the loops linking the β-sheets are variable in length and sequence [146], [174]. In contrast to heparinase I from *P. heparinus*, for which two Ca^{2+} binding motifs have been reported, a single Ca^{2+} bound in the hinge region between the β-jellyroll fold and the thumb domain was observed in the heparinase structure of *B. thetaiotaomicron*. This ion seems to organize residues involved in heparin binding and thus support the structural integrity of this enzyme. By analyzing three-dimensional structures captured at various stages of the reaction, the 3-step catalytic mechanism of heparinase I from *B. thetaiotaomicron* was proposed with H151 and Y357 residues acting as a general acid and base, respectively [168], [174].

Crystal structures of other PLs have revealed that, although these enzymes have different structural folds, each one of them contains an elongated deep cleft to accommodate the polysaccharide substrate prior to catalysis [175]. Residues along this cleft interact with the substrate via hydrogen bonds, salt bridges and van der Waals contacts. These interactions are not only necessary for substrate binding but also to precisely orient the polysaccharide molecule for catalysis. Therefore, substrate specificity is derived from the shape of the active site cleft to precisely fit a given polysaccharide as well as the bonds that are formed with the different sugar rings [159], [176]. Like heparinase I, heparinase II is also able to depolymerize heparin and despite having very different overall architectures these enzymes show similarities in the arrangement of their catalytic residues [174]. One of the consensus sequences in heparinase II shares homology with a consensus sequence in the heparin binding site of heparinase I with the conservation of the positive charge and the catalytically histidine residue [148]. This represents a good example of convergent evolution of two different protein folds that adopted a similar enzymatic mechanism [146].

1.3.3 Heparinase I as a biochemical tool and its applications

Heparinases, including heparinase I, have a wide range of applications in health care. Besides demonstrating potential for treatment of various diseases such as metastasis in various types of cancers, inflammation and angiogenesis, these enzymes have been valuable tools in the elucidation of heparin structures, production of therapeutically important LMWHs and blood deheparinization [136], [158], [177].

In comparison to chemical depolymerization, enzymatic depolymerization of heparin by heparinases for production of LMWHs as anticoagulants and other heparin-derived oligosaccharides has a high selectivity. In addition, it offers the advantage that reaction conditions are milder and the reaction can be conducted without altering the fine structure of the fragments [4]. Up to date, *P. heparinus* is the major source for production of heparinase I. However, due to low yield and poor stability of this enzyme, costs for its application in industrial production of LMWHs are too high. Even though recombinant heparinases have been produced, obtaining large amounts of enzyme in an active and soluble form remains a major challenge [178], [179].

Moreover, although it is recognized that specific sequences of heparin and HS, typically from tetra- to decasaccharide in length, are responsible for modulation of biological activities of proteins, generation of such oligosaccharides is still difficult [180]. GAGs with a higher degree of polymerization, such as dp6 and dp8, can target proteins with a higher level of specificity due the possibility of reduced non-specific interactions [181]–[183]. Partial digestion of heparin and HS chains has been useful in generating small amounts of longer oligosaccharides. However, this strategy has a low reproducibility and requires individualized processing for every batch of depolymerization product [137]. An alternative is the modification of heparinase's activity to decrease production of some oligosaccharides and increase the production of others, which are also usually present in very small amounts. For heparinase I, two examples have been reported, including the use of histamine in the depolymerization reaction as well as the enzyme immobilization on a solid support. Such approaches were able to either block potential cleavage sites or alter the properties of heparinase I and thus yield a different product profile distribution of heparin fragments [137], [184].

A detailed understanding of the mechanism of action during heparin depolymerization by heparinase I together with the residues involved in its catalytic activity would provide a strong basis towards the development of enzymes with improved properties for a variety of applications. Heparinase engineering could create possibilities to produce more favorable LMWHs or longer oligosaccharides with specific sequences crucial for the biological activities of proteins. Development of rational approaches to protein engineering can lead to the design of novel tailored molecules with improved properties regarding stability, activity and potency. Structure-based protein engineering is a potential strategy to effectively incorporate site-specific changes at selected regions of the proteins [104], [137], [163]. This protein engineering method requires the combination of multiple techniques, including gene cloning, site-directed mutagenesis, protein expression and biochemical characterization, as well as a detailed level of structural information [185], [186].

1.4 X-ray crystallography

Detailed knowledge of protein structures and macromolecules assemblies has been vital to study numerous biological processes from enzymatic reactions to immune evasion by viruses. Three-dimensional structures have revealed how enzymes interact with their respective substrates and how inhibitors can bind and disrupt activity, which is essential in the fields of protein engineering and drug discovery. To determine three-dimensional structures of proteins at atomic resolution, electromagnetic radiation with a shorter wavelength than the visible light, such as X-rays, is required. To date, one of the most powerful methods to determine protein structures is X-ray crystallography. In the so-called X-ray diffraction experiments, macromolecules arranged in a regular array, called crystal lattice, are exposed to electromagnetic waves with wavelengths of 0.5-2.0 Å [187].

1.4.1 Protein crystallization and cryoprotection

Protein crystals are highly ordered three-dimensional arrays comprising regular repetitions of protein molecules, which are held together by intermolecular interactions. In practice, initial protein crystals are a result of a high-throughput screening approach, performed with an automated crystallization setup and employing highly purified protein samples. Different precipitants, such as salts, alcohols or water-soluble polymers like polyethylene glycols (PEGs) are used to reduce protein solubility and initiate crystal growth. Protein crystallization is commonly achieved using the hanging or sitting drop vapor diffusion methods. A drop of pure protein is mixed with a precipitant-containing reservoir solution and equilibrated against the reservoir solution under sealed conditions. Since the protein-precipitant mixture is less concentrated than the reservoir solution, water molecules diffuse from the drop into the reservoir. Protein and precipitant concentrations gradually increase until the protein's solubility limit is exceeded and the protein drop enters the supersaturated phase. When the precipitant or protein concentration is too high, the protein precipitates. However, under optimal crystallization conditions, the supersaturated solution reaches the nucleation zone, where crystallization nuclei can spontaneously form and reach a critical size to support subsequent crystal growth (Figure 1.8). Once crystals are in equilibrium with the saturated protein solution, they have reached their maximum size and growth terminates [187], [188].
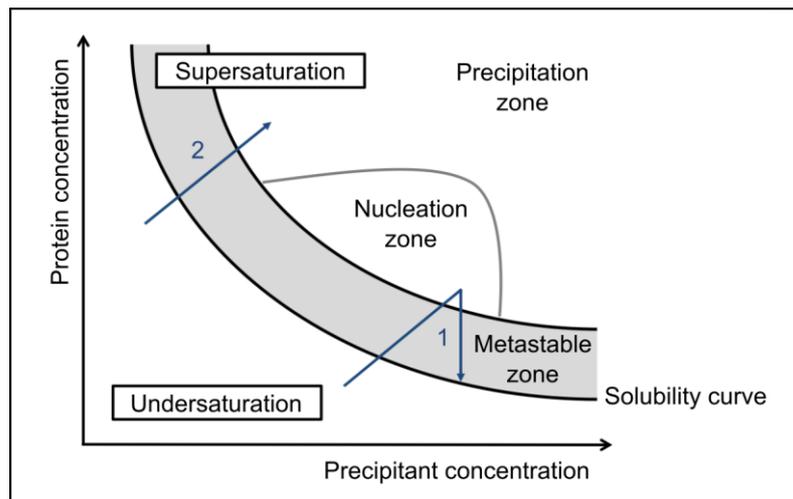
Figure 1.8 Phase diagram illustrating protein crystallization by vapor diffusion. Equilibration of the protein-precipitant mixture against the reservoir solution increases both protein and precipitant concentrations. Crystal formation only occurs in the nucleation zone with subsequent crystal growth in the metastable zone. Arrow 1 illustrates a successful crystallization experiment where initial protein crystals are formed in the nucleation zone with subsequent crystal growth in the metastable zone. Arrow 2 represents an unsuccessful crystallization experiment where protein precipitation occurs Figure adapted from [196].

Protein crystallization and crystal growth are influenced by parameters such as temperature, pH, protein and precipitant concentrations as well as drop and reservoir volumes. Seeding is an alternative approach to induce nucleation and facilitate crystallization. Thereby, solid material, such as small or crushed crystals and spherulites can be used as nucleating agents. These nucleating agents commonly called seeds are transferred to a crystallization drop to bypass the nucleation zone and therefore reduce the level of supersaturation that is required for crystal growth. Based on the size of the seeds, seeding is classified into micro- or macroseeding. In microseeding, tiny crystalline fragments are added to a crystallization drop of reduced supersaturation, whereas in macroseeding, a small single crystal (usually 5-50 μm) is transferred into a new crystallization solution with identical composition to continue crystal growth [189], [190]. Protein crystals are very fragile since they are loosely packed and contain a high fraction of solvent, accounting for about 30% to 70% of the crystal volume. Solvent channels allow diffusion of ligand molecules through the crystal, often allowing ligand molecules to bind to the protein molecules in the crystalline state. During diffraction data collection, protein crystals are exposed to intense ionizing X-ray radiation causing crystal damage due to the formation of free radicals. To minimize crystal decay during X-ray exposure, crystals are cooled to cryogenic temperatures (about 100 K). Prior to flash-freezing with liquid nitrogen, crystals are transferred into cryoprotection conditions to prevent formation of crystalline ice that would result in strong diffraction of the ice crystals [191].

1.4.2 X-ray diffraction and data collection

In a diffraction experiment, X-rays are scattered when the electromagnetic radiation interacts with the electrons of molecules in the protein crystal. Since the scattering of X-rays by a single molecule is too small to be detected, an arrangement of many molecules in a periodic lattice is required and the sum of all scattering events generates a diffraction pattern which reflects the crystal symmetry. The crystal lattice is formed by periodic repetition of a three-dimensional unit cell defined by the length of unit cell axes a, b, c and the unit cell angles α, β and γ. If multiple protein molecules with identical conformation are present in the unit cell, they can be transformed into each other by translational and rotational symmetry operations, and these symmetry operations are required to fully describe the protein crystal. The complete unit cell can thus be generated by applying crystallographic symmetry operations to the smallest unit present in the crystal, denoted as the asymmetric unit (ASU). Because biological molecules are chiral, some symmetry elements such as inversion centers or mirror planes, which occur in inorganic salt crystals, are not possible in their crystals. There are 65 possible space groups for crystals containing proteins [187].

The interpretation of X-ray diffraction data was first described by Sir William Lawrence Bragg as a reflection of X-rays by imaginary planes in a crystal lattice [192]. According to Bragg's law, constructive interference of elastically scattered X-rays by a crystal lattice can give rise to a diffracted beam and can be observed in the form of discrete diffraction spots on a suitable detector (originally a photographic film) (Figure 1.9A). These spots are called reflections and the dimensions of the unit cell in the crystal lattice are inversely proportional to the distances that separate the recorded reflections. During the X-ray diffraction experiment, the crystal is continuously rotated, giving rise to multiple diffraction images ('frames'). The virtual lattice generated from the overall diffraction set is defined as the reciprocal lattice. Lattice planes in the reciprocal lattice and their directions in three-dimensional crystals are specified by the Miller indices h,k,l and vectors $\vec{h}, \vec{k}, \vec{l}$. As described by Bragg's equation (1.1), a reflection (h,k,l) is obtained when the path difference between X-ray waves with wavelength λ diffracting from parallel crystal planes, separated by a distance d under an angle of reflection θ, is equal to an integer n.

$$n\lambda = 2d_{hkl}sin\theta \qquad\qquad (1.1)$$

The maximum angle θ corresponds to the minimum distance d_{min} between a set of lattice planes that can still be resolved, and therefore describes the resolution limit of a protein crystal.

Bragg's law can also be expressed through a geometrical construction, designated the Ewald sphere (Figure 1.9B). In this construction, a sphere with a radius 1/λ is drawn around the

crystal origin. The incident X-ray beam is directed towards the center of the sphere and when reciprocal lattice points (h,k,l) lie on the Ewald sphere Bragg's law is obeyed and reflections can be recorded on a detector. The reciprocal space origin (0,0,0) is placed at the intersection of the Ewald sphere with the incident beam. During data collection only a few reflections can be observed at a given crystal orientation. Therefore, to record all reflections of a crystal, the reciprocal lattice points are brought to intersection with the Ewald sphere by rotating the crystal in the center of the sphere, and its reciprocal lattice around its own origin [193]. The number of recorded reflections is dependent on the crystal symmetry and diffraction limit.
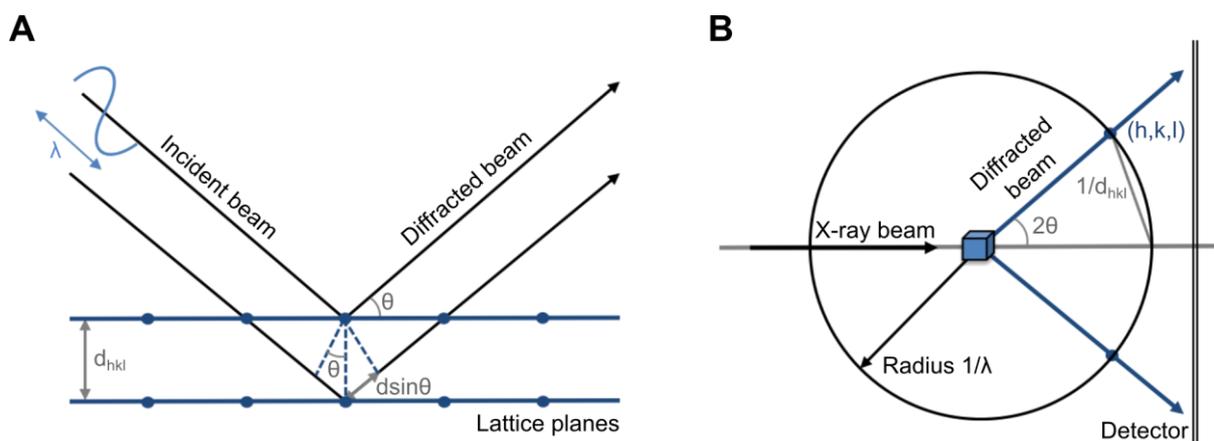
Figure 1.9 Bragg's law and Ewald sphere construction. A) Illustration of Bragg's law. In diffracting conditions, constructive interference occurs when the path difference of two waves diffracted at parallel crystal lattice planes equals an integer multiple of the wavelength of the incident X-ray beam. The scattered X-rays are then in phase and a reflection spot is produced. B) Ewald sphere construction with a radius of $1/\lambda$ around the crystal. Reflections with coordinates h, k, l are recorded when Bragg's law is fulfilled and reciprocal lattice points lie on the Ewald sphere. Figure adapted from [187], [194].

The observed diffraction spots can be assigned to the corresponding reciprocal space coordinates (h,k,l) during so-called indexing of the diffraction data. The information about the position of these reflections together with the distance of the detector and the wavelength used during data collection allows the calculation of the unit cell parameters. Once these parameters are correctly determined, intensities of all measured reflections can be obtained by integration. Accurate diffraction data measurements require high redundancy, which can be achieved by multiple independent measurements of equivalent reflections. Further data processing includes the application of scaling factors to render the data internally consistent and minimize differences between equivalent reflections, i.e. account for instance for the increasing decay of a crystal's diffraction power due to inelastic scattering events. Symmetry-related reflections are then merged and converted into structure factors. When more than a single molecule is present in the asymmetric unit, phenomena such as twinning,

and pseudo-translational or rotational symmetry can occur, making structure determination more difficult and sometimes impossible [195].

1.4.3 Structure determination and refinement

To determine the distribution of electrons in the crystal and to obtain an electron density map, information of both amplitude and phase of each recorded reflection are required. This information is used to define a complex number, termed the structure factor F_{hkl}, which represents the sum of the scattering contributions from each and every atom in the unit cell (equation 1.2).

$$F_{hkl} = |F_{hkl}| \cdot \exp\left(i\varphi_{hkl}\right) \tag{1.2}$$

The structure factor is a vector in the Argand diagram with the amplitude $|F_{hkl}|$ and phase angle φ_{hkl} that describes each scattered beam that leads to a reflection (h,kl,) in reciprocal space. By determining the amplitude and phase for each F_{hkl} and applying a Fourier transformation, the electron density can be calculated for the crystal unit cell (equation 1.3) and the structure of the underlying diffracting molecule that 'decorates' the unit cell can be determined.

$$\rho(x, y, z) = \frac{1}{V} \Sigma_{h,k,l} F_{hkl} \cdot \exp\left[-2\pi i(hx + ky + lz)\right] \tag{1.3}$$

In this equation, x,y,z are the coordinates in real space and V corresponds to the volume of the unit cell. In a diffraction experiment, structure factor amplitudes can be determined directly from the measured intensities for each reflection since the intensity of a reflection is proportional to the square of the structure factor amplitude. However, the phase information cannot be extracted from the diffraction data, and as consequence, the electron density function ρ(x,y,z) cannot be directly obtained giving rise to the so-called phase problem [187], [196].

To overcome the phase problem and determine three-dimensional structures of proteins from diffraction patterns, several methods can be employed such as molecular replacement (MR), single or multiple isomorphous replacement (SIR/MIR), and single or multiple- wavelength anomalous dispersion (SAD/MAD). In case of small proteins and very high data resolution (at least 1.2 Å), structures can also be determined *ab initio* via direct methods [197].

MR is widely used when a structural model similar to a molecule in the target crystal structure is available (for proteins usually more than 25% sequence identity is required). The known structure can be used as a preliminary model for the unknown structure. As structurally related protein molecules share similar Patterson maps, these maps can be

aligned by rotation and translations functions to place the known structure in the correct orientation and position in the unit cell of the unknown structure. In contrast to the electron density map, which requires both phase and amplitude information, a Patterson map can be calculated directly from the observed diffraction data, as phase information is not required. The Patterson function is the Fourier transform of the squared structure factor amplitude $|F_{hkl}|^2$ with phases set to zero. The Patterson map derived from the observed data is a map of interatomic vectors containing information about the contents of the crystal and the obtained peaks in Patterson space correspond to the positions of the vectors between atoms in the unit cell. Once a MR solution is found, the initial phases derived from the known structure are combined with the measured structure factor amplitudes of the unknown structure and an initial electron density map with preliminary phases is calculated [198]. Model building is an iterative process for which the structure factors calculated from the model (F_{calc}) are constantly compared to the experimental structure factors (F_{obs}). Two types of maps are commonly calculated: the electron density difference map (F_{obs}-F_{calc}) and the double difference map ($2F_{obs}$-F_{calc}) from which the molecular model is stepwise improved by modeling the amino acids into the density according to the primary sequence of the protein. This model is validated and improved through refinement cycles against the experimental data using the maximum-likelihood method.

During structural refinement atomic coordinates and atomic displacement parameters are improved to maximize the agreement between the built model and the experimental X-ray data [199]. The refinement progress is monitored by crystallographic R-factors which compare measured (F_{obs}) and calculated (F_{calc}) structure factor amplitudes (equation 1.4).

$$R = \frac{\Sigma_{hkl}||F_{obs}(hkl)|-|F_{calc}(hkl)||}{\Sigma_{hkl}|F_{obs}(hkl)|} \tag{1.4}$$

During refinement of a protein model, the R-factor value should decrease indicating the convergence of the model with the experimental data. Data over-fitting can be observed by using a cross-validation in the form of a free R-factor. The R_{free} is used to assess model and refinement quality. A small subset of reflections (5 to 10% of the data set) is flagged as 'free' and is excluded during refinement. This set of reflections allows an unbiased estimate of the improvement of the structure model and is usually higher than the R_{work}, which comprises the reflections used during refinement [200]. Remaining bias from the MR model can be removed by calculating simulated annealing-omit maps. Phases can be further improved by applying a variety of density modification methods, like solvent flattening and non-crystallographic symmetry (NCS).

1.5 Inhibition of c-Jun N-terminal 3 kinase

Elucidating atomic details of molecular interactions of protein-ligand complexes is essential for structure-based design and optimization of potential therapeutic drugs. Over the past decades, protein kinases became attractive targets in drug discovery. The importance of the c-Jun N-terminal kinase (JNK) family in the pathogenesis of numerous diseases has triggered an extensive development of medicinal chemistry strategies. Some of these strategies were based on crystal structures to identify and optimize JNK inhibitors suitable for clinical application [201], [202]. However, despite a significant amount of publications concerning JNK inhibitors and several candidates that have reached clinical trial phases, no JNK inhibitor has been approved by the U.S. Food and Drug Administration (FDA) to this date [6], [7], [203].

1.5.1 Protein kinases

Protein kinases are signaling enzymes that mediate protein phosphorylation, an essential step of signaling cascades, and regulate nearly every aspect of cellular function including gene expression, cell growth and proliferation, differentiation and apoptosis. Protein phosphorylation is carried out by transferring the terminal phosphoryl group from a nucleoside triphosphate, typically adenosine triphosphate (ATP), to the hydroxyl group of a serine, threonine or tyrosine residue of the target substrate protein or the kinase itself [204]-[206]. The phosphorylation process can be reversed by corresponding phosphatases, which are able to catalyze the removal of the phosphate group from the target protein [207], [208]. The superfamily of human protein kinases, generally described as the kinome, comprises 518 genes, accounting for approximately 1.7% of the human genome-encoded proteins [209].

Most of the protein kinases share a common fold and have the same essential structural features but phosphorylate different target proteins and can be notably different in how their catalytic activity is regulated [210]. Differences in core sequence and flanking regions of kinases allow a fine-tuning of their activity, which results in a unique response of each kinase to a set of cell signaling events. Aberrant phosphorylation is implicated in an increasing number of diseases, making kinases potential drug targets in different therapeutic approaches. Treatment of such diseases is focused on selectively blocking the activity of the disease-associated kinase, which in turn prevents the effect of potential aberrant cell signaling events [211], [212].

1.5.2 Mitogen-activated protein kinases

Eukaryotic cells recognize and respond to extracellular stimuli by engaging specific intracellular processes, such as the signaling cascade that leads to activation of the mitogen-activated protein kinases (MAPKs). This evolutionary highly conserved family of enzymes transmits external signals to various cytoplasmic and nuclear targets and regulates diverse cellular processes, such as proliferation, differentiation, cell cycle and cell death [213].

Mammals express at least four distinctly related groups of MAPKs, including the extracellular signal-regulated kinases (ERK 1 and 2), the p38 MAP kinases (p38 α, β, γ and δ), the c-Jun N-terminal kinases (JNK 1, 2 and 3) and the ERK5. Each of the MAPKs can regulate several distinct and sometimes overlapping cellular processes, while differing in their main biological functions [214], [215]. MAPKs pathways are organized into kinase signaling modules, which are regulated by sequential phosphorylation (Figure 1.10). In most cases, MAPK signaling originates from the plasma membrane where transmembrane receptors are activated by extracellular ligands or stimuli. In general, ERK is preferentially activated in response to growth factors or mitogenic stimuli, whereas JNK and p38 kinases are most responsive to inflammatory cytokines and environmental stress, such as ionizing radiation, heat and oxidative stress [216], [217].

MAPK activity is regulated through a three-tiered cascade composed of MAPKs, MAPK kinases (MAPKKs) and MAPKK kinases (MAPKKKs) [218]. MAPKs are activated by dual phosphorylation of threonine and tyrosine residues located in the activation loop adjacent to the active site. Once activated, these kinases transmit signals by phosphorylating serine and threonine residues on cytoplasmic and nuclear proteins. Specificity for different targets is conferred by amino acids surrounding the phosphorylation site and through interactions mediated by another site within the kinase that recognizes a distinct site on the substrate [219], [220].
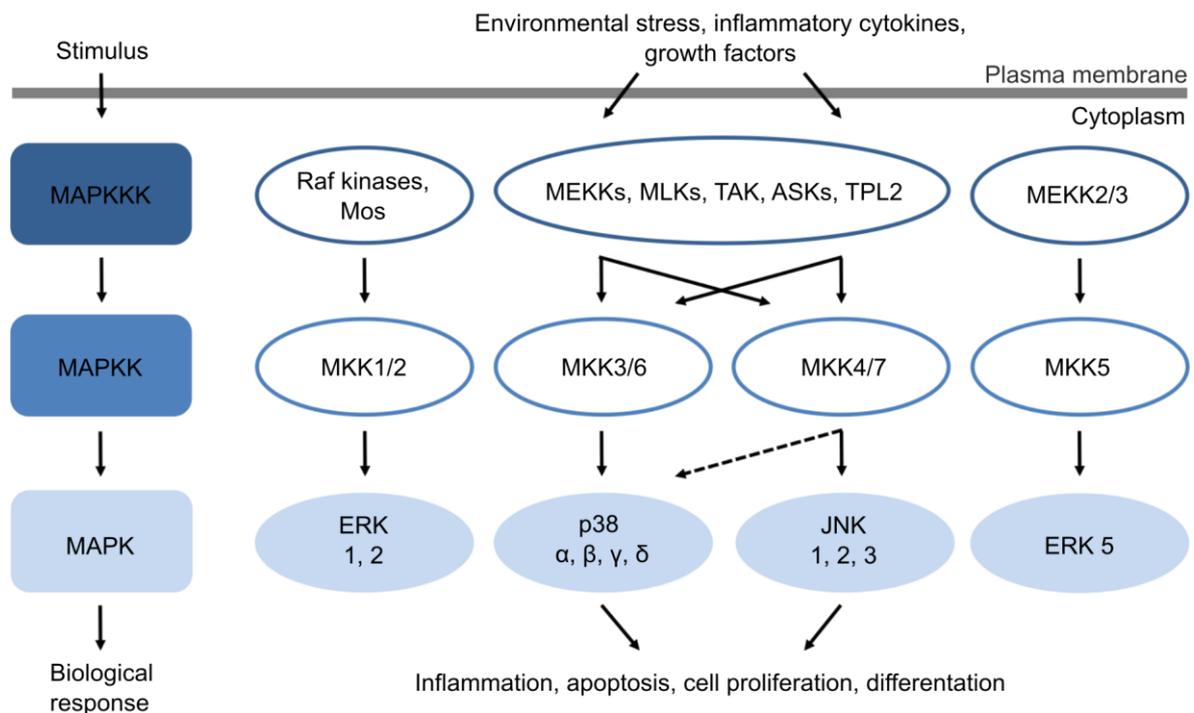
Figure 1.10 Overview of MAP kinase signaling cascades with focus on activation of p38 MAPKs and JNKs. Each family of MAPKs is composed of three evolutionary conserved, sequentially acting kinases. Extracellular stimuli activate MAPKKK, which leads to phosphorylation and activation of MAPKK and, subsequently, stimulates MAPK activity through dual phosphorylation. Once activated MAPKs phosphorylate target substrates. The dashed arrow represents the phosphorylation of p38 by MKK4. Figure adapted from [215].

In addition, several other regulators, activators and scaffolding proteins can influence the regulation of MAPK-mediated signaling pathways. Scaffolding proteins are multidomain proteins that can bind to multiple components of the MAPK cascade and thus increase the local concentration of cascade components, contribute to the spatial and temporal regulation of cascade activation, and/or localize the signaling module to a specific cellular site or substrate [221], [222].

1.5.3 The JNK protein family

JNKs, also known as stress-activated protein kinases (SAPK), were originally identified in 1990 by their activation in response to a variety of environmental stimuli including growth factors, cytokines, heat shock, hyperosmolarity and UV-radiation [223], [224]. The name of this protein family was then given after their ability to phosphorylate the N-terminal domain of the transcription factor c-Jun [225]. In mammals, three genes encoding for JNKs have been identified: *jnk1*, *jnk2* and *jnk3*, and each JNK gene has alternative splicing forms that result in at least 10 different isoforms [226]. Despite their structural homology, JNK genes follow a different tissue distribution pattern. *jnk1* and *jnk2* genes are ubiquitously expressed, whereas *jnk3* gene has a more limited pattern of expression and is predominantly restricted to the

central nervous system (CNS), cardiac smooth muscle and the testis. In addition to this, JNK1, 2 and 3 are characterized by distinct substrate specificities, suggesting that different JNK isoforms could play specific roles in multicellular organisms [218], [227], [228].

JNK activation requires dual phosphorylation of tyrosine and threonine residues within the conserved tripeptide motif T-P-Y located in the activation loop. This phosphorylation is mediated by two MAPK kinases, the MKK4 and MKK7. Although both kinases can activate JNK on both threonine and tyrosine residues, they seem to have different biochemical properties and significant differences in substrate specificity. First, even though full activation of JNK requires dual phosphorylation, MKK4 and MKK7 appear to preferentially phosphorylate JNK on tyrosine and threonine, respectively, indicating that these proteins function in a synergistic manner. Second, although, MKK3 and MKK6 are the main activators of p38 MAPKs, MKK4 can also phosphorylate p38. MKK7 functions instead as a specific activator of JNK [228]–[230]. An additional level of regulation of the JNK activation pattern is constituted by scaffolding and by other protein-protein interactions, which are crucial to the function of JNK in the cell. Examples of these scaffolds are JNK-interacting proteins (JIPs 1, 2, 3 and 4) and β-arrestins [228], [231].

Activated JNKs phosphorylate multiple targets, including nuclear and non-nuclear substrates. The main cellular substrate activated by JNK-mediated phosphorylation is c-Jun, which causes increased transcription activity. In addition, JNK can also phosphorylate the transcription factors JunB, JunD, c-Fos, and ATF-2 that constitute, together with c-Jun, the activator protein-1 (AP-1) transcription factor family, and regulate the expression of several stress-responsive genes. JNKs also phosphorylate other transcriptions factors such as Elk-1, c-myc, NFAT, and the p53, which are not part of the AP-1 complex [218], [232], [233]. JNK mediates apoptosis not only through its effects on gene transcription, but also through phosphorylation of both pro- and anti-apoptotic proteins, such as members of the Bcl-2 family (Bcl-2, Bcl-xL, Bim and BAD), and regulation of their activity [234], [235].

JNK is a multifunctional protein kinase involved in many physiological and pathological processes that modulate both cell survival and apoptosis depending on the cellular context. The JNK pathway plays critical roles in the pathogenesis of neurological disorders, including Alzheimer's disease and Parkinson's disease [236]. In addition to this, the JNK signaling pathway is also implicated in other diseases, such as type 2 diabetes, some types of cancer, stroke, heart disease, and inflammatory diseases [237]. Given the role of the JNK pathway in several pathological states, JNKs are promising drug targets for therapeutic intervention. Because of the complex cross-talk within the signaling cascade, as well as its response-specific modulation, it is, however, difficult to predict potential adverse events that might arise from JNK pathway inhibition [202], [228], [238].

1.5.4 Structural features of JNKs

Determination of three-dimensional structures of members of the JNK family and also other MAPKs has helped to understand the regulation of these enzymes and has also been crucial in the development of potential drug targets [202]. The first structure at atomic resolution of a JNK protein was determined in 1998 by Xie *et al.* [239]. This crystal structure containing inactive JNK3 in complex with the ATP analogue adenylyl-imidodiphosphate (AMP-PNP) and Mg^{2+} was determined at a resolution of 2.3 Å. The overall three-dimensional fold of JNK3 is similar to other protein kinases and it comprises all typical kinase motifs (Figure 1.11).
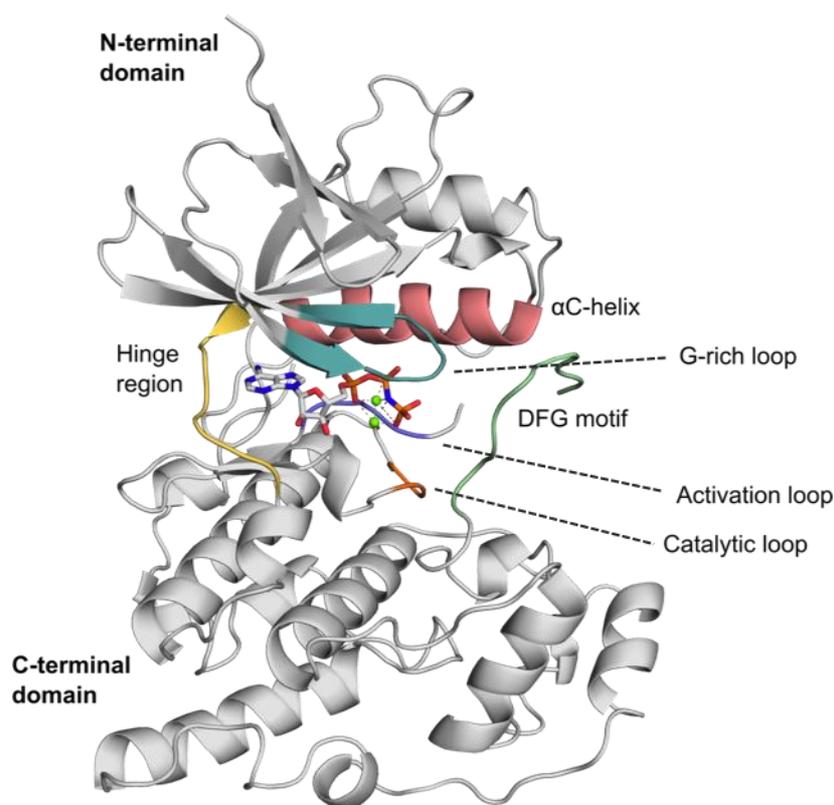
Figure 1.11 Three-dimensional structure of the inactive form of JNK3 in complex with AMP-PNP. The structure was reported for the first time in 1998 (PDB 1JNK) [239]. The active site, located between the N- and C-terminal lobes, is occupied with by the non-hydrolyzable ATP analogue coordinated by two Mg^{2+} ions (green spheres). The protein backbone is depicted in gray, while the characteristic kinase elements are highlighted in different colors. Part of the activation loop is missing, a common phenomenon for X-Ray structures of protein kinases since this loop is highly flexible.

Protein kinases contain 12 conserved subdomains, previously identified by sequence analysis, which are separated by divergent amino acid stretches that characterize each class of these enzymes [224], [240]. The catalytic domain of protein kinases, including JNKs, consists of a small, mostly β-stranded N-lobe connected by a hinge region to a large α-helical C-lobe. The cleft formed between the two lobes harbors the kinase activation site and the ATP-binding site. The adenine group of ATP is positioned between hydrophobic

residues and interacts with the hinge region via a bidentate hydrogen bond. The N-lobe or N-terminal domain comprises an antiparallel five-strand β-sheet motif and a single α-helix, denominated the αC-helix. This helix represents an important regulatory element for the catalytic activity, as it contains a glutamate residue that forms a salt bridge with a conserved lysine present in the β3 strand, which in turn coordinates the α- and β-phosphate groups of ATP. The most flexible part of this domain is a loop that lies between β1 and β2 strands, known as the glycine-rich loop or glycine-rich phosphate anchor loop, which comprises the conserved motif G-X-G-X-X-G-X-X. This loop is formed by hydrophobic residues that constitute the active site roof and contributes to the coordination of the triphosphate group of ATP via backbone interactions [241], [242]. In contrast, the C-terminal domain is mainly composed of α-helices and four short β-strands. The catalytic loop, located between β6 and β7 strands, comprises the highly conserved histidine-arginine-aspartate (HRD) motif, which plays a key role in the reaction catalysis. The aspartate residue takes part in the catalytic mechanism by deprotonating a hydroxyl group of the substrate. Within this loop there is also an asparagine residue that binds a Mg^{2+} ion coordinating the α- and γ-phosphate groups of ATP. The activation loop, located in the C-lobe, constitutes the primary site of phosphorylation and thereby plays a critical role in the kinase activation and deactivation processes. The length of this loop varies between different kinases and it usually contains one or more phosphorylation sites, which allow the activation of the enzyme by upstream kinases. In the case of JNKs, this sequence comprises the tripeptide motif T-X-Y. The activation loop is part of an extended activation segment, which begins at the DFG motif, a highly conserved sequence of three amino acids asparagine-phenylalanine-glycine that are part of the ATP binding site in the active kinase. The aspartate within this sequence chelates a Mg^{2+} ion that positions the β- and γ-phosphates for phosphotransfer, whereas the phenylalanine contributes to the correct positioning of the αC-helix in the N-lobe via two hydrophobic interactions. The activation segment plays the most critical role in regulation of the conformational changes in many kinases [243]–[245].

The flexibility of the activation loop enables each kinase to adopt two catalytically important conformations, which are dependent on the protein phosphorylation state. The active conformation, in which the substrate can bind, is called the DFG-in conformation. Upon phosphorylation, the activation loop moves away from the catalytic center allowing substrate binding and catalysis. In the inactive state, known as the DFG-out conformation, the activation loop collapses into the active site, preventing the substrate from binding to the kinase domain. In this case, the aspartate residue points towards the ATP binding site, which leads to a mispositioning of the Mg^{2+} ion that coordinates the ATP phosphates. Additionally, the phenylalanine residue is perturbed and thus influences the αC-helix, which in turn affects the position of the catalytically important glutamate and alters the interface between the

N- and C-lobes. Structures of protein kinases share very similar features in their catalytic active state due to the high conservation of residues involved in the phosphorylation reaction. However, there are a variety of conformations accessible to different protein kinases in the inactive site, making their structural characterization by X-ray crystallography a challenge. In the case of JNKs, only the DFG-out conformation of the JNK2 structure is available up to date [211], [246], [247].

1.5.5 Development of JNK inhibitors

During the past decades, a combination of high-throughput screening, kinase-specific libraries and structure-based drug design has helped in the discovery of new selective kinase inhibitors. Since deregulation of the JNK pathway is involved in the pathogenesis of several diseases, a large variety of small-molecule inhibitors of JNK have been generated but only a limited number of them has been evaluated in clinical trials [202]. Most of the chemical entities targeting JNKs are type I inhibitors that are defined by their ability to compete with and displace ATP within the ATP-binding site. These inhibitors form non-covalent interactions with the surrounding residues of the ATP binding site and can exert their activity regardless of the conformation adopted by the DFG motif in the activation loop. Despite the fact that this region shares significant amino acid sequence homology and a conserved core structure among kinases, there are still minor residue differences in and near the ATP binding pocket that enable a certain degree of selectivity [248], [249].

A pharmacophore model developed by Traxler *et al*. to describe the ATP binding site has been successfully used for rational design of type I inhibitors [250]. This model defines five distinct subsites within the ATP binding site, which are characterized by different chemical environments and local sequence differences (Figure 1.12):

1) **Adenine binding region.** All ATP-competitive inhibitors bind in this hydrophobic area and interact with the backbone amino acids of the hinge region via hydrogen bonds, mimicking ATP binding.

2) **Ribose binding pocket**. This region has a hydrophilic character in most kinases and it can be exploited to accommodate hydrophobic substituents. As it is not conserved among kinases, it can be targeted to achieve selectivity and affinity.

3) **Phosphate region**. This hydrophilic region has high solvent exposure and is generally not addressed to increase inhibitor binding affinity.

4) **Hydrophobic region I (HR1)**. Also known as hydrophobic back pocket or selectivity pocket, this region is not involved in ATP binding. A specific residue of the N-terminal domain, termed gatekeeper, controls the access to this region. Since the gatekeeper and

other residues surrounding this pocket are not highly conserved among kinases, this area can be targeted to improve inhibitor selectivity and binding affinity.

5) **Hydrophobic pocket II (HR2)**. This solvent-exposed area is also not address by ATP and can be exploited to gain binding affinity and selectivity since its amino acid composition is not conserved.
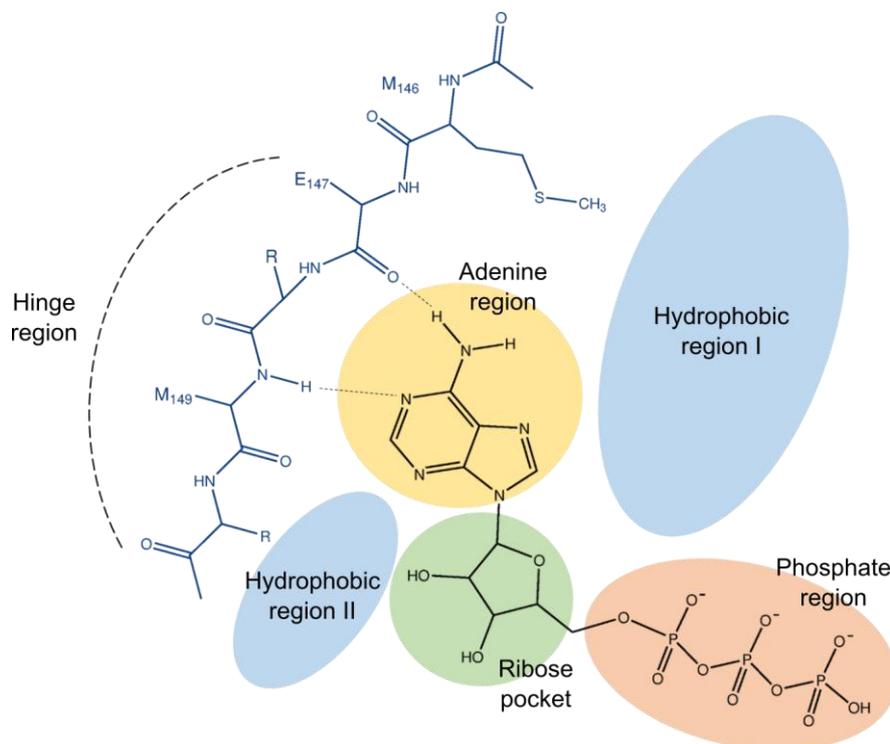
Figure 1.12 Representation of Traxler's pharmacore model used for the rational design of type I inhibitors. The ATP binding site of JNK3 is illustrated with the five distinct subsites located within this pocket highlighted in different colors. The protein backbone is displayed in blue and the compound in black. The only side chain shown is that of the gatekeeper residue of JNK3, M146. This residue alters the features of the hydrophobic region I, which can be addressed to achieve inhibitor selectivity or binding affinity. Figure adapted from [250].

The success of type I inhibitors demonstrates that, despite the highly conserved ATP binding site, it is possible to optimize inhibitor selectivity towards specific kinases. X-ray crystallography together with homology modeling enabled the characterization of structural differences within the vicinity of the ATP-binding site between JNK1, JNK2 and JNK3 and other MAPKs. The gatekeeper residue (corresponding to a methionine either in position 108 in JNK1 and JNK2 or in position 146 in JNK3) restricts access to the selectivity pocket [236], [241]. Effective strategies for the design of selective type I inhibitors are based on differences in size, shape and polarity of the gatekeeper residue and can be used to increase selectivity and binding affinity. Unfortunately, these strategies are not effective to obtain selectivity among the JNK family itself, since these enzymes share a very high sequence identity. Therefore, the design of isoform-selective type I JNK inhibitors remains a

significant challenge [251].

Compound SP600125 (from Celgene) was the first potent ATP-competitive kinase inhibitor possessing specificity for JNKs [252]. Although later shown to lack specificity by inhibiting all three JNK isoforms and also other kinases, a number of *in vivo* experiments using SP600125 have demonstrated the potential of directly inhibiting JNK for treatment of autoimmune, anti-inflammatory and neurodegenerative diseases (Figure 1.13). This compound still serves as a common reference in biological JNK assays [7], [253].

	IC$_{50}$ [µM]
JNK1	0.04
JNK2	0.04
JNK3	0.09
p38α MAPK	>10
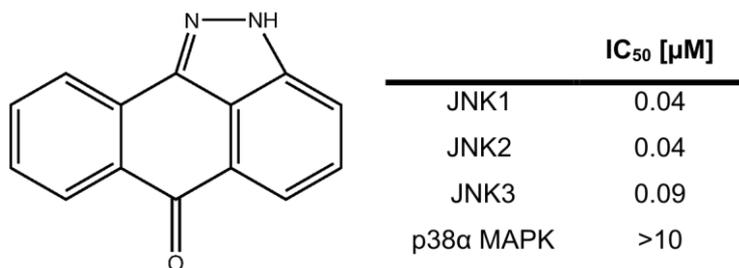
Figure 1.13 Structure and biological activity of compound SP600125 [252]. The potency of a kinase inhibitor is typically expressed as the half maximal inhibitory concentration (IC$_{50}$). This parameter determines the concentration of the drug candidate at which 50% of the kinase activity is inhibited [255].

A different approach to inhibit JNK activity is the use of peptide inhibitors that target the substrate-binding site or regulatory protein sites of the kinase, as for example the JIP1-derived peptides [254]. Targeting other sites on protein kinases than the ATP binding pocket avoids the possible problem of competing with high cellular ATP levels (1-5 mM) and also shows the highest degree of selectivity by exploiting sites that are unique to a particular kinase. However, a rational approach can hardly be used for the design and optimization of these type of inhibitors. Although the vast majority of efforts in the identification of inhibitors for the JNK pathway have focused on the JNK proteins themselves, JNK activity could also be suppressed by inhibiting the upstream MAPKKs, MKK4 and MKK7. Knowledge and control of inhibitor selectivity is important since off-target kinase inhibition potentially increases the risk of unwanted effects and toxicity [201], [255], [256].

1.5.6 Inhibition of JNK3

JNK3, predominantly found in neuronal cells, is a widely studied target in drug discovery to treat a variety of neurological disorders such as Alzheimer's disease, Parkinson's disease, Huntington's disease, multiple sclerosis and ischemia/reperfusion injury [257]. JNK3 activity is usually low in the brain but increases when irregularities in the cerebral metabolism occur. In the past years, significant progress has been made in the development of small-molecule inhibitors of JNKs, however achieving JNK3 selectivity is still a challenge due to the high

similarity among the JNK isoforms. Besides, compounds targeting CNS disorders have to overcome the additional challenge of penetrating the blood-brain barrier [6]. The first isoform-selective JNK3 inhibitor was reported by Zheng *et al.* [258]. This aminopyrazole-based inhibitor exhibited a high potency and selectivity against JNK3 with an IC_{50} value in the subnanomolar range and showed an isoform selectivity of >500/200-fold over JNK1/JNK2 and >20000-fold against p38α MAPK. Despite this success, isoform selective JNK1 and JNK2 inhibitors are still unavailable and there are no approved drug targets against any JNK to date [7].

2 Objectives

The work presented here is focused on three topics, namely structural and functional features of three proteins in their respective biological context. These are human properdin, the positive regulator of the alternative complement pathway, heparinase I from *Pedobacter heparinus*, a bacterial enzyme selective for heparin depolymerization and c-Jun N-terminal kinase 3, a human MAP kinase that is expressed primarily in the brain and that is implicated in several neurodegenerative disorders. The individual aims are divided in three different sections corresponding to each project.

1. The role of properdin in the activation of the alternative complement pathway

Properdin, the only positive regulator of the alternative complement pathway, has been a focus of controversy within the scientific community since its discovery by Louis Pillemer in 1954. In the last decade new evidence accumulated for Pillemer's hypothesis that the positive AP regulator may act as pattern-recognition molecule and initiate AP, apart from its widely accepted role in the stabilization of AP convertases. The three-dimensional structure of properdin has not been determined to atomic resolution to date and its role in glycan recognition and AP initiation remains unclear [2], [3]. The aim of this project consisted in the application of structural biology techniques, namely X-ray crystallography, EM, STD-NMR and glycan microarrays to study this protein at a molecular and atomic level and identify potential PAMPs that could activate the complement AP. To unravel the role of properdin as initiator of C3 convertase formation and AP activation, the following aims were pursued:

- Develop and establish the production of recombinant human properdin in bacterial or mammalian cells and analyze its stability, folding and oligomerization state.
- Evaluate the biological activity of recombinantly produced properdin in comparison to commercially available properdin via the standard hemolytic complement assay using red blood cells (RBCs).
- Elucidate the glycan specificity of properdin by identifying glycans that may activate the complement AP via direct recruitment of properdin.
- Determine the crystal structure of properdin to study the involvement of its domains in glycan binding at an atomic level.
- Compare the glycan specificity of properdin to the specificity of Factor H, the negative regulator of AP and as such a properdin antagonist.

2. Production of recombinant heparinase I from *Pedobacter heparinus* for structure-based protein engineering

GAGs are heterogeneous polysaccharides that mediate important biological processes through interaction with proteins. Due to their high chemical and structural complexity, deciphering the precise GAG sequences that are responsible for protein interactions and biological events remains a challenge since defined and pure GAG oligosaccharides are barely commercially available. Enzymatic depolymerization of GAGs by polysaccharide lyases, such as heparinase I, followed by gel filtration and anion exchange chromatography is a common strategy to obtain size-defined or chemical homogeneous oligosaccharides. However, large amounts of heparinase I are required if GAG oligosaccharides are to be used in structural and functional studies. Besides, the most relevant oligosaccharides have a higher degree of polymerization than the main disaccharide product obtained via enzymatic cleavage. Although heparinases from *P. heparinus* are the most studied, no three-dimensional structure of its heparinase I has been reported to date [146], [174]. To develop a structure-based protein engineering approach to rationally design a novel enzyme capable of producing larger amounts of longer heparin-derived oligosaccharides that are necessary to identify relevant interactions of proteins, heparinase I from *P. heparinus* was investigated with the following aims:

- Establish the production of recombinant heparinase I from *P. heparinus* in order to establish GAG production via enzymatic depolymerization for structural and functional studies of GAG-binding proteins.
- Determine crystal structures of heparinase I without substrate and in complex with heparin oligosaccharides to elucidate the catalytic mechanism of this enzyme.
- Develop a structure-based engineering approach to introduce site-specific modifications in heparinase I that result in a new product profile distribution during enzymatic depolymerization yielding higher amounts of longer oligosaccharides.

3. Structural-based optimization of selective type I c-Jun N-terminal kinase 3 inhibitors

Inhibition of JNK3 has become a promising strategy for treatment of several pathological states, including neurodegenerative disorders such as Alzheimer's disease and Parkinson's disease. In the past years, the number of JNK3 inhibitors and publications concerning this topic significantly increased. However, it is still challenging to achieve selectivity within the kinome, and especially between the different JNK isoforms. Up to date, there are no clinically approved inhibitors targeting JNK3, which highlights the need for new drug candidates and

further optimization [6], [7]. The use of X-ray crystallography for structure-guided optimization of lead compounds is nowadays essential in drug discovery and development. Three-dimensional structures of JNK3 in complex with inhibitors accelerate inhibitor design and help to address issues such as selectivity, binding affinity and pharmacokinetics. In collaboration with Dr. Francesco Ansideri and Prof. Dr. Pierre Koch from the Institute of Pharmaceutical Sciences from the University of Tübingen, new selective JNK3 inhibitors were investigated with the following aims:

- Elucidate the binding mode of two pyridinylimidazole-based compounds that act as selective ATP-competitive JNK3 inhibitors.
- Confirm the possibility of a flip of the imidazole core of the compound class depending on its substitution pattern.
- Gain insights into the role played by the different substituents around the pyridinylimidazole scaffold, namely the *S*-methyl group, in inhibitory activity and selectivity.
- Devise further strategies from the three-dimensional structures of JNK3-inhibitor determined by X-ray crystallography to further increase binding affinity and selectivity of inhibitors.

3 Materials and Methods

3.1 General

All standard chemicals and reagents used in this work were of analytical grade and obtained from Calbiochem, Merck, Roth and Sigma-Aldrich. Buffers and solutions were prepared with MilliQ water and stored at room temperature or at 4°C. Buffers were sterile filtered (0.22 µm) and degassed at 4°C. Reducing agents such as Dithiothreitol (DTT), β-mercaptoethanol and Tris (2-carboxyethyl) phosphine pH 7 (TCEP) were freshly added to the buffers before usage.

3.1.1 Bacterial strains

Two *E. coli* strains were used in this work (Table 3.1). *E. coli* strain XL10 gold (Stratagene) was used for cloning and plasmid amplification and *E. coli* BL21 (DE3) strain (Novagen) was used for recombinant protein production.

Table 3.1 Bacterial strains used in this work.

E. coli strain	Genotype
XL10 gold	$Tet_R\Delta$(*mcr*A)183Δ(*mcr*CB-*hsd*SMR-*mrr*)173 *end*A1 *sup*E44 *thi*-1 *rec*A1 *gyr*A96 *rel*A1 lac Hte [F´ proAB *lacl*qZΔM15 Tn10 (Tet_R) (Kan_R) Amy]
BL21 (DE3)	F- *ompT hsdS*B(rB- mB-) *gal dcm* (DE3)

3.1.2 Glycerol stock preparation

E. coli culture glycerol stocks were prepared by mixing 700 µL of bacterial overnight culture with 300 µL of a 50% (v/v) sterile glycerol solution. Cultures were transferred to cryovials, flash frozen with liquid nitrogen and stored at -80°C.

3.1.3 DNA isolation

Plasmid DNA from 10 mL *E. coli* cultures was isolated using a miniprep kit (Promega). To isolate plasmid DNA from 0.5-2 L of *E. coli* cultures a maxiprep kit (Sigma-Aldrich) was used. Both kits were used according to manufacturers' instructions.

3.1.4 DNA concentration determination

DNA concentration and purity were determined via UV absorption at 260, 230 and 280 nm using a NanoDrop ND-1000 spectrophotometer (PEQLAB).

3.1.5 Transformation of competent bacterial cells

Transformation of chemically competent cells was carried out by addition of 50-100 ng of target DNA to 50 µL competent bacterial cells and incubation on ice for 30 min. To induce transformation, cells were exposed to a heat shock of 42°C for 20-30 s followed by incubation on ice for 2 min. After addition of pre-warmed (37°C) 900 µL of antibiotic-free Luria broth (LB) or Super optimal broth with catabolite repression (SOC) medium (Table 3.2), cells were allowed to grow for 1 h at 37°C and 400 rpm. Transformed cells were then plated on LB-Agar dishes (Table 3.2) supplemented with the respective antibiotics and incubated overnight at 37°C.

Table 3.2 Media composition used for *E. coli* cell cultivation.

LB	LB-agar	SOC
1% (w/v) tryptone	1% (w/v) tryptone	0.5% (w/v) yeast extract
0.5% (w/v) yeast extract	0.5% (w/v) yeast extract	10 mM NaCl
1% (w/v) NaCl	1% (w/v) NaCl	2.5 mM KCl
	6% (w/v) agar	10 mM $MgCl_2$
		10 mM $MgSO_4$
		20 mM glucose

3.1.6 Sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE)

SDS-PAGE was used to analyze protein expression, purity and stability by separating charged molecules according to their molecular weight in an electrical field. Discontinuous acrylamide gels consisting of a stacking gel and a 10, 12 or 15% total acrylamide resolving gel were prepared, depending on the molecular weight of the target protein (Table 3.3). Protein samples were mixed with 4x SDS-PAGE loading buffer (Table 3.4), incubated at 95°C for 5 min and centrifuged at 14 000 rpm for 5 min before loading onto the gel. PageRuler unstained protein ladder (Thermo Fisher Scientific) was used as size standard and electrophoresis was carried out at 50 mA for 45-60 min. Electrophoresis buffer was prepared from a 10x stock (Roth) by dilution with MilliQ water. Gels were stained either with Coomassie staining solution (Table 3.4) or with Instant Blue solution (Expedeon) using an orbital shaker.

Table 3.3 Composition of 4x SDS-PAGE gels.

Gel composition	Resolving gel			Stacking gel
	10%	12%	15%	
30% Acrylamide-Bisacrylamide	5 mL	4.5 mL	7.5 mL	1.3 mL
1.5 M Tris pH 8.8	3.75 mL	3.75 mL	3.75 mL	-
1.5 M Tris pH 6.8	-	-	-	2.5 mL
10% (w/v) SDS	150 µL	150 µL	150 µL	100 µL
H_2O	6 mL	5 mL	3.5 mL	6.1 mL
10% (w/v) APS[1]	150 µL	150 µL	150 µL	100 µL
TEMED[2]	7.5 µL	7.5 µL	7.5 µL	10 µL

[1] Ammonium persulfate

[2] Tetramethylenediamine

Table 3.4 Composition of SDS-PAGE loading buffer and Coomassie staining solution.

4x SDS-PAGE loading buffer	Coomassie staining solution
Glycerol (20 mL)	Coomassie Brilliant Blue G250 (0.25 g)
1 M Tris pH 6.8 (20 mL)	Ethanol (100 mL)
10% (w/v) SDS (10 mL)	H_2O (900 mL)
0.5 M EDTA[1] pH 8 (1.63 mL)	Concentrated HCl (2.5 mL)
β-mercaptoethanol (4 mL)	Stir overnight at room temperature
Bromophenol blue (20 mg)	
H_2O (13 mL)	

[1] Ethylenediaminetetraacetic acid

3.1.7 Western blot

Western blot was performed to confirm the presence of properdin and heparinase proteins in cell culture supernatant or after protein purification. Buffers and solutions used for western blots were prepared according to Table 3.5. Antibodies specific for the different target proteins are listed in Table 3.6. Protein samples were first subjected to SDS-PAGE where a prestained protein ladder (Thermo Fisher Scientific) was used. The gel was equilibrated in western blot transfer buffer for 15 minutes while a polyvinylidene fluoride (PVDF) membrane (0.45 µm pore size, Roth) was equilibrated for 10 minutes in methanol, 5 minutes in water, and finally another 10 minutes in transfer buffer. Whatman papers (VWR International) were briefly moistened with transfer buffer and the blot was assembled as follows from top to bottom: 4 Whatman papers, PVDF membrane, SDS-PAGE gel and 4 Whatman papers. Protein bands were transferred from the gel to the PVDF membrane by applying a current of 20 V for 1h. Subsequently, the membrane was incubated for 1 h using milk solution at RT,

rinsed three times in Tris-buffered saline solution containing 0.5% (v/v) Tween-20 (TBS-T) and incubated overnight at 4°C with the respective primary antibody. On the following day, the membrane was rinsed again with TBS-T buffer and incubated for 1 h at RT with the respective secondary antibody. The membrane was rinsed once more with TSB-T and developed using the Amersham ECL detection reagents following the recommended protocol (GE Healthcare). Chemiluminescent detection was conducted with a ChemicDoc MP imaging system (Bio-Rad) and visualized using the Image Lab Touch Software (Bio-Rad). Colorimetric detection was used to visualize the prestained protein ladder and composite figures were generated by overlapping the results from both chemiluminescent and colorimetric detections.

Table 3.5 Composition of buffers and solutions used for western blotting.

Transfer buffer	TBS-T buffer	Milk solution
Trizma base (3.025 g)	200 mM Tris pH 7.5	5% (w/v) milk powder
Glycine (14.4 g)	1.5 M NaCl	TBS-T buffer
MeOH (200 mL)	0.5% (v/v) Tween-20	Freshly prepared
H_2O (800 mL)	Stored at RT	
Stored at 4°C		

Table 3.6 Primary and secondary antibodies used for western blotting. Dilutions performed in milk solution prepared in TSB-T as listed in Table 3.5 are indicated for each antibody.

Primary antibody	Dilution	Secondary antibody	Dilution
Anti-c-myc (9E10) (Santa Cruz Biotech)	1:1000	Anti-mouse HRP conjugate (Novagen)	1:5000
Anti-His tag (Merck)	1:5000		
Anti-mCherry (Thermo Fisher Scientific)	1:3000	Anti-rabbit HRP conjugate (Jackson Immuno Research)	1:10 000
Anti-Properdin (Abcam)	1:2000		

3.1.8 Protein concentration determination

Protein concentration was determined by measuring UV absorption at 280 nm using a NanoDrop ND-100 spectrophotometer (PEQLAB). The theoretical extinction coefficient (ε) at 280 nm and the molecular weight (MW) of each protein were predicted using the ExPASy ProtParam computational tool (Table 9.1) and were used to calculate the protein concentration in mg/mL units [259].

3.1.9 Analytical size exclusion chromatography (SEC)

Analytical SEC was used to analyze the oligomerization state, homogeneity and purity of protein samples. A precision column (PC), as listed in Table 3.7, was mounted on an ÄKTA Ettan LC system (GE Healthcare) and equilibrated with the respective SEC buffer (indicated in the following sections). 30 µL of protein sample were injected via a sample loop previously rinsed with SEC buffer, the protein was eluted with an isocratic flow of 0.04-006 mL/min and fractions of 50 µL were collected in a 96-well plate. Molecular weights were calculated from the elution volumes of the peaks observed in the resulting chromatogram based on standard equilibration curves obtained with the GE Healthcare SEC low molecular weight (LMW) and high molecular weight (HMW) calibration kits.

Table 3.7 Columns used for high-resolution SEC. The respective molecular separation range is indicated in Daltons (Da). Columns were purchased from GE Healthcare.

Column	Separation range (Da)
Superdex Peptide PC 3.2/300	100 - 7 000 (peptides)
Superdex 75 PC 3.2/300	3000 - 70 000 (globular proteins)
Superdex 200 increase PC 3.2/300	10 000 - 600 000 (globular proteins)
Superose 6 PC 3.2/300	5000 - 5 000 000 (globular proteins)

3.1.10 Circular dichroism (CD) spectroscopy

CD spectroscopy is a widely used technique to predict the secondary structure content of proteins and monitor their folding in different solution conditions, e.g. various combinations of pH, buffer and salts. Optically active molecules, such as proteins, exhibit differential absorption of circular polarized light in the far-UV (180 nm to 250 nm) mainly due to absorption by peptide bonds. The peptide backbone forms characteristic secondary structures such as α-helices, β-sheets, turns, and disordered sections, which exhibit distinctive CD spectra in the far-UV [260], [261]. CD spectra of protein and respective buffer samples as blanks were recorded on a J-720 spectropolarimeter (JASCO). Measurements were carried out at room temperature in a 0.1 cm path length quartz cuvette (Hellma Analytics) using a protein concentration of 0.4 mg/mL. Spectra were measured in a 195 nm to 250 nm wavelength range with an increment of 0.2 and bandwidth of 1 nm. Each spectrum reported in degrees of ellipticity (θ) was baseline corrected and represents spectra from ten individual scans. The data were converted to mean residue weight ellipticity (MRE) using equation (3.1).

$$[\theta]_{MRW,\lambda} = \frac{MRW\ \theta_\lambda}{10dc} \tag{3.1}$$

The mean residue weight of a protein (MRW) is the molecular weight of the protein (in Da) divided by the number of peptide bonds (number of amino acids (N)-1). θ corresponds to the observed ellipticity at a given wavelength λ, d is the path length of the cuvette (cm) and c is the molar concentration of the measured sample (mol/L) [262]. Secondary structure analysis was carried out with the algorithm Beta Structure Selection (BeStSel), which assigns proportions of individual secondary structure elements to polypeptide and protein CD spectra [263].

3.1.11 Differential scanning fluorimetry (DSF)

DSF monitors the thermal unfolding of a protein in the presence of a fluorescent dye and is typically performed using a real-time PCR instrument. The fluorescence intensity of the dye increases in a non-polar environment, such as the hydrophobic core of a protein that becomes unfolded and exposed upon heating. Fluorescence is plotted as a function of temperature and the unfolding or melting temperature (T_m) of the protein can be derived from the inflection point of the resulting melting curve. DSF is mostly use to screen parameters that influence protein thermal stability such as protein mutations, ligand binding and buffer formulations (pH, salts and additives) [264], [265]. In this work, DSF was performed using a LightCycler 480 Real Time PCR (Roche) and the SYPRO ORANGE dye (Invitrogen) or a StepOnePlus Real Time PCR and the Protein Thermal Shift dye (Thermo Fisher Scientific). DSF was performed in a 96-well format and samples were heated starting from either 4°C or 20°C to 95°C using a heating rate of 0.02°C/s or 1.6°C/s. SYPRO ORANGE dye was excited at 498 nm and the emission wavelength of 610 nm served as a readout. In the case of Protein Thermal Shift dye excitation was at 580 nm and the readout at 623 nm. The total reaction volume was 20 µL and final protein concentrations varied from 0.15 mg/mL to 0.4 m/mL with final dye concentrations of 5x for SYPRO ORANGE and 8x for Protein Thermal Shift dye. Protein samples were either analyzed in individually prepared buffer conditions or after being diluted with solutions from commercial buffers screens (JBScreen Thermofluor FUNDAMENT, Jena Bioscience and Solubility & Stability Screen 2, Hampton Research). These screens allow the identification of protein-stabilizing buffer conditions, which are essential for protein purification, characterization and crystallization.

3.2 Project I

3.2.1 Plasmids

Plasmid design and site-directed mutagenesis for human properdin production, planned using SnapGene or SnapGene Viewer software (GSL Biotech), were based on a publication by Pedersen *et al.* [71]. The gene sequence was codon optimized for expression in mammalian cells, synthesized by BioCat and subcloned using HindIII and BamHI restriction sites into the pCEP4 episomal mammalian expression vector (Invitrogen). This vector uses the cytomegalovirus (CMV) immediate early enhancer/promoter for high-level expression of recombinant protein. Human properdin constructs, listed in Table 3.8, were designed with the endogenous Kozak sequence and signal peptide. In one of the expression constructs, properdin was fused C-terminally to mCherry after removal of mCherry's endogenous start methionine. The mCherry fluorescent protein is frequently used as a reporter gene in the expression of cytosolic proteins and was used to detect protein expression and secretion via flow cytometry, fluorescence spectroscopy and western blotting. In a further expression construct, mCherry was replaced by the c-myc epitope, solely used for detection via western blotting. A poly-histidine tag (His_{10}-tag) located at the C-terminal end of both constructs was used for purification via Nickel affinity chromatography and a Tobacco Etch Virus (TEV) site was inserted before the tag sequences, including mCherry, to allow for proteolytic cleavage.

Table 3.8 Expression plasmids of human properdin. Protein sequence, cleavage site and fusion tags are indicated for each construct. The detailed protein sequence is shown in section 9.1.

Plasmid name	Protein sequence	Specifications
Properdin_mCherry	Residues 1-469[1]	C-terminal TEV cleavage site, mCherry protein and His_{10}-tag
Properdin_c-myc	Residues 1-469	C-terminal TEV cleavage site, c-myc epitope and His_{10}-tag
Properdin_E244K	Mutation E244K	C-terminal TEV cleavage site, mCherry protein and His_{10}-tag

[1]UNIPROT entry P27918

3.2.2 Site-directed mutagenesis

Properdin_mCherry substitution mutant E244K was generated by site-directed mutagenesis. Forward (5'-GGCCTGGCTTACAAGCAGAGGAGATGCACAGGCCTGC-3') and reverse (5'-CATCTCCTCTGCTTGTAAGCCAGGCCGGGGCAAGGCTTTC-3') primers were obtained from biomers.net. Due to the high melting temperature of the primers ($T_m > 69°C$) a 2-step polymerase chain reaction (PCR) was performed in 32 cycles to amplify the DNA. A total

reaction volume of 50 μL was prepared containing 25 ng of template DNA, 10 mM deoxyribonucleotide triphosphate (dNTPs), 25 μM primers and 0.5 μL ExactRun-DNA polymerase (Genaxxon bioscience). The time for the extension step at 72°C was calculated based on the length of the template DNA (15 s per 1 kb). PCR products were incubated with 1 μL DpnI (Thermo Fischer Scientific) for 2 h at 37°C to digest template DNA followed by purification using the Wizard SV Gel and PCR Clean-up System (Promega) according to the manufacturer's recommendation. The obtained DNA samples were concentrated using a vacuum centrifuge (Heto) and transformed in *E. coli* XL10 gold cells. Single colonies were selected, DNA was amplified and the mutation was verified by Sanger sequencing (Microsynth Seqlab).

3.2.3 Mammalian cell culture

FreeStyle 293-F cells purchased from Thermo Fisher Scientific are derived from 293 cells, a permanent cell line established from primary embryonal human kidney transformed with sheared human adenovirus type 5 DNA and adapted for serum-free suspension culture. Cells were cultured in suspension in FreeStyle 293 medium (Gibco) without serum and were maintained at 37°C on an orbital shaker platform (Celltron, Infors HT) rotating at 125 rpm in a humidified incubator containing 8% CO_2. Cell viability and density was determined according to the trypan blue dye exclusion method (Trypan blue stain 0.4%, Gibco) using a Neubauer chamber [266]. Cells were thawed, cultured and frozen according to the manufacturer's recommendations (Invitrogen). Disposable, sterile Erlenmeyer flasks of 125, 250 or 500 mL with vented cap (Corning) were used for cell cultivation.

3.2.4 Transient transfection of expression constructs

The protocol described here is representative for a transfection in one flask of 500 mL containing a cell suspension of 160 mL. This protocol was improved by reducing the cell density on the day of transfection and by using a ratio of 1:4 of transfection reagent to DNA for transient transfection. Approximately 24 h before transfection, cells were passaged at 0.6×10^6 or 0.7×10^6 cells/mL. On the day of transfection, the cell density was about 1.5×10^6 cells/mL and cells were diluted with pre-warmed medium at 37°C to 1×10^6 cells/mL. Viability of cells was usually higher than 85% and cells were transiently transfected using final concentrations of 0.5 mg/L plasmid DNA and 2 mg/L polyethylenimine (PEI-25 kDa, Polysciences). Mixtures containing plasmid DNA and the transfection reagent PEI diluted into OptiPro SFM medium (Gibco) were prepared separately with a total volume of 2.4 mL and incubated for 5 min at room temperature. The two mixtures were subsequently combined and incubated for 15 min at room temperature to allow complex formation. The DNA-lipid mixture was then slowly added to the cell suspension while swirling the flask.

Transfected cells were incubated at 37°C, 8% CO_2 on an orbital shaker platform rotating at 125 rpm. After 7 days of incubation, medium containing the secreted protein was harvested and cell debris was removed by centrifugation at 4 000 x g for 10 min. Harvested supernatant was either frozen at -20°C or directly used for purification of target proteins. Volumes of cell suspension and DNA-lipid mixture were proportionally adapted when using flaks of 125 mL or 250 mL.

3.2.5 Immunoprecipitation of His-tagged proteins

To detect secreted properdin via western blotting, His-tagged proteins present in the harvested supernatant were enriched by immunoprecipitation. Immunoaffinity resin coated with anti-His monoclonal antibody (GenScript) was resuspended and 100 µL of the slurry was transferred to a 1.5 mL tube. 500 µL of TBS buffer were added to the resin and the mixture was centrifuged at 10 000 rpm for 60 s at 4°C. Afterwards, buffer was discarded and this washing step was repeated twice. 100 µL of concentrated cell supernatant, previously concentrated 10-fold with a Vivaspin 20 centrifugal concentrator with a molecular weight cutoff (MWCO) of 50 kDa (Sartorius), were added to the tube containing the washed resin and the solution was mixed gently. Subsequently, the anti-His affinity resin was incubated with the supernatant for 4 h at 4°C while rotating the tube on a shaker platform to allow binding of the histidine residues of the target protein to the resin. After incubation, the tube was centrifuged once more at 10 000 rpm for 60 s at 4°C, supernatant was discarded and the washing step was repeated three times. For visualization of immunoprecipitated proteins, resin with the bound target protein was mixed with 30 µL of SDS-PAGE loading buffer and the mixture was heated for 10 min at 95°C followed by centrifugation at 10 000 rpm for 60 s at 4°C to spin down the resin. Supernatant was collected and analyzed by western blotting.

3.2.6 Fluorescence measurement

Expression and secretion of properdin fused to mCherry was followed within a period of 7 days post-transfection by measuring the mCherry fluorescence as previously reported by Duellman *et al.* [267]. 50 µL of transfected cell suspension were mixed with the same amount of water and centrifuged at 15 000 rpm for 1 min. 90 µL of the mixture were transferred to a black 96-well F-bottom plate (Greiner Bio-One) and fluorescence intensity was recorded with an infinite M200 microplate reader (Tecan). mCherry fluorescence was measured using an excitation wavelength/bandwidth of 590/9 nm and an emission wavelength/bandwidth of 645/20 nm. Measurements were performed in duplicate and non-transfected cells were used to determine the background signal.

3.2.7 Fluorescence-activated cell sorting (FACS) analysis

Flow cytometry can be used to assess and optimize transfection efficiency of transient gene expression at the cell level. In this work, the fluorescence of mCherry fused to properdin was measured to quantify target protein expression. Samples of 500 µL of cell suspension were collected post-transfection and centrifuged at 300 x g for 10 min using a table centrifuge 5415D (Eppendorf). Medium was discarded and transfected cells were washed with 500 µL of phosphate-buffered saline (PBS) and the previous centrifugation step was repeated. After removing the PBS solution, cell pellet was resuspended in 150 µL PBS and transfection efficiency was analyzed by flow cytometry using CytoFlex S (Beckman Coulter). To measure fluorescence-expressing cells, the flow cytometer was set to 10 000 events for the different experiments. Non-transfected cells were used as a negative control for appropriate gating and to measure background fluorescence. Data were analyzed using the FACSDiva software (Beckman Coulter) and dot plots of forward scatter (FSC) versus side scatter (SSC) of mCherry-producing cells were generated. A gate was designated on each dot plot to include as many mCherry-positive cells as possible. In addition, a histogram plot of cell counts versus fluorescent signal was generated using the CytoFlex channel ECD. The fluorescent dye was excited with a 561 nm laser and the emitted light was detected with a 610/20 band pass (BP) fluorescent channel. In this plot, the % of fluorescent cells (successfully transfected) within the total cell population was reported.

3.2.8 Ni^{2+} affinity chromatography

Ni^{2+} affinity chromatography was used as a first purification step for His-tagged proteins. Harvested supernatant containing target secreted protein (500 mL to 1L) was diluted in buffer containing 20 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole with the pH adjusted to 7.4. Diluted supernatant was loaded onto a 5 mL HisTrap excel column (GE Healthcare) at a flow rate of 1.5 mL/min using a peristaltic pump. After loading, the column was mounted on a ÄKTA purifier system (GE Healthcare) and washed with 25% washing buffer containing 20 mM sodium phosphate pH 7.4, 0.5 M NaCl, 10 mM imidazole. After removing unspecifically bound proteins, a stepwise gradient of elution buffer containing 500 mM Imidazole was applied to elute the target protein. A first step of 85% elution buffer was applied for 3 mL and the pump was paused for about 30 min to allow the dissociation of histidine residues from the nickel resin. A flow rate of 1 mL/min was used to elute properdin fused to mCherry or c-myc and 1 mL fractions were collected. A step of 100% elution buffer was subsequently applied to elute the remaining protein bound to the column. Collected fractions were analyzed by SDS-PAGE and fractions containing target protein were pooled and concentrated using a Vivaspin 20 centrifugal concentrator with a MWCO of 50 kDa (Sartorius).

3.2.9 Tag cleavage and reverse Ni^{2+} affinity chromatography

To remove the C-terminal His-tag as well as the mCherry or the c-myc epitope, concentrated protein fractions were incubated with TEV protease (1 µg of TEV /10 µg of target protein) previously diluted 1:10 in cleavage buffer containing 20 mM sodium phosphate pH 7.4, 500 mM NaCl and 0.5 mM Ethylenediaminetetraacetic acid (EDTA). The mixture containing His-tagged protein and His-tagged TEV protease was dialyzed overnight at 4°C against the cleavage buffer using a dialysis membrane with a MWCO of 6-8 kDa (Spectrum Labs). Dialyzed protein was filtered and applied onto a 1 mL HisTrap ExCel column (GE Healthcare) at a flow rate of 0.8 mL/min. Washing and elution buffers, used in the first Ni^{2+} affinity chromatography step, were used once again in this purification step to separate cleaved target protein. Cleaved properdin was collected in the flow-through, whereas mCherry or c-myc fused to His-tag as well as the TEV protease were eluted from the column with 100% buffer B. Fractions containing cleaved properdin were pooled, concentrated and dialyzed overnight at 4°C against SEC buffer containing 10 mM HEPES pH 7.3 and 150 mM NaCl. Protein stability and oligomerization were analyzed via analytical SEC using a Superose 6 PC 3.2/300 column. To further characterize properdin, DSF and CD measurements were performed using SEC buffer, as described in section 3.1.

3.2.10 Protein deglycosylation

Deglycosylation of properdin was attempted with the use of endoglycosidases Endo Hf and PNGase F (New England Biolabs) to release the N-linked oligosaccharides produced by the mammalian cells. Protein samples were separately incubated with an excess of enzyme to cleave off the N-linked oligosaccharides overnight at 4°C. Samples were analyzed by SDS-PAGE on the following day.

3.2.11 Isolation of sheep erythrocytes and hemolytic complement assay

To assess the biological activity of different properdin preparations, the complement activity hemolytic assay was performed using sheep erythrocytes (RBCs) as previously reported by Blaum *et al.* [101]. This assay was originally designed by Sánchez-Corral *et al.* to detect Factor H related complement regulatory defects in patients [268]. RBCs were obtained from defibrinated sheep blood (TCS Biosciences), whereas commercial properdin (Factor P), properdin-depleted serum and normal human serum (NHS) were obtained from Complement Technologies. 1 mL of sheep RBCs was transferred to a 15 mL falcon tube and diluted with 10 mL of 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin from pork skin (Fluka) and 5 mM EDTA. EDTA was used to chelate and remove Ca^{2+} ions from the cell suspension in

order to inhibit the Ca^{2+}-dependent classical pathway during hemolysis. The cell suspension was gently mixed and centrifuged at 500 x g for 10 min at 4°C. The supernatant and a layer of leukocytes were discarded and the procedure was repeated twice with the same buffer and subsequently three times with 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin, 5 mM $MgCl_2$ and 5 mM ethylene glycol-bis(β-aminoethyl ether)-N,N,N′,N′-tetraacetic acid (EGTA). EDTA was replaced by EGTA to selectively bind Ca^{2+} from serum but not Mg^{2+}, since this ion is a cofactor to complement Factor B. The sample was centrifuged at 1 000 x g. After discarding the supernatant, cells were stored at 4°C overnight and the washing step was repeated once more with 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin, 5 mM $MgCl_2$ and 5 mM EGTA to remove lysed cells before proceeding with the hemolytic assay. The amount of cells was determined by measuring the absorption at 412 nm (A_{412}). An A_{412} value of 0.87 at 1 cm path length corresponds to a cell concentration of 5 x 10^8 RBCs/mL (Kerry Pangburn, Complement Technologies, personal communication). Cells were treated with neuraminidase from *Clostridium perfringens* (New England Biolabs) to remove sialic acid. Approximately 6 x 10^9 cells were resuspended in a total reaction volume of 100 µL using buffer containing 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin and 500 units of neuraminidase and the reaction was incubated for 30 min at 37°C. Afterwards, 150 µL of 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin, 5 mM $MgCl_2$ and 5 mM EGTA were added and the cells were centrifuged at 500 x g for 1 min at 4°C. The supernatant was removed and the washing step was repeated twice. For the hemolytic assay, approximately 1.7 x 10^8 neuraminidase-treated cells and cells that were treated identically except for the omission of the neuraminidase treatment were used. Cells were mixed with 20 µL of NHS (intact, heat-inactivated for 30 min at 56°C or properdin-depleted) and buffer 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin, 5 mM $MgCl_2$ and 5 mM EGTA and/or properdin (10 µg/mL and 15 µg/mL) to obtain a total reaction volume of 40 µL. The reaction, prepared in a V-shaped bottom 96-well plate (Greiner Bio-One) was incubated for 20 min at 37°C. Hemolysis was stopped by addition of 150 µL cold buffer containing 20 mM HEPES pH 7.3, 145 mM NaCl and 5 mM EDTA, followed by centrifugation at 1500 x g for 5 min at 4°C. 150 µL of the supernatant was transferred to a F-bottom UV-star transparent 96-well plate (Greiner Bio-One) and the absorbance at 415 nm was measured using an infinite M200 microplate reader (Tecan). Results are expressed relative to the total osmolysis of RBCs in water. The experiment was repeated two times and a standard error was calculated for each sample.

3.2.12 Saturation transfer difference nuclear magnetic resonance (STD-NMR)

Ligand-based NMR spectroscopy techniques such as STD-NMR have been widely used for directly monitoring ligand binding to protein receptors in solution. STD-NMR can be used to study weak protein-ligand interactions (K_d in mM to µM range) via non-scalar magnetization

transfer from a large protein to smaller ligands, e.g. glycans. Since signals from small ligands are observed, the molecular weight of the protein is not a limiting factor in an STD experiment, unlike in protein-observed NMR techniques. Mixtures containing different ligands can be efficiently screened as long as resonances from the ligands do not overlap [269], [270]. The STD-NMR experiment is based on the nuclear Overhauser effect between the protein and the ligand in the case of complex formation with fast off-rates. In STD experiments, a spectrum obtained with selective saturation of protein proton resonances (on-resonance spectrum) is subtracted from a spectrum recorded without such selective protein saturation (off-resonance spectrum). In the resulting difference spectrum only proton resonances of the ligand that received saturation indirectly from the protein (saturation transfer) will remain. Non-bound ligands that are present in the same solution do not experience such saturation transfer, their signals are, therefore, of equal intensity in the on- and off-resonance spectra and will not appear in the resulting difference spectrum. For a ligand that binds to the receptor only resonances belonging to those protons of the ligand that are in close contact to the protein (up to 5 Å) receive saturation transfer and appear in the difference spectrum, with their relative intensities reflecting the positioning with respect to the protein. Based on this procedure the binding epitope on the ligand can be determined from STD experiments [271].

NMR spectra were recorded between 283 K and 287 K using 3 mm tubes (200 µL sample volume) on a Bruker AVIII-600 MHz spectrometer equipped with a room temperature probe head. Data were processed with TOPSPIN 3.0 (Bruker). Samples for STD-NMR spectra contained 1 mM of each of the different glycans (Table 3.9) and 15-15.7 µM of properdin fused to mCherry or properdin after tag removal. Before NMR experiments, proteins were buffer exchanged with 10 kDa MWCO Slide-A-Lyzer Mini devices (Thermo Fisher Scientific) to NMR buffer containing 20 mM potassium phosphate, pH 7.4, 150 mM NaCl in D_2O. Glycan samples of the same concentration but with no protein present were used as controls to verify that no direct excitation of the glycans occurred during the on-resonance irradiation step of the STD pulse program. Off- and on-resonance irradiation frequencies were set to −30 ppm and, depending on the glycan, to 0 ppm or 6.9 ppm, respectively. The irradiation power of the selective pulses was 57 Hz, the saturation time was 2 s, and the total relaxation delay was 3 s. A 50-ms continuous-wave spin-lock pulse with a strength of 3.2 kHz was used to suppress residual protein signals. A total number of 512 scans were recorded. Spectra were multiplied with a Gaussian window function prior to Fourier transformation.

Table 3.9 Glycans used for STD-NMR experiments.

Glycan	Sequences
Arixtra or Fondaparinux (Mylan)	GlcNS6S-GlcA-GlcNS3S6S-IdoA2S-GlcNS6S-OMe
TF antigen (Elicityl)	Galβ1-3GalNAc disaccharide
HepMer[1] M06 S03a (Iduron)	(GlcNS-GlcA)$_3$-*p*-nitrophenyl
HepMer[1] M09 S06a (Iduron)	(GlcA-GlcNS)$_2$-(GlcA-GlcNS6S)$_2$-GlcA-*p*-nitrophenyl
Hep dp4 (Iduron)	Heparin tetrasaccharide
Hep dp6 (Iduron)	Heparin hexasaccharide
Polysialic acid dp5[2]	α2,8-linked Neu5Ac dp5
Lactose (Elicityl)	Galβ1-4Glc
Asialo-GM1 (Elicityl)	Galβ1-3GalNAcβ1-4Galβ1-4Glc

[1]HS/Heparin oligomers of defined sequence

[2] Provided by Prof. Harald Neumann, University of Bonn

3.3 Project II

3.3.1 Plasmids

Plasmid design was initially performed based on a publication from Han *et al.,* reporting the first crystal structure of bacterial heparinase I from *B. thetaiotaomicron* [174]. All expression constructs were designed using the SnapGene or SnapGene Viewer software. Genes were synthesized either by Eurofins Genomics (Ebersberg) or BioCat (Heidelberg) and subcloned using NdeI and BamHI or NdeI and XhoI restriction sites into pET28a and pET21b vectors, respectively. All sequences were codon optimized for expression in *E. coli* (DE3) strains. With the exception of construct HepIv2, which contained part of the native signal peptide at the N-terminus, in constructs HepIv1 and HepIv3 the signal peptide was omitted (Table 3.10). A His$_6$-tag located at the N- or C-terminal end was used for purification via Ni^{2+}-affinity chromatography. Thrombin and human rhinovirus 3C protease (HRV3C) cleavage sites were inserted into HepIv1 and HepIv2 constructs, respectively.

Table 3.10 Plasmids used to express recombinant heparinase I from *P. heparinus*. The detailed protein sequence is shown in section 9.1

Plasmid name	Protein sequence	Vector	Antibiotic	Specifications
HepIv1	Residues 22-384[1]	pET28a	Kanamycin	N-terminal His$_6$-tag, thrombin cleavage site
HepIv2	Residues 12-384	pET28a	Kanamycin	N-terminal His$_6$-tag, HRV3C cleavage site
HepIv3	Residues 22-384	pET21b	Ampicillin	C-terminal His$_6$-tag

[1]UNIPROT entry Q05819

3.3.2 Expression of heparinase constructs in *E. coli* cells

For gene expression, overnight cultures containing 10-20 mL of LB medium supplemented with 50 µg/mL Kanamycin or 100 µg/mL Ampicillin were either inoculated with bacteria from glycerol stocks or a single bacteria colony from LB-Agar plates. Starter cultures were grown overnight at 37°C and 110 rpm for 12-16 h and added to larger expression cultures containing 1-2 L of LB medium supplemented with the respective antibiotics. Incubation at 37°C and 90 rpm was performed until an optical density at 600 nm (OD$_{600}$) of 0.5-0.7 was reached. Target gene expression was induced by addition of 0.1 mM to 1 mM of isopropyl-β-D-thiogalactopyranoside (IPTG) and overnight cultivation at 20°C and 90 rpm. The next morning, cells were harvested by centrifugation at 7 000 rpm and 4°C for 20 min (Sorvall RC 6+, rotor SLC-4000). Cell pellets were either directly used for purification or stored at -80°C.

3.3.3 Cell disruption

Cell pellets were resuspended in 10 mL of lysis buffer/g of wet cells and supplemented with one cOmplete, Mini, Ethylenediaminetetraacetic acid (EDTA)-free tablet (Roche) and 0.1 unit of Benzonase (Novagen) per approximately 50 mL of cell suspension to inhibit serine and cysteine proteases and degrade nucleic acids, respectively. Different buffers were used for cell disruption depending on the expressed construct (Table 3.11). Lysis was performed either by homogenizing the cells on an Avestin EmulsiFlex C3 instrument at a pressure of 700-800 bar for 20-30 min or by sonication using a Branson Digital Sonifier. Lysis of culture from a large-scale expression was conducted with an amplitude of 40% and an overall pulse time of 4-6 min (cycles of 0.5 s pulse on and 0.5 s pulse off) while cooling the cells in an ice bath. Cell debris were removed by centrifugation at 17 000 rpm and 4°C for 30 min (Sorvall RC 6+, rotor: SS-34). The supernatant was collected and sterile filtered with a sterile 0.22 µm nitrocellulose membrane filter (Millipore).

Table 3.11 Lysis buffer used for cell disruption.

HepIv1 construct	HepIv2/HepIv3 constructs
50 mM Tris-HCl pH 7.5/ pH 8	50 mM HEPES pH 8
500 mM NaCl	150 mM NaCl
10 mM imidazole	5 mM β-mercaptoethanol
	125 mM L-Arginine hydrochloride
	0.01% (w/v) Triton X-100

3.3.4 Ni^{2+} affinity chromatography

Ni^{2+} affinity chromatography was used as a first purification step for all expressed heparinase I constructs due to the presence of a His$_6$-tag at the N- or C-terminal end of the protein. Protein extracts were loaded onto a 5 mL HisTrap FF crude column (GE Healthcare) mounted either on a ÄKTA prime plus or ÄKTA purifier system (GE Healthcare) at a flow rate of 1 mL/min, and pre-equilibrated with washing buffer. The column was washed with 5-20% of elution buffer to remove unspecifically bound proteins. For the purification of HepIv2, an additional washing step was included with a buffer containing 1 M NaCl and 20% (v/v) glycerol to disrupt ionic and hydrophobic interactions formed with the target protein. His-tagged target protein was eluted from the column using linear or stepwise gradients of elution buffer (Table 3.12). Collected fractions were analyzed by SDS-PAGE and fractions containing target protein were pooled and concentrated using a Vivaspin 20 centrifugal concentrator with a MWCO of 10 kDa (Sartorius).

Table 3.12 Buffers used for the first purification step of heparinase I. Washing and elution buffers are indicated for the purification of each construct.

Buffer	HepIv1 construct	HepIv2/HepIv3 constructs
Washing	50 mM Tris-HCl pH 8 500 mM NaCl 10 mM imidazole	50 mM HEPES pH 8 150 mM NaCl 5 mM β-mercaptoethanol 20 mM imidazole
Elution	Washing buffer containing 500 mM imidazole	Washing buffer containing 500 mM imidazole

3.3.5 His-tag cleavage and reverse Ni^{2+} affinity chromatography of HepIv1

To remove the N-terminal His-tag from HepIv1 construct, bovine thrombin protease (Calbiochem) was added to the pooled fractions containing target protein (1 unit of thrombin/ 0.5 mg of His-tagged protein). The mixture was dialyzed against a buffer containing 50 mM Tris-HCl pH 8, 150 mM NaCl, 2.5 mM CaCl$_2$ and 1 mM DTT overnight at 4°C using a dialysis membrane with a MWCO of 6-8 kDa (Spectrum Labs). Dialyzed samples were filtered and applied onto a 5 mL HisTrap FF crude column connected to a 1mL HiTrap Benzamidine FF column (GE Healthcare) at a flow rate of 1 mL/min. Cleaved protein was collected in the flow-through fractions and dialyzed into a buffer containing 5 mM HEPES pH 7.5, 150 mM NaCl and 1 mM DTT. His-tagged protein as well as cleaved His-tag and thrombin protease remained bound to the HisTrap and HiTrap Benzamidine columns.

3.3.6 Preparative size exclusion chromatography

SEC was used as final purification step to separate molecules according to their size/oligomerization state and exchange the buffer of HepIv1. A Superdex 200 Increase 10/300 GL column (GE Healthcare) mounted on a BioLogic DuoFlow chromatography system (Bio Rad) was equilibrated with SEC buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl and 1 mM TCEP or a buffer containing 20 mM Super buffer (citric acid:HEPES:CHES 2:3:4) pH 4, 250 mM NaCl and 1 mM TCEP. 500 µL of filtered protein sample was injected via a sample loop previously rinsed with water and SEC buffer. Protein was eluted with an isocratic flow of 0.6 mL/min and collected fractions were analyzed by SDS-PAGE. Fractions containing target protein were pooled and concentrated using a Vivaspin 20 centrifugal concentrator with a MWCO of 10 kDa (Sartorius). Purified protein was either kept at 4°C or flash frozen and stored at -80°C.

3.3.7 Ion exchange chromatography (IEX)

Cation exchange chromatography was used as a second purification step for the HepIv2 and HepIv3 constructs since heparinase I has a theoretical isoelectric point above 9 and is, therefore, positively charged at the pH used for the purification (Table 9.1). Pooled and concentrated fractions from the Ni^{2+} affinity column were applied onto one or more gravity-flow PD-10 desalting columns (GE Healthcare), depending on the protein volume. Buffer was exchanged to IEX start buffer containing 50 mM HEPES pH 8, 20 mM NaCl and 5 mM β-mercaptoethanol to decrease the sample ionic strength prior to IEX. The recovered protein solution was filtered at 0.22 µm and applied with a flow rate of 0.8 mL/min onto a 1 mL HiTrap SP Sepharose FF (GE Healthcare), previously equilibrated with start buffer. After washing the column with start buffer, steps of 5 % and 10% of elution buffer containing 1 M NaCl were applied followed by a linear gradient up to 100% of elution buffer. Protein-containing fractions were analyzed by SDS-PAGE and pooled accordingly. The combined eluted fractions containing heparinase were further dialyzed against a buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl, 2 mM DTT and stored at 4°C or -80°C.

3.3.8 Crystallization and crystal testing

Initial crystallization trials were performed on 96-well sitting drop crystallization plates (Intelli plate, Art Robbins Instruments) with commercially available screens (from Qiagen, Molecular Dimensions and Hampton Research) and a reservoir volume of 50 µL or 100 µL. His-tagged or cleaved HepIv1 as well as His-tagged HepIv2 at final concentrations ranging from 0.7 mg/mL to 8 mg/mL were used. Crystal trial drops were set using an automated crystallization robot (Freedom evo, Tecan or Gryphon, Art Robbins Instruments) by pipetting 0.2-0.3 µL of protein solution and 0.2-0.3 µL of reservoir solution for each crystallization condition. Crystallization plates were sealed and incubated at 4°C or 20°C. To distinguish between salt and protein crystals and verify their diffraction quality, crystals were flash frozen in liquid nitrogen and tested with an in-house X-ray diffraction system equipped with a rotating copper anode X-ray generator (MicroMax-007HF, MSC Rigaku) and a mar345 dtb image plate detector (Marresearch) at CuKα-radiation (λ=1.5418 Å).

3.3.9 Heparinase activity assay

Heparinase activity was analyzed by measuring formation of the C4-C5 double bond at the non-reducing end of oligosaccharide products indicated by absorption at 232 nm. The experimental procedure used was based on the activity assay protocol recommended by R&D systems. The activity assay was performed using commercial heparinase I purchased from R&D systems and recombinantly produced HepIv1 (1.15 mg/mL) stored in buffer

containing 50 mM HEPES pH 7.5, 150 mM NaCl and 2 mM DTT. A stock solution of 20 mg/mL of sodium heparin substrate (Bioiberica) was initially prepared in water. 300 µL reactions were prepared in a F-bottom UV-star transparent 96-well plate (Greiner Bio-One) with heparinase I (0.5 µL, 5 µL or 10 µL) and 0.75 mg/mL of sodium heparin diluted in assay buffer containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 2 mM, 5 mM or 10 mM $CaCl_2$. A substrate blank prepared with heparin substrate and assay buffer was included to determine the background absorbance. All assays were performed at 35°C based on findings from Lohse and Linhardt, who determined this to be the optimal temperature for activity of heparinase I from *P. heparinus* [145]. The enzymatic depolymerization was monitored by UV absorption measurements at 232 nm with an infinite M200 microplate reader (Tecan) at regular intervals of 20 s. Reactions were stopped by incubation at 100°C for 3 min and the depolymerization solution was then filtered and analyzed via analytical SEC. Heparin-derived oligosaccharides were size-fractionated on a Superdex Peptide PC 3.2/300 column and eluted with assay buffer containing 10 mM $CaCl_2$. Separation of heparin-derived oligosaccharides was monitored by UV absorption at 232 nm.

3.4 Project III

3.4.1 Samples

Pure JNK3 protein and 4-methyl-5-(pyridine-4-yl) imidazole derivatives (compounds **38** and **44)** used for crystallization and nanoDSF experiments were provided by Prof. Dr. Frank M. Böckler and Dr. Francesco Ansideri, respectively, from the Institute of Pharmaceutical Sciences from the University of Tübingen. JNK3 was purified according to the experimental procedure described by Lange *et al.* [272] and stored in 50 mM HEPES pH 7, 100 mM NaCl, 2 mM $MgCl_2$, 10 mM β-mercaptoethanol buffer at -80°C. Solutions of compounds **38** and **44** were prepared in dimethyl sulfoxide (DMSO) to final concentrations of 100 mM and stored at -20°C.

3.4.2 Crystallization

Initial JNK3 crystals were obtained following the experimental procedure described by Lange *et al.* [272]. A mixture containing pure JNK3 in 50 mM HEPES pH 7, 100 mM NaCl, 2 mM $MgCl_2$, 10 mM β-mercaptoethanol buffer as well as 1 mM AMP-PCP (β,γ-methyleneadenosine 5′-triphosphate disodium salt (Sigma-Aldrich)), 0.4 mM Zwittergent (Calbiochem) and 10% (v/v) ethylene glycol was prepared and incubated on ice for 30 min. Final protein concentration was 10 mg/mL and a reservoir solution containing 0.1 M Bis-Tris pH 5.5, 0.2 M NaCl and 28-31% (v/v) PEG 3350 was used for crystallization. Drops were set up with 1 µL of protein mixture and 1 µL reservoir solution, and crystals were grown at 20°C using the sitting drop vapor diffusion method. To improve the initial crystal quality, lower protein concentrations and microseeding were used. Best diffracting crystals were obtained using a final JNK3 concentration of 5 and 2.5 mg/mL together with microseeding and a reservoir solution containing 29% (v/v) of PEG 3350. A seed stock for microseeding was prepared using initial crystals that were crushed in 50 µL reservoir solution by vortexing with a Polytetrafluoroethylene (PFTE) bead (Hampton Research). Serial tenfold dilutions of the seed stock were prepared with reservoir solution and 0.2 µL of seed solution were transferred to a freshly prepared crystallization drops. After one week, AMP-PCP containing crystals were incubated with compounds **38** and **44** for 36 h. The drop solution was gradually exchanged by reservoir solution supplemented with 10 mM of either compound **38** or **44** previously prepared in DMSO. For data collection, crystals incubated with compounds were cryoprotected by stepwise incubation in reservoir solution containing 10 mM of the respective compound and 15% (v/v) glycerol. For structural comparison, AMP-PCP containing crystals without compounds were   also analyzed. For cryoprotection, 15% (v/v) glycerol was added to the reservoir solution. Cryoprotected crystals were collected in nylon loops and flash-frozen in liquid nitrogen for synchrotron data collection.

3.4.3 Data collection and structure determination

Diffraction quality of protein crystals was evaluated using the in-house X-ray diffraction system. Full data sets were collected at the X06DA beamline at the Swiss Light Source (Paul Scherrer Institute, Villigen, Switzerland) at 100 K using a Pilatus 2M-F detector and a wavelength of 1 Å. Diffraction images were indexed, integrated and scaled using XDS [273]. Phase information was obtained by molecular replacement with MOLREP [274] using a search model derived from another JNK3-inhibitor complex structure (PDB ID 4X21) [272]. For structural refinement alternating model building and reciprocal refinement cycles were performed. Model building was carried out with Coot [275] and restrained reciprocal refinement including transition-libration-screw (TLS) parameterization was performed with REFMAC5 [276]. After the final refinement step, ligand molecules were removed from the model and a simulated annealing omit difference map was calculated with PHENIX [277]. The final structures were validated using MolProbity [278].

3.4.4 Nano differential scanning fluorimetry (nanoDSF)

To analyze the influence of ligand binding on JNK3 stability, nanoDSF measurements were conducted using a Prometheus NT.48 instrument (NanoTemper Technologies). With this dual-UV technology, protein unfolding can be followed by monitoring changes in intrinsic fluorescence of tryptophans and tyrosines as a result of their exposure to aqueous buffer environment.

Protein and ligand samples were prepared in protein buffer containing 50 mM HEPES pH 7, 100 mM NaCl, 2 mM $MgCl_2$, 10 mM β-mercaptoethanol. Compounds (stock concentration of 10 mM in DMSO) and AMP-PCP (stock concentration of 19 mM in protein buffer) were diluted to 100 µM in protein buffer. Final concentrations of 5 µM JNK3 and 50 µM ligand were used. 10 µL of each sample were loaded into glass capillaries and measured in triplicates. Capillaries were heated from 20 to 70°C with a heating rate of 1°C/min. Changes in fluorescence were monitored as the ratio of the emission wavelengths 350 and 330 nm as a function of temperature. First derivative analysis was used to determine unfolding temperatures (T_m).

4 Project I

The role of properdin in the activation of the alternative complement pathway

Contributions to this work

Dr. Katharina Hipp from the electron microscopy facility of the Max Planck Institute for Developmental Biology, Tübingen, analyzed purified properdin samples by negative-stain transmission electron microscopy (TEM). Dr. Lisete Silva and Dr. Yan Liu from the Glycosciences laboratory, Department of Medicine from the Imperial College London, performed and analyzed a glycan array screening of purified properdin. Dr. Bärbel Blaum from the Interfaculty Institute of Biochemistry, University of Tübingen, carried out STD-NMR measurements.

4.1 Results

4.1.1 Design of properdin expression constructs

Expression constructs of human properdin for recombinant production in mammalian cells were designed based on a publication by Pedersen *et al.* using a pCEP4 expression vector [71]. These constructs contained the endogenous Kozak sequence (CAA<u>ATG</u>A) followed by the native signal peptide of properdin required for secretion, which comprises 27 amino acids. The designed sequence lacked the native properdin stop codon, TAA, reported to terminate the properdin gene [279]. Instead, a c-myc tag or the fluorescent protein mCherry tag were C-terminally fused to properdin after a TEV cleavage site and glycine/serine linker to facilitate detection of protein expression and protein purification (Figure 4.1).
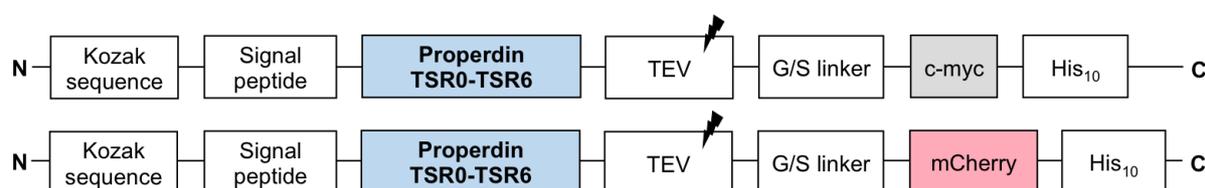
Figure 4.1 Expression constructs designed for the recombinant production of human properdin in mammalian cells. The TEV cleavage site is indicated and c-myc and mCherry tags used for detection are highlighted in different colors. The different box length indicates the relative size of the proteins. TSR stands for thrombospondin repeat.

Although the gene reporter mCherry is often used for detection of cytosolic proteins, in this work mCherry was C-terminally fused to properdin (properdin_mCherry) and protein expression was analyzed by fluorescence measurements. In a second construct, the mCherry sequence was replaced with the c-myc epitope (properdin_c-myc), which is commonly used to detect protein expression via western blots. Additionally, both constructs comprised a His_{10}-tag at the C-terminus for purification via Ni^{2+} affinity chromatography.

4.1.2 Transient transfection and expression of properdin

The mammalian FreeStyle 293-F cell line adapted to suspension culture in serum-free medium was used to express both the properdin_mCherry and the properdin_c-myc constructs. Cells were transiently transfected with 0.5 mg/L plasmid DNA and 2 mg/L PEI and incubated at 37°C, 8% CO_2 on an orbital shaker platform rotating at 125 rpm. Supernatant containing secreted properdin was collected 3 days post-transfection, 10-fold concentrated using a Vivaspin 20 centrifugal concentrator and immunoprecipitated using an anti-His affinity resin in order to increase the concentration of target protein for detection. Supernatant from transfected cells without plasmid DNA was treated identically and used as a negative control to identify unspecific bands in SDS-PAGE and western blot analysis. Samples of the collected supernatant before and after concentration and immunoprecipitation were analyzed by SDS-PAGE and identical bands were observed for all samples with the exception of a faint band at approximately 100 kDa within the concentrated supernatant of cells transfected with the properdin_mCherry construct (Figure 4.2A). This band could correspond to secreted properdin fused to mCherry although the calculated molecular weight based on the primary amino acid sequence would be only 78.1 kDa (Table 9.1). Since western blotting is more sensitive than SDS-PAGE and specific antibodies were available to identify the target proteins, samples of collected supernatant were analyzed using anti-mCherry and anti-c-myc as primary antibodies, respectively. For the western blot incubated with anti-mCherry no significant bands were observed for the negative control and samples containing properdin fused to c-myc. However, supernatant samples of properdin_mCherry showed two identical strong bands at about 100 kDa, which most likely corresponded to the secreted and glycosylated properdin fused to mCherry (Figure 4.2B). Instead, for the western blot incubated with anti-c-myc antibody, multiple bands were observed for all concentrated and immunoprecipitated supernatant samples, including the negative control (Figure 4.2C). These bands observed at MWs lower than 55 kDa resulted from unspecific binding of the antibody to the anti-His affinity resin used for immunoprecipitation. Protein bands with higher intensity were observed at about 67 kDa for supernatant samples of properdin_c-myc and most likely corresponded to the secreted protein fused to the c-myc epitope despite the calculated theoretical MW of 52.6 kDa (Table 9.1).
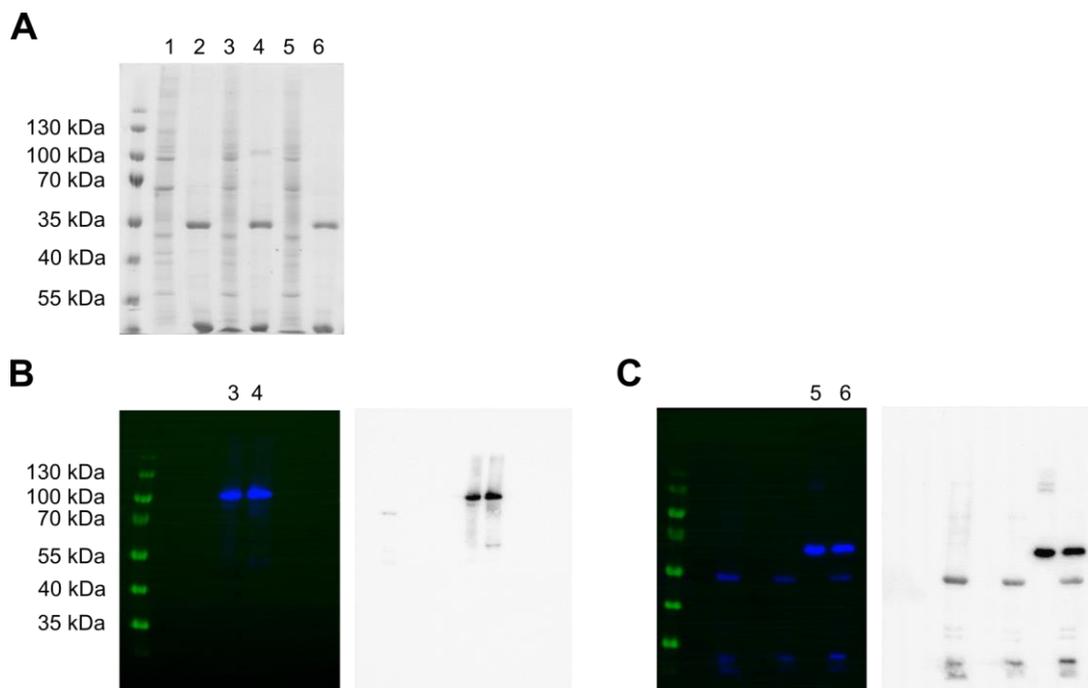
Figure 4.2 Expression of properdin constructs in FreeStyle-293F cells. A) SDS-PAGE of supernatant samples stained with Instant Blue solution. Lanes assigned with odd and even numbers correspond to supernatant before and after concentration and immunoprecipitation, respectively. (1, 2) negative control, (3, 4) supernatant of properdin_mCherry and (5, 6) supernatant of properdin_c-myc. B) Western blot obtained using anti-mCherry as primary antibody. Bands at approximately 100 kDa correspond to expressed properdin fused to mCherry. C) Western blot obtained with anti-c-myc antibody. Bands visible at approximately 65 kDa correspond to expressed properdin fused to the c-myc epitope, whereas bands at lower MWs probably resulted from unspecific binding of the antibody to the resin.

Since properdin is heavily glycosylated, the higher molecular weight observed by SDS-PAGE and western blot in comparison to the predicted molecular weight based on the amino acid sequence could arise through glycosylation [280]. However, as the glycosylation pattern of properdin produced in FreeStyle-293F cells is unknown, the theoretical MW of the secreted proteins could not be calculated precisely.

4.1.3 Direct expression detection of properdin fused to mCherry

Expression of properdin fused to the fluorescent protein mCherry was further analyzed by fluorescence measurements over a period of 7 days post-transfection. Cells were transfected with two different ratios of plasmid DNA vs. transfection reagent PEI (1:4 and 1:2, respectively) with DNA concentrations varying between 0.5 mg/L and 1 mg/L and a PEI concentration of 2 mg/L. Fluorescence intensities of supernatant samples collected every day during a period of 7 days after transfection were measured and correlated with the amount of expressed and secreted mCherry (Figure 4.3A). The background fluorescence of serum-free medium was determined by measuring supernatant from non-transfected cells and showed that the measured fluorescence signal was only derived from secreted mCherry

protein and not from the medium used to grow the cells. For all supernatant samples, an increase of the mCherry fluorescence signal was observed over time, indicating that properdin fused to mCherry was still being secreted 7 days after transfection. This fluorescence signal was higher for supernatant from cells transfected with a DNA:PEI ratio of 1:4 than for those transfected with a 1:2 ratio. The transfection efficiency of cells incubated with both DNA:PEI mixtures was determined by FACS analysis using the gene reporter mCherry (Figure 4.3B). Transfected cells harvested over a period of 7 days post-transfection showed an increasing number of fluorescent cells, indicating that properdin fused to mCherry was successfully expressed. On day 7, about 42% of the cells transiently transfected with a ratio of 1:4 were fluorescent, whereas only about 28% of the cells were fluorescent within the total cell population when transfected with a ratio of 1:2.
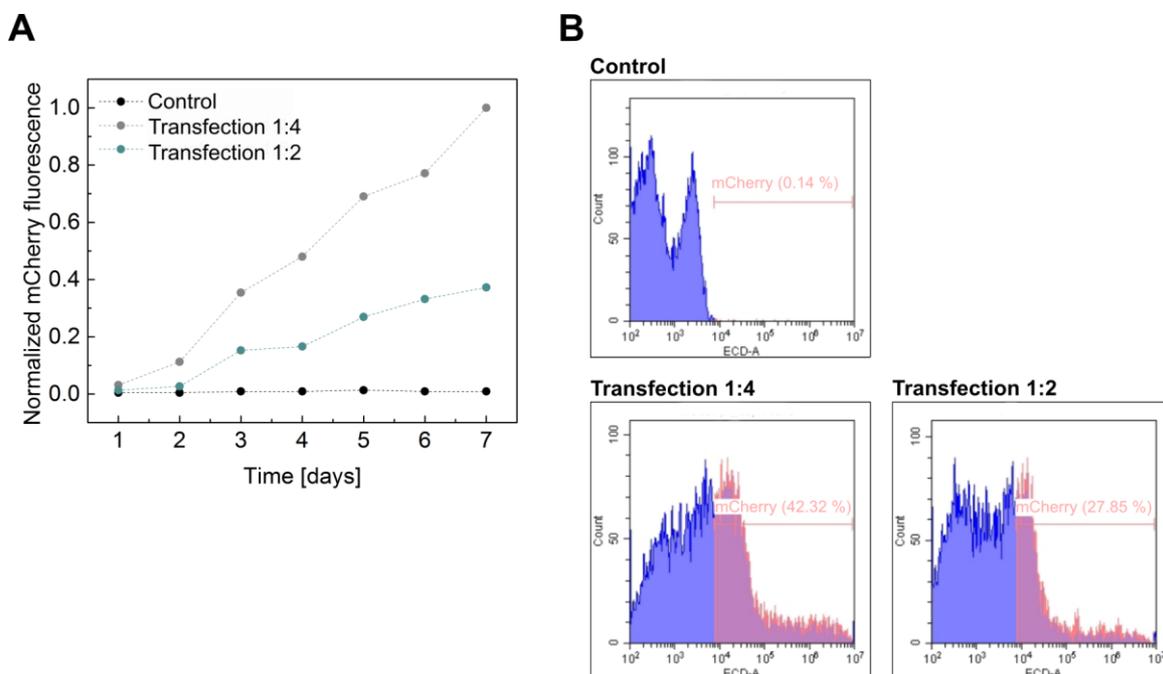
Figure 4.3 Expression of properdin_mCherry followed by fluorescence of secreted mCherry and FACS analysis over a period of 7 days post-transfection. Cells were transiently transfected with mixtures containing different ratios of PEI:plasmid DNA. Non-transfected cells were used as negative control. A) Normalized fluorescence signals of secreted mCherry followed over time. Supernatant samples were collected and analyzed each day and the correspondent fluorescent signal is depicted as spheres. Serum-free medium from non-transfect cells (control) is displayed in black and supernatants from transfection with ratios of PEI:DNA 1:4 and 1:2 are highlighted in gray and teal, respectively. B) Transfection efficiency analyzed by FACS on day 7 post-transfection. Expressed mCherry was used as a reporter gene to follow transfection of the cells. Transfection efficiency is shown as the % of fluorescence cells indicated for each plot.

Considering that both experiments showed a higher fluorescence signal for cells transfected with a lower amount of DNA (DNA:PEI ratio of 1:4), this ratio was selected to transiently transfect cells with the properdin_mCherry construct for properdin production as well as an incubation period of 7 days post-transfection. The same DNA:PEI ratio and incubation period

were also used for transfection of properdin_c-myc since no fluorescent measurements could be performed with this construct. Expression of properdin fused to mCherry was improved by reducing the amount of cells present in suspension from 2×10^6 cells/mL to 1×10^6 cells/mL on the day of transfection (Figure 4.4). FACS analysis showed that, under these conditions, 66% of the cells were fluorescent 7 days after transfection, meaning that higher transfection efficiency was achieved when a smaller number of cells was used for transfection.
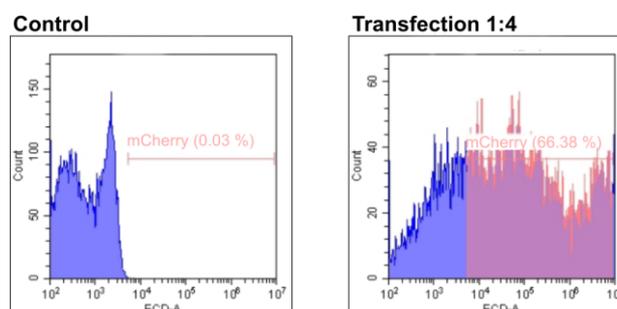
Figure 4.4 Improved expression of properdin_mCherry. Transfection was performed with a lower cell density (1×10^6 cells/mL). Non-transfected cells were used as a control and a transfection efficiency of about 66% was determined by FACS after an incubation period of 7 days post-transfection.

These results were confirmed by fluorescence measurements performed using supernatant samples, which also showed a higher fluorescence signal compared to previous experiments. Therefore, transfection of a lower number of cells in suspension increases the amount of expressed and secreted protein that can be used for purification.

4.1.4 Purification of recombinant properdin

Secreted properdin fused to mCherry followed by a His_{10}-tag was purified by Ni^{2+} affinity chromatography. Approximately 1 L of harvested supernatant diluted 1:1 in buffer containing 20 mM sodium phosphate, 500 mM NaCl, 5 mM imidazole adjusted to pH 7.4 was loaded onto a 5 mL HisTrap excel column (Figure 4.5A). A washing step was performed with 125 mM imidazole to remove unspecifically bound proteins from the cultivation medium and His-tagged properdin was eluted with 425 mM imidazole. Fractions containing the target protein were pooled and 1 µg of self-produced TEV protease per 10 µg of His-tagged properdin was added. The mixture was dialyzed into a buffer containing 20 mM sodium phosphate pH 7.4, 500 mM NaCl and 0.5 mM EDTA and subsequently purified by reverse Ni^{2+} affinity chromatography using a 1 mL Histrap excel column (Figure 4.5B). Cleaved properdin was collected in the flow-through and the mCherry_His_{10}-tag as well as the TEV protease (also fused to a His-tag) bound to the column and were eluted with 500 mM imidazole. Purified properdin was dialyzed against a buffer containing 10 mM HEPES pH 7.3 and 150 mM NaCl, concentrated and subsequently analyzed by analytical SEC using a Superose 6 column. Properdin is known to oligomerize and form cyclic

tetramers, trimers and dimers in a 20:54:26 ratio [59]. Size exclusion chromatograms of properdin before and after tag removal revealed the presence of properdin in three different oligomerization states (Figure 4.5C). The three elution peaks were observed at 1.22 mL, 1.31 mL and 1.47 mL, which corresponded to approximate molecular weights of 331 kDa, 203 kDa and 85 kDa, respectively (column calibrated with globular proteins). Considering that SDS-PAGE bands corresponding to properdin were observed at higher MWs, the three elution peaks most likely corresponded to tetrameric, trimeric and dimeric properdin. However, these oligomeric forms of properdin were present in a different ratio compared to previously reported data. Properdin samples collected at the different purification stages (prior to the first purification step to cleaved and purified protein) were analyzed by SDS-PAGE and compared to properdin (Factor P) purchased from Complement Technology Inc., which has been purified from human serum (Figure 4.5D). Protein bands at approximately 58 kDa showed that recombinantly produced properdin had an identical molecular weight to properdin purified from human serum. Besides, the observed molecular weight shift from approximately 100 kDa to 58 kDa as well as the presence of single bands in sample of properdin after mCherry and His-tag removal also indicated that properdin was successfully purified.

On average, about 0.5 mg of purified mature properdin were obtained from approximately 500 mL of harvested supernatant when using a cell density of 2×10^6 cells/mL for transfection. However, when the cell density was reduced to 1×10^6 cells/mL, a 2-fold increase in the protein yield was observed, which correlated with the higher transfection efficiency measured for these cells.
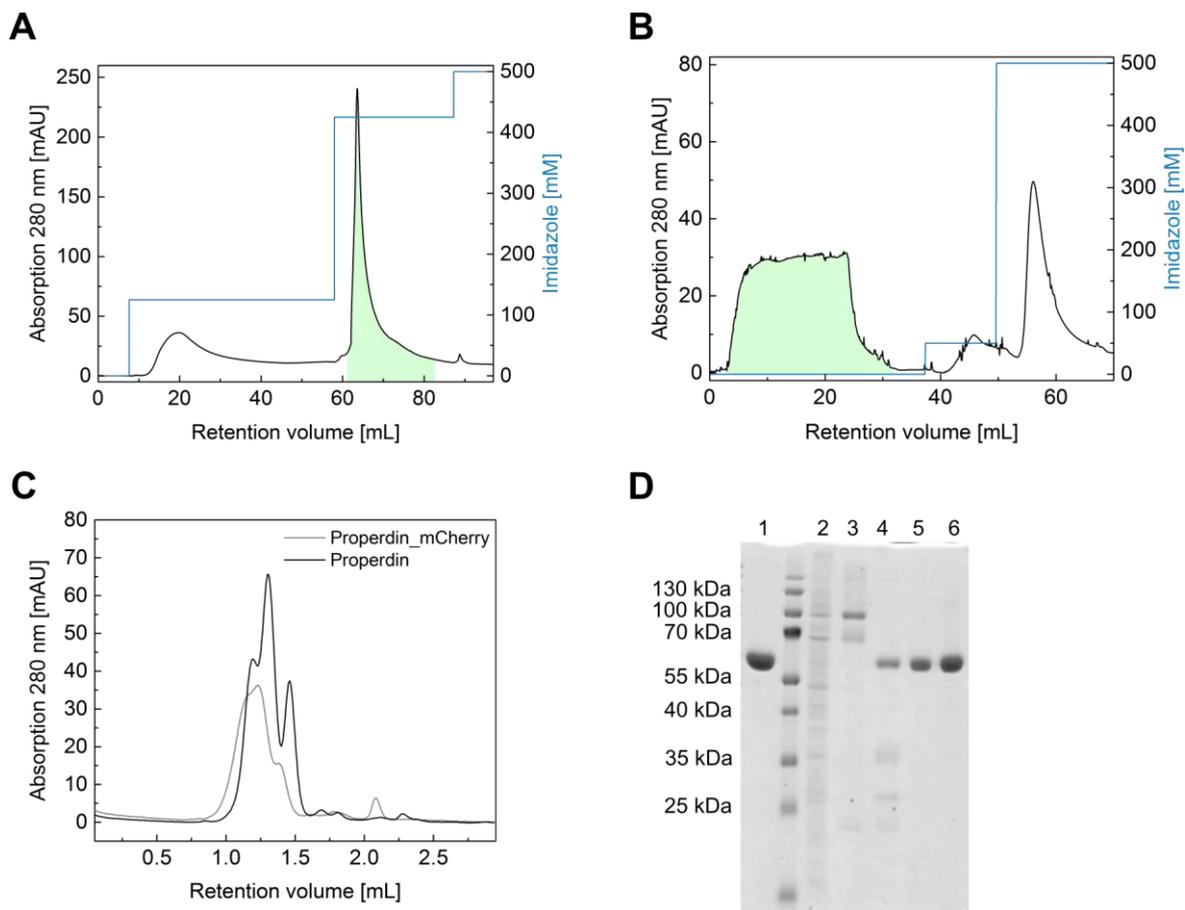
Figure 4.5 Purification of properdin fused to mCherry. A) Ni^{2+} affinity chromatography. Properdin eluted with 400 mM imidazole and pooled fractions are highlighted in green. B) Reverse Ni^{2+} affinity chromatography. Cleaved properdin was collected in the flow-through and pooled fractions are highlighted in green. C) Analytical size exclusion chromatograms obtained using a Superose 6 PC 3.2/300 column. Chromatograms of properdin fused to mCherry after the first purification step and cleaved properdin after reverse Ni^{2+} affinity are highlighted in gray and black, respectively. D) SDS-PAGE analysis of properdin samples from the different purification steps, including commercial properdin (Factor P) from Complement Technology as positive control. (1) Factor P, (2) harvested supernatant containing secreted target protein, (3) pooled fractions after the first purification step, (4) dialyzed protein containing the TEV protease and the cleaved tag and (5) purified properdin after tag removal (6) concentrated purified properdin used for SEC analysis. Proteins bands at approximately 58 kDa correspond to Factor P and purified properdin after tag removal.

Analysis of properdin samples by western blot using anti-mCherry (Figure 4.6A) and anti-properdin antibodies (Figure 4.6) confirmed the identity of bands observed by SDS-PAGE. The western blot developed with an anti-mCherry antibody demonstrated that cleavage of properdin and tag removal was successfully achieved, and a western blot obtained with anti-properdin antibody confirmed the presence of the target protein in the purified samples.
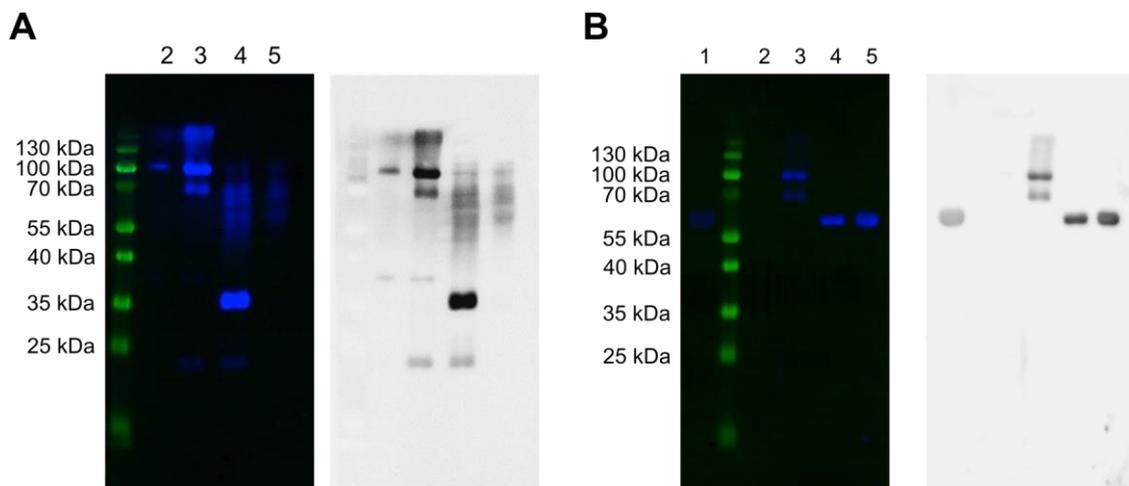
Figure 4.6 Western blot analysis of properdin samples collected throughout the purification. Samples correspond to (1) commercial Factor P, used as positive control, (2) harvested supernatant containing secreted target protein, (3) pooled fractions after the first purification step, (4) dialyzed protein containing the TEV protease and the cleaved tag and (5) pure properdin after tag removal A) Western blot obtained with anti-mCherry as primary antibody. B) Western blot obtained with anti-properdin as primary antibody.

Purification of supernatant containing secreted properdin fused to c-myc was also performed by Ni^{2+} affinity chromatography. After the first purification step, His-tagged properdin fused to c-myc was analyzed by analytic SEC using a Superose 6 column (Figure 4.7A) and SDS-PAGE (Figure 4.7B). Both results showed that the production of properdin fused to c-myc was identical to the production of properdin fused to mCherry and that the corresponding oligomeric species were present in solution.
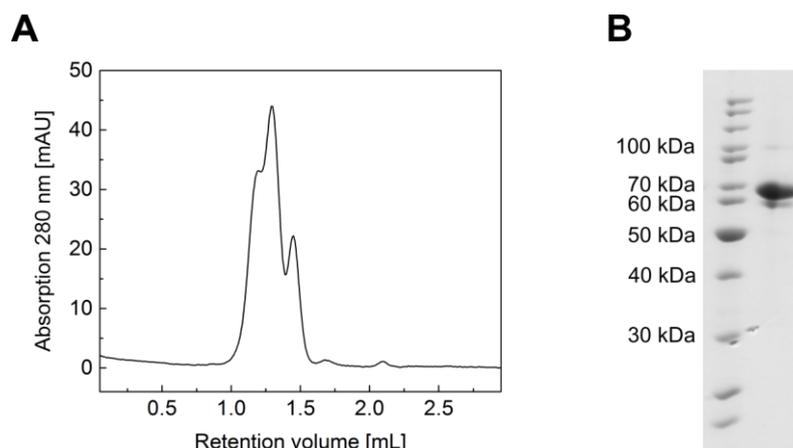
Figure 4.7 Analysis of His-tagged properdin fused to c-myc. A) Size exclusion chromatogram obtained using a Superose 6 PC 3.2/300 column showing a similar elution profile to cleaved properdin. B) SDS-PAGE analysis of His-tagged properdin fused of c-myc after the first purification step.

Based on the results from both purifications, mCherry did not seem to influence the biophysical properties of purified properdin. Besides the advantages of using this protein to

detect transfection and protein expression, this red fluorescent protein is also useful for purification since its presence can be easily observed by the color of the solution.

4.1.5 Biochemical and biophysical characterization of properdin

After purification, properdin samples fused to either mCherry (Figure 4.8A) or c-myc (Figure 4.8B) or after tag removal (Figure 4.8C) were examined by negative-staining TEM to structurally determine the effect of the different tags on the formation of properdin oligomers. Electron micrographs showed a similar morphology for all samples, demonstrating that tags fused to properdin did not impact its structural properties. Properdin was mostly composed of ring structures heterogeneous in size. Structures resembling a triangle or a rectangle were also observed and most likely corresponded to trimeric and tetrameric properdin [56]. Although aggregated species were not observed by SEC analysis, EM micrographs showed large amorphous aggregates (not shown here), which could be caused during sample preparation. Negative staining using uranyl acetate might have affected the structural integrity of properdin and contributed to its aggregation. Electron micrographs of freshly purified samples or after a freeze-thaw cycle showed that the aggregate content was not altered. However, a higher amount of aggregates was observed in electron micrographs of carbon-coated grids that were incubated with the different proteins, negative-stained and stored for a few weeks before analyses, suggesting that the formation of aggregates could be an artifact of sample preparation and grid storage.

A

B

C
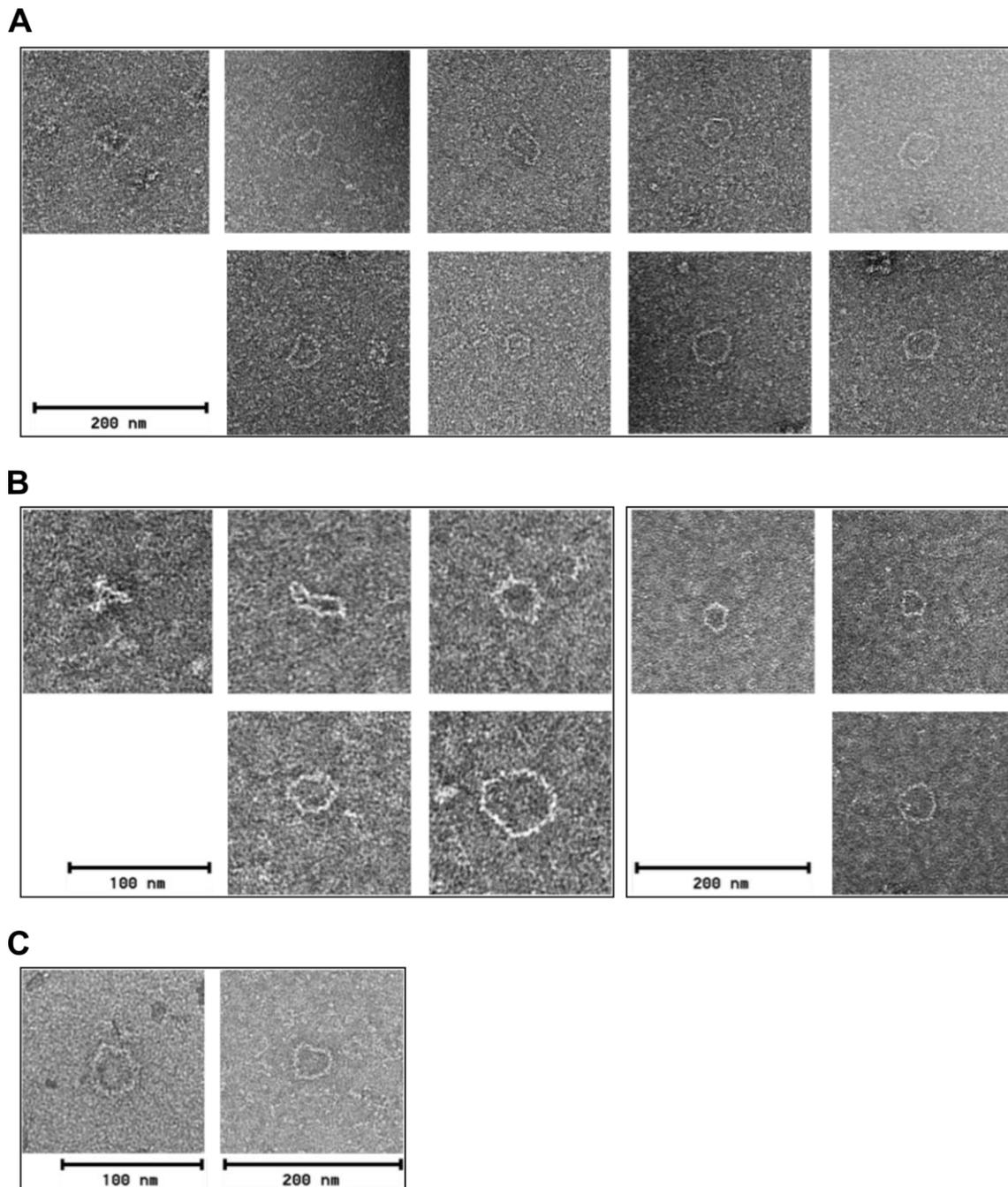
Figure 4.8 Gallery of raw images of properdin oligomers selected from electron micrographs. Glow-discharged carbon-coated grids were incubated with purified properdin variants (concentrations ranging from 40 µg/mL to 60 µg/mL), negatively stained with 1% (w/v) uranyl acetate and examined with a Tecnai G2 Spirit BioTwin transmission electron microscope operated at 120 kV. Images were collected on a Gatan Ultrascan 4 000 CCD-camera at final magnifications of 30 000x or 49 000x. 100 nm or 200 nm scale bars are included next to the respective selected images of properdin oligomers. Electron micrographs were provided by Dr. Katharina Hipp from the Max Planck Institute for Developmental Biology, Tübingen. A) Electron micrographs of properdin fused to mCherry. B) Electron micrographs of properdin fused to c-myc. C) Electron micrographs of properdin after tag removal.

Since EM results further supported the notion that mCherry did not have an impact on the biophysical properties of properdin, the properdin_mCherry expression construct was

selected to recombinantly produce properdin for further experiments, including biochemical and biophysical characterization and interaction studies.

Mature properdin contains a single *N*-glycosylation site in the sixth TSR, which is not essential for its activity or oligomerization [66]. Since the presence of such a long, flexible glycan chain can interfere with the formation of crystal lattices, enzymatic deglycosylation of properdin was attempted using endoglycosidases Endo Hf and PNGase F (Figure 4.9A) [281]. These two enzymes cleave *N*-linked glycans to different extents, with PNGase F having the broadest specificity, cleaving nearly all types of *N*-linked glycans. SDS-PAGE analysis of purified properdin samples incubated with both enzymes showed that Endo Hf resulted in partial enzymatic deglycosylation, with only a small reduction in size. For properdin incubated with PNGase F a molecular weight shift of 5 kD was observed, suggesting that removal of the *N*-linked glycan was successfully accomplished. Besides the bands corresponding to properdin, observed between 50 kDa and 60 kDa, the additional bands visualized on the SDS-PAGE gel corresponded to Endo Hf and PNGase F enzymes, which have apparent molecular weights of 70 kDa and 36 kDa, respectively. Purified properdin was further characterized by comparing its oligomeric distribution by analytical SEC after a freeze-thaw cycle and after an incubation at 4°C for 7 days (Figure 4.9B). Similar to freshly purified properdin, size exclusion chromatograms demonstrated the presence of different oligomeric species in solution, which were previously presumed to be tetramers, trimers and monomers of properdin. According to these results, storage of purified protein samples for a short period of time at different temperatures did not influence the structural stability nor the oligomerization profile of properdin.
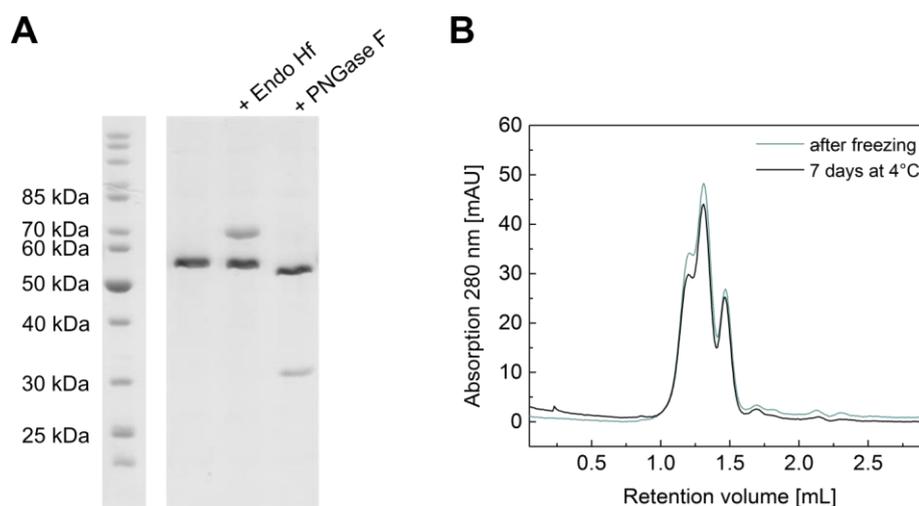
Figure 4.9 Biochemical characterization of purified properdin. A) Enzymatic deglycosylation of purified properdin after tag removal. Endo Hf and PNGase F were incubated with properdin samples and differences in MWs observed by SDS-PAGE correlate with the *N*-glycosylation state. B) Analysis of properdin stability after a storage period of 7 days at different temperatures. Size exclusion chromatograms were obtained using a Superose 6 PC 3.2/300 column. SEC profiles of properdir stored at 4°C or after freeze and thawing are depicted in teal and black, respectively.

Thermal unfolding of purified properdin was monitored by measuring the fluorescence intensity of an aromatic fluorescent dye upon binding to properdin during a temperature gradient (Figure 4.10A). Results from the first derivative analysis of fluorescence profiles measured in triplicate by DSF showed that purified properdin has a moderate stability with a determined T_m of 57.4±0.3°C. CD spectroscopy was used to analyze the secondary structure content of purified properdin but recorded spectra were untypical and did not display characteristics associated with the different secondary structural elements (Figure 4.10B). A prominent ellipticity maximum was consistently observed at about 230 nm for different recorded spectra but ellipticity measured at lower wavelengths was variable. This maximum most likely resulted from the presence of disulfide bonds, which are optically active and contribute to CD [60], [282]. Native properdin contains an absolute number of 20 disulfide bridges.
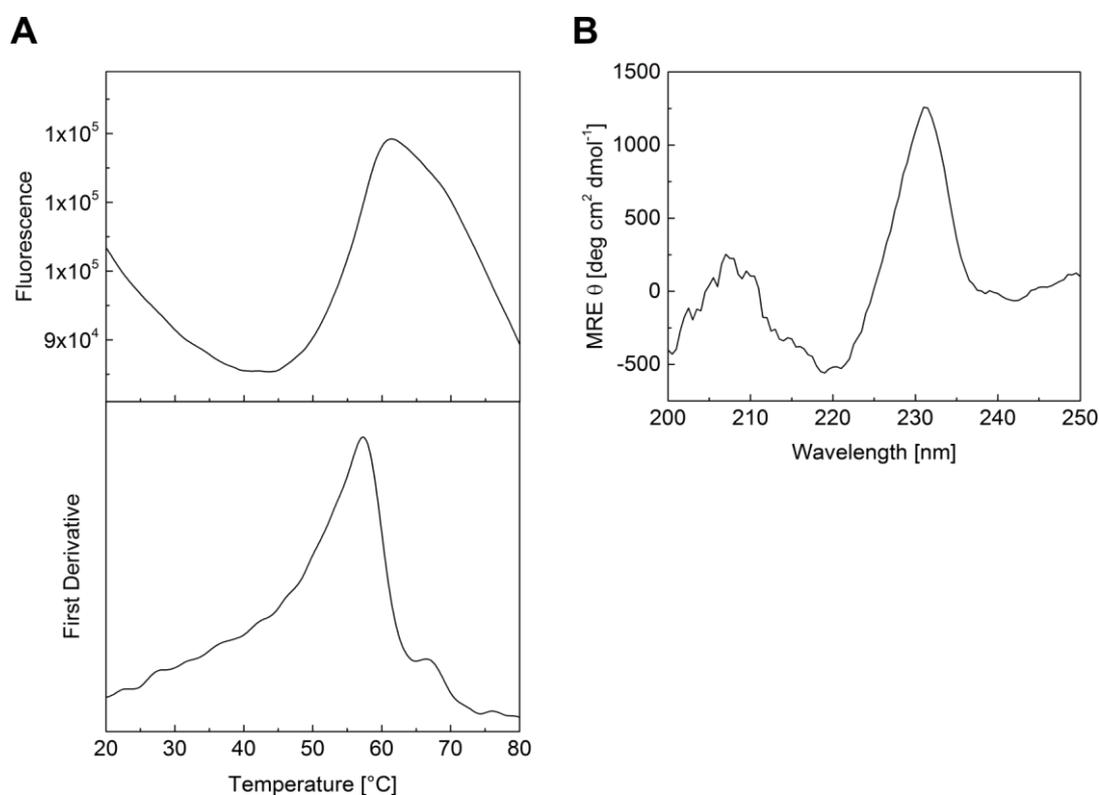
Figure 4.10 Biophysical characterization of purified properdin by DSF and CD spectroscopy. A) Thermal unfolding profile of cleaved properdin determined by monitoring the fluorescence intensity of a specific dye upon binding to properdin and its first derivative are plotted as a function of temperature. A T_m of 57.4±0.3°C was determined for the first derivative analysis. B) Far-UV CD spectrum of cleaved properdin. The signal is expressed as mean residue weight ellipticity plotted as a function of wavelength. A maximum of ellipticity was observed at about 230 nm.

4.1.6 E244K properdin mutant

Monomeric properdin was recently reported to occur as a result of a single point mutation in TSR3. This mutation, consisting of the replacement of a glutamate residue at position 244 by

a lysine (E244K), was discovered in a patient with a type II properdin deficiency [71]. Since heterogeneous protein samples are not favorable for the formation of high-quality crystals for structural determination, site-directed mutagenesis was used to introduce this single point mutation in the properdin_mCherry construct. FreeStyle 293-F cells were transiently transfected as previously described for wild-type properdin fused to mCherry but a lower transfection efficiency as well as a lower fluorescence signal of the supernatant containing secreted protein were observed. Harvested supernatant was purified using Ni^{2+} affinity chromatography followed by tag cleavage using TEV protease. In contrast to previously published results, size exclusion chromatograms of the recombinantly produced E244K properdin mutant revealed a similar elution profile as observed for wild-type properdin (Figure 4.11A). Bands observed for purified E244K properdin prior and after cleavage by SDS-PAGE showed a lower molecular weight for properdin mutant fused to mCherry (85 kDa) compared to the wild-type mCherry fusion protein (100 kDa), which could be a result of reduced glycosylation (Figure 4.11B). Bands that appeared at lower molecular weights for the cleaved sample corresponded to the cleaved mCherry_His_{10}-tag and the TEV protease as observed by SEC analysis. The shift in molecular weight to about 55 kDa observed for the E244K properdin mutant confirmed that cleavage of the tag was successful.
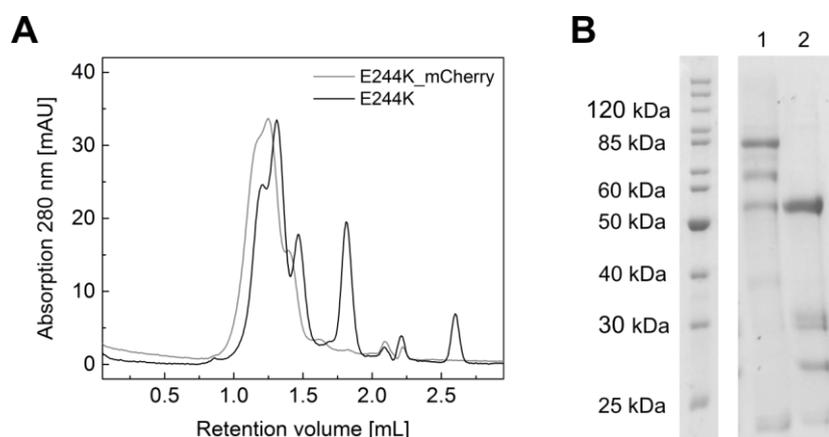
Figure 4.11 Analysis of the recombinantly produced E244K properdin mutant. A) Size exclusion chromatograms of E244K mutant before (gray) and after mCherry cleavage (black) obtained using a Superose 6 PC 3.2/300 column showed elution profiles similar to wild-type protein. The three non-separated elution peaks are presumed to be tetramers, trimers and monomers of properdin. For the cleaved E244K properdin mutant, additional elution peaks observed at higher retention volumes corresponded to TEV protease and the cleaved mCherry tag. B) SDS-PAGE analysis of (1) E244K properdin mutant fused to mCherry and (2) cleaved E244K properdin mutant containing the cleaved mCherry and TEV protease. Bands at approximately 85 kDa and 55 kDa correspond to the E244K properdin mutant before and after cleavage, respectively.

Following these unexpected observations, which did not confirm the reported findings that the E244K point mutation results in monomeric properdin, DNA sequencing of the mutated expression construct was repeated and the introduction of the single point mutation was

confirmed. The results indicated that no monomers were formed and the oligomerization state of properdin was identical to wild-type protein under the conditions used for production of this properdin mutant.

4.1.7 Alternative pathway-mediated hemolytic assay

To assess the functional activity of the purified properdin variants, a hemolytic assay was performed by exposing sheep erythrocytes (RBCs) treated with neuraminidase to normal human serum (NHS) or properdin-depleted serum (Figure 4.12). Unmodified sheep RBCs usually do not activate the human AP [283]. However, enzymatic removal of sialic acid by neuramidase treatment triggers or augments the activation of the human AP and leads to RBCs lysis in the presence of NHS but not when exposed to properdin-depleted serum [101].
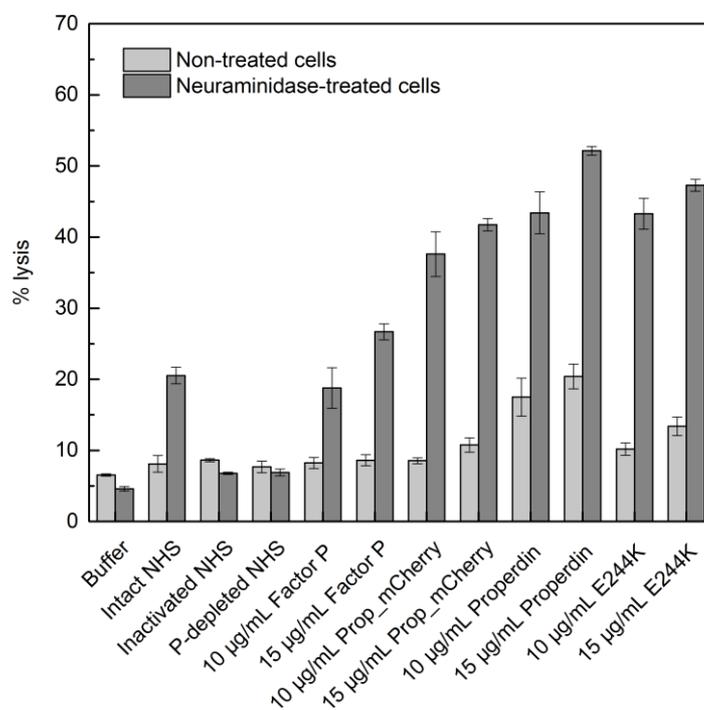
Figure 4.12 Hemolysis of sheep RBCs by purified properdin variants. Lysis of neuraminidase treated RBCs exposed to properdin-depleted NHS was stimulated through addition of properdin in a concentration-dependent manner. Reactions containing only buffer, NHS, heat-inactivated NHS, P-depleted NHS and Factor P from Complement Technology Inc. were included as controls. Lysis of non-treated RBCs is shown in light gray, whereas lysis of neuramidase treated RBCs is shown in dark gray. Lysis extent was quantified via absorption at 415 nm and are expressed relative to the total osmolysis of RBCs in water. The experiment was carried in duplicate and data shown represent mean values and standard deviations. Recombinant properdin was quantified based on the absorbance at 280 nm while the Factor P concentration was based on the manufacturer's product information.

All reactions were performed with neuraminidase treated and non-treated sheep RBCs, incubated at 37°C for 2 min and absorption at 415 nm was measured. Lysis of RBCs was normalized with respect to RBCs lysed in water. RBCs exposed to buffer, intact NHS,

heat-inactivated NHS and P-depleted NHS were used as controls. Sheep RBCs incubated with buffer containing 20 mM HEPES pH 7.3, 145 mM NaCl, 0.1% (w/v) gelatin, 5 mM $MgCl_2$ and 5 mM EGTA showed that about 5% of spontaneous lysis occurred under the chosen conditions. Reactions containing non-treated RBCs showed a low percentage of spontaneous lysis and increased lysis of treated RBCs was only observed when exposed to normal NHS but not in the presence of inactivated NHS or properdin-depleted serum. However, properdin-depleted serum supplemented with properdin triggered AP activation and stimulated lysis of treated sheep RBCs in a concentration-dependent manner. Purified recombinant properdin fused to mCherry or the c-myc tag, as well as, properdin after tag cleavage showed a higher level of RBCs lysis in comparison to Factor P purified from NHS. The highest activity was observed for purified properdin after tag removal, which also caused increased lysis of non-treated RBCs. Addition of 15 µg/mL of cleaved properdin resulted in about 55% of lysed RBCs. Purified E244K properdin mutant was also tested and lysis of RBCs was comparable to lysis of RBCs caused by other recombinantly produced properdin variants.

4.1.8 Analysis of properdin-glycan interactions

There is on-going controversy concerning the question whether properdin can act as a direct activator of the alternative complement pathway or merely act as a promotor of C3 convertase activity [2]. Specific glycans are likely candidates for the direct activation hypothesis [3]. To identify glycans that might bind to properdin and lead to AP activation, a glycan array screening containing 885 lipid-linked oligosaccharide probes, neoglycolipids (NGLs) and glycolipids was performed (Figure 4.13). Both purified properdin after mCherry removal and rabbit anti-properdin antibody were included in the array. The rabbit anti-properdin was used to detect properdin binding in the glycan array of properdin and served as background control. Glycan binding was detected with biotinylated anti-rabbit IgG followed by overlay with streptavidin-Alexa Fluor 647 and results were plotted as histogram charts of relative binding intensities. Both glycan arrays showed a low overall fluorescence signal and binding of properdin was observed only to glycan probes that were also bound by the antibody-based detection system in the absence of properdin. Apart from these binding signals, which displayed a good signal-to-noise ratio, negligible or no binding signals were detected in the glycan array screening of properdin under the conditions tested. These results suggested that either properdin does not bind to glycans under the tested conditions, that the glycan-affinity is beyond detection limit or that no specific-properdin glycans were present on the micro-array slides.
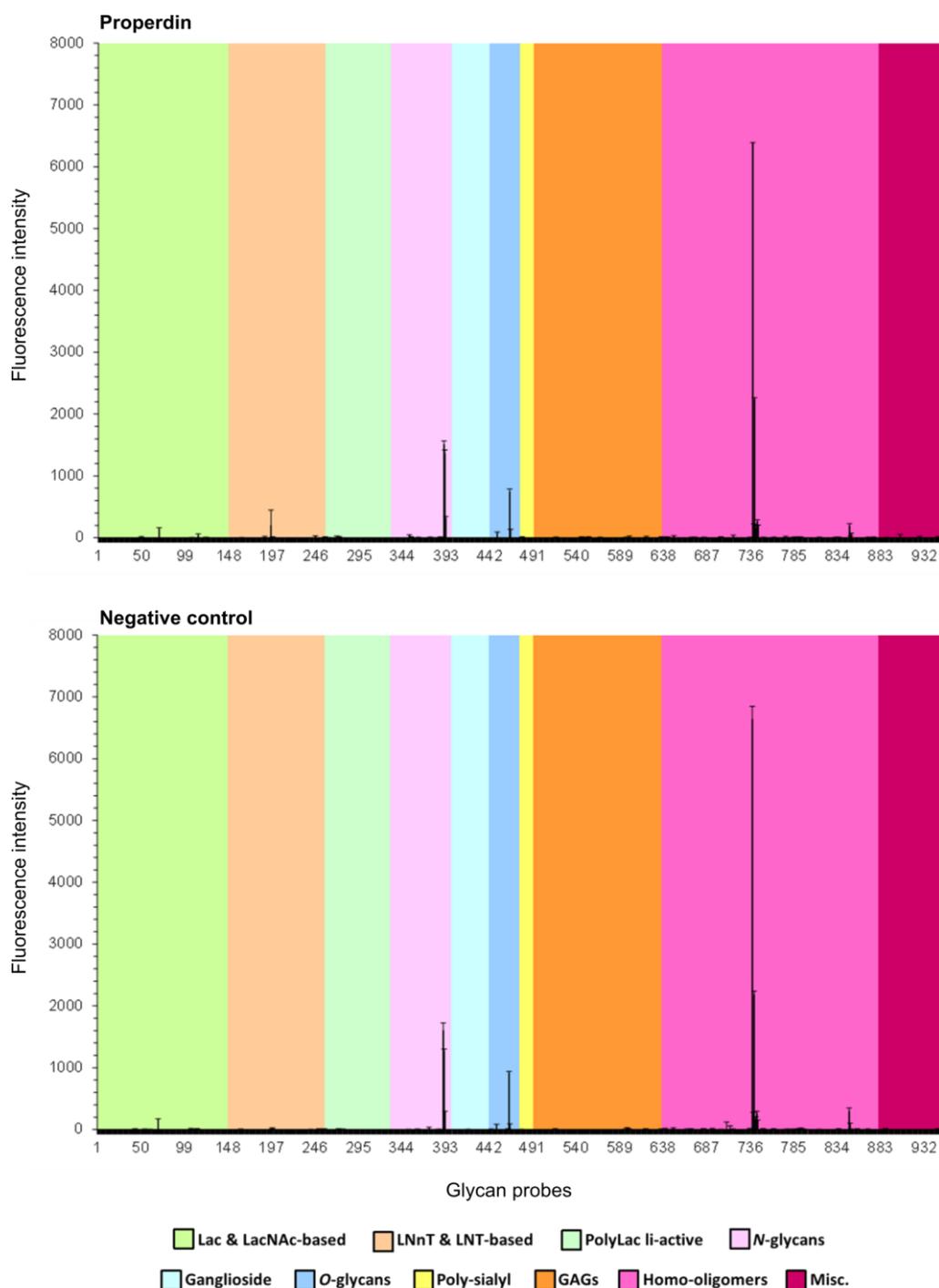
Figure 4.13 Glycan microarray analysis of properdin and anti-properdin. Properdin and rabbit anti-properdin antibody from Abcam (used as detection system and in the absence of protein as negative control) were diluted in blocking solution containing 10 mM HEPES pH 7.4, 150 mM NaCl, 10 mM $CaCl_2$, 1% (w/v) bovine serum albumin and 0.02% (w/v) casein and overlaid onto the arrays at room temperature for 1.5 h followed by 1 h of incubation with the corresponding detection system (rabbit anti-properdin for properdin and biotinylated anti-rabbit IgG for the anti-properdin). Binding was detected with biotinylated anti-rabbit IgG, incubated for 1 h, followed by 30 min overlay with Alexa Fluor 647-labeled streptavidin. Both samples were analyzed in a wide spectral screening and the binding signals are shown as means of the fluorescence intensity at 5 fmol/probe spot. Error bars represent half of the difference of signal intensities of duplicate spots for each glycan probe. A) Histogram chart of properdin (100 μg/mL). B) Histogram of the negative control rabbit anti-properdin antibody (1:200 dilution). The assay was performed by Dr. Yan Liu and Prof. Ten Feizi, Imperial College London.

Representatives of different types of glycans were tested for properdin binding by STD-NMR using purified properdin samples either fused to mCherry or after mCherry and His$_{10}$-tag removal. All spectra were recorded in D$_2$O.

Initial STD-NMR measurements were performed using properdin fused to mCherry and a mixture of two specific glycans, namely Arixtra, a synthetic heparin mimetic pentasaccharide, and Thomsen-Friedenreich (TF)-antigen, a disaccharide that resembles de-sialylated mammalian cell-surface glycans (Figure 4.14). In the STD-NMR difference spectrum, mostly resonances belonging to Arixtra were observed indicating that preferentially this glycan interacted with properdin.
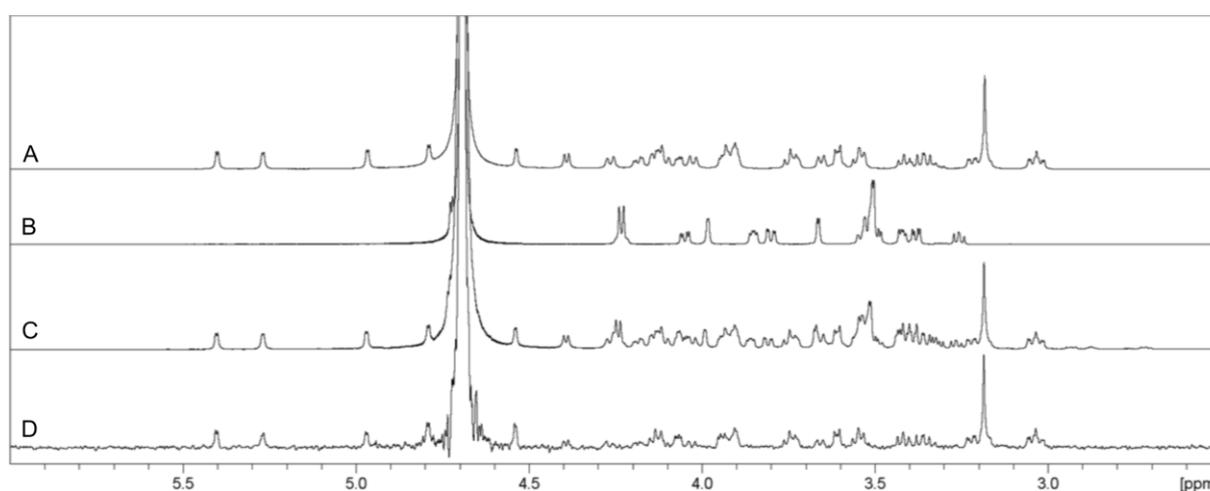
Figure 4.14 STD-NMR spectra of purified properdin fused to mCherry. Spectra were recorded using a protein concentration of 16.7 µM and glycans at a concentration of 1 mM each. The truncated peak at 4.7 ppm corresponds to the truncated HDO signal. A) ^{1}H NMR reference spectrum of Arixtra (GlcNS6S-GlcA-GlcNS3S6S-IdoA2S-GlcNS6S-OMe). B) ^{1}H NMR reference spectrum of TF-antigen (Galβ1-3GalNAc). C) ^{1}H NMR reference spectrum of Arixtra and TF-antigen mixture. D) STD-NMR difference spectrum of Arixtra and TF-antigen mixture with properdin_mCherry.

To verify if the mCherry_His$_{10}$-tag part of the properdin fusion protein could have influenced binding of glycans, binding of Arixtra was probed once more using purified properdin after tag removal (Figure 4.16). Similar to properdin_mCherry, the difference STD-NMR spectrum showed binding of Arixtra to cleaved properdin, suggesting that mCherry and His$_{10}$-tag did not interfere with glycan binding and that tag removal is most likely not necessary to perform such interactions studies and identify glycans that may bind to properdin. STD-NMR measurements with additional GAG molecules comprising different degrees of polymerization (dp4 and dp6) as well as a polysialic acid (dp5) were subsequently performed to evaluate the substantial negative charge effect on binding (Figure 4.15). Difference STD-NMR spectra showed binding of both GAGs, tetra- and hexasaccharides, to properdin, with several weak but unambiguous resonances from either molecule. Instead, in the STD-NMR spectrum of polysialic acid and properdin only resonances from the sialic acid

methyl groups were observed between 1 ppm and 2 ppm. These resonances were found to arise from relaxation artifacts by a control STD-NMR experiment performed without protein.
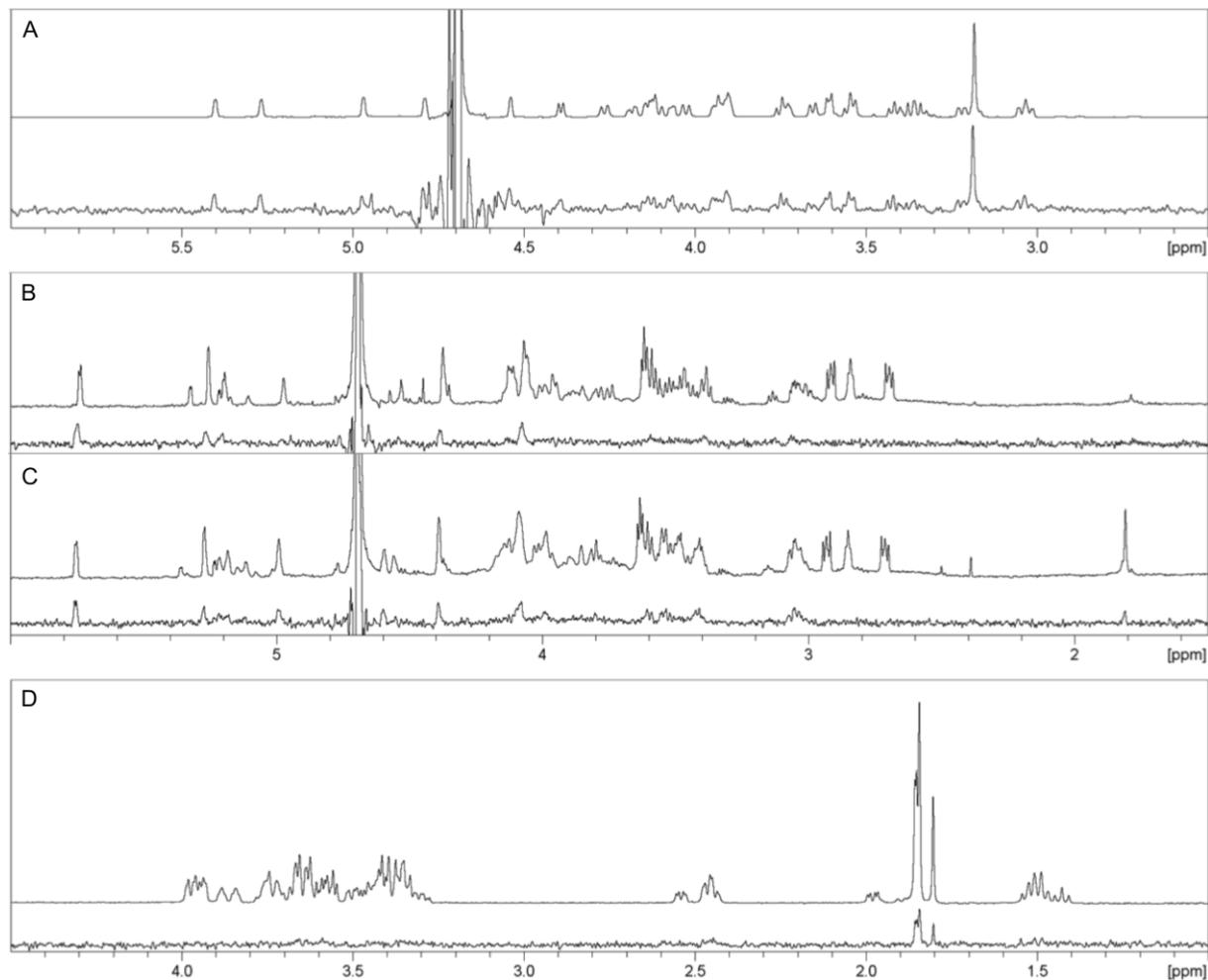
Figure 4.15 STD-NMR spectra of purified properdin after tag removal with size-defined charged glycans. The truncated peak at 4.7 ppm corresponds to the truncated HDO signal. A) ^{1}H NMR reference spectrum of Arixtra (top) and STD-NMR difference spectrum of 15.7 µM cleaved properdin with 1 mM Arixtra (bottom). B) ^{1}H NMR reference spectrum of heparin dp4 (top) and STD-NMR difference spectrum (bottom). C) ^{1}H NMR reference spectrum of heparin dp6 (top) and STD-NMR difference spectrum (bottom). D) ^{1}H NMR reference spectrum of polysialic acid (α2,8-linked Neu5Ac) dp5 (top) and STD-NMR difference spectrum (bottom). NMR measurements with dp4, dp6 and polysialic acid were performed using a glycan concentration of 1 mM and cleaved properdin at 15 µM each.

Apart from the TF-antigen, binding of two additional non-charged glycans such as lactose and asialo-GM1 oligosaccharide was also tested (Figure 4.16) However, no resonance signals were observed in either difference STD-NMR spectrum, which indicates that neither glycan bound to properdin and negative charge is most likely required for binding to properdin.
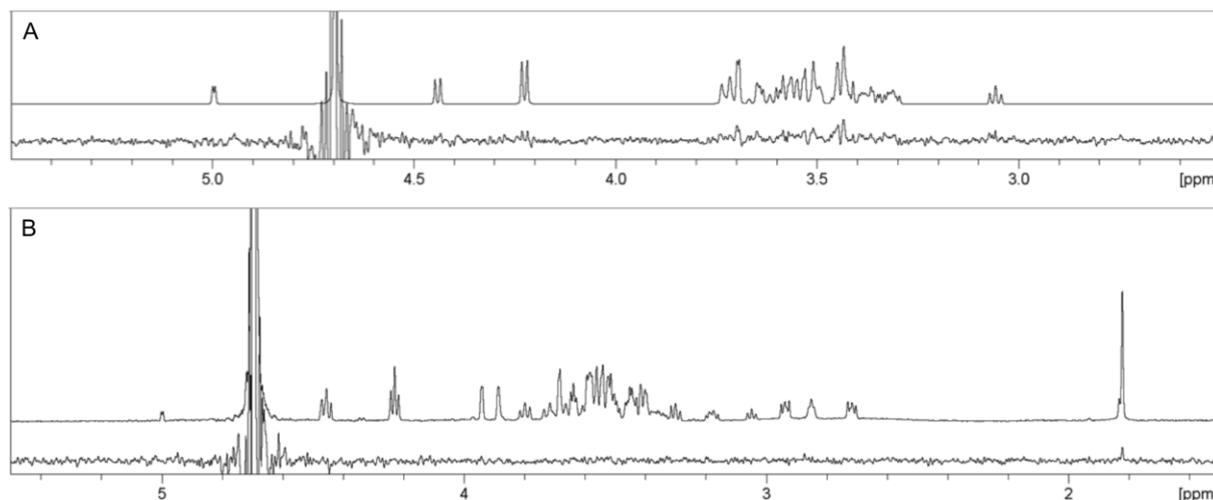
Figure 4.16 STD-NMR spectra of purified properdin after tag removal with non-charged, non-sialylated glycans. Spectra were recorded using a protein concentration of 15 µM and glycans at a concentration of 1 mM. The truncated peak at 4.7 ppm corresponds to the truncated HDO signal. A) ^{1}H NMR reference spectrum of lactose (Galβ1-4Glc) and STD-NMR difference spectrum (top and bottom, respectively). B) ^{1}H NMR reference spectrum of asialo-GM1 oligosaccharide (Galβ1-3GalNAcβ1-4Galβ1-4Glc) and STD-NMR difference spectrum (top and bottom, respectively).

Influence of glycan length and sulfation to properdin binding was also evaluated by STD-NMR using a mixture of defined GAGs, HepMers M06 S03a and M09 S06a from Iduron (Figure 4.17). These GAGs differ in length (6 vs. 9 pyranoses) and in sulfation. Resonances observed in the difference STD-NMR spectrum suggested that HepMer M09 S06a was most likely the only glycan binding to properdin. However, to confirm this observation, STD-NMR measurements should be repeated using the HepMers separately, given that both glycan spectra exhibit extensive spectral overlap.
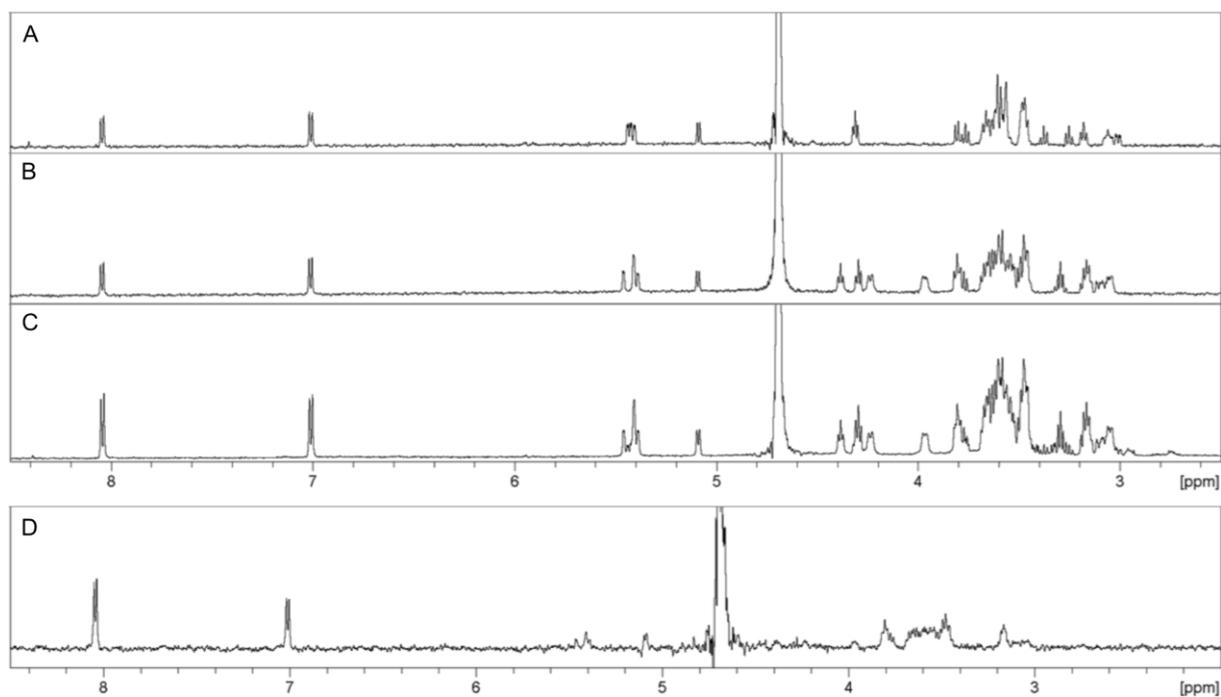
Figure 4.17 STD-NMR spectra of purified properdin after tag removal with sequence-defined heparin/HS oligosaccharides (HepMers). Spectra were recorded using a protein concentration of 15.7 µM and HepMers at a concentration of 1 mM. The truncated peak at 4.7 ppm corresponds to the truncated HDO signal. A) ^{1}H NMR reference spectrum of HepMer M06 S03a ((GlcNS-GlcA)$_3$-p-nitrophenyl). B) ^{1}H NMR reference spectrum of HepMer M09 S06a ((GlcA-GlcNS)$_2$-(GlcA-GlcNS6S)$_2$-GlcA-p-nitrophenyl). C) ^{1}H NMR reference spectrum of HepMers M06 S03a and M09 S06a mixture. D) STD-NMR difference spectrum of cleaved properdin with a mixture containing both HepMers.

Overall, STD-NMR experiments showed that some glycans bind to properdin in solution and that factors like glycan charge, length and most likely sulfation are critical for the interaction.

4.2 Discussion

Since its discovery in 1954, human properdin has been a subject of controversy within the scientific community. Although extensive research has been done, both the structure and crucial aspects of properdin function remain unclear and oligomerization is still a major obstacle for a detailed analysis.

Determination of three-dimensional structures requires high amounts of highly pure protein, which makes *E. coli* the most favorable expression organism for structural studies. However, recombinant production of proteins with correctly folded disulfide bonds, such as necessary for properdin, in *E. coli* cytoplasm remains a challenge due to its reductive environment. A construct comprising two domains of the structurally-related serum protein thrombospondin has been produced as glutathione *S*-transferase (GST) fusion protein in *E. coli* and its crystal structure has been determined, demonstrating that well-folded TSR domains proteins can be recombinantly produced in *E. coli* [284]. Based on these results, a similar approach was initially followed to express properdin constructs comprising TSR domains 4 to 6 fused to a N-terminal GST-tag or His-tag in *E. coli* cells, since these domains were shown to be involved in glycan binding and C3bBb-stabilizating functions [66]. The engineered *E. coli* strain named SHuffle was selected for properdin expression within the cytoplasm since it has been shown to promote folded and active proteins comprising disulfide-bonds. This is due to mutations in the *E. coli* cytoplasmic reductase genes (*trxB* and *gor*) and expression of the perisplasmic disulfide isomerase (DsbC) lacking its signal sequence in the cytoplasm [285]. Despite multiple attempts with varying parameters during protein expression and purification resulting properdin constructs were either insoluble or aggregates were observed after tag removal. One possibility is that besides correctly formed disulfide bonds, glycosylation might also be essential for the correct folding of properdin.

As an alternative, mammalian cells containing the complex machinery required for post-translational modifications can be used to produce correctly folded recombinant proteins. Based on a publication from Pedersen *et al.*, the FreeStyle 293F mammalian cells, derived from human embryonic kidney (HEK) 293 cells, were selected for the expression of human full-length properdin. By employing an approach using transient transfection, these cells can be used to generate significant amounts of recombinant protein in a short period of time. Since FreeStyle 293F cells are adapted to grow in suspension culture and reach high densities using serum-free medium, scalability of this process is not limited by cell culture surface availability and purification of secreted proteins is facilitated in absence of fetal bovine serum [286], [287].

Two expression constructs differing in their C-terminal part were designed, both comprising full-length properdin including its native signal peptide necessary for protein secretion and the endogenous Kozak sequence. Although detection of secreted proteins in the cell culture supernatant typically includes methods such as western blot or an enzyme-linked immunosorbent assay (ELISA), a detection method based on fluorescence measurements of a mCherry fusion reporter protein has also been shown to be applicable [267]. Therefore, one of the expression constructs was C-terminally fused to a c-myc epitope commonly used for protein detection via western blot, whereas the second expression construct was fused to the monomeric red fluorescent mCherry protein. Expression and secretion of both constructs was successfully observed by immunoprecipitation of cell culture supernatant samples followed by western blot analysis. To further improve protein expression and increase the yield of secreted protein, parameters including concentration of the cationic polymer PEI used as transfection reagent, concentration of DNA, incubation time after transfection as well as cell density at the time of transfection were analyzed and optimized using fluorescence signal of expressed properdin fused to mCherry. Cells and supernatant assayed up to 7 days post-transfection for the presence of fluorescence signal via FACS analysis and fluorescence measurements showed a higher transfection efficiency for cells transiently transfection with DNA:PEI complexes formed using a 1:4 and an increase in signal intensity over time. Transfection efficiency was optimized and the yield of secreted protein improved by transfecting cells with a viability greater than 90% and a cell density of 1×10^6 cells/mL instead of 2×10^6 cells/mL at the time of transfection. These results demonstrate the applicability of mCherry as a convenient reporter of successful production and secretion of complex human proteins in FreeStyle 293F cells.

Addition of a His$_{10}$-tag at the C-terminus of properdin expression constructs facilitated the purification of secreted protein by Ni^{2+} affinity chromatography. Removal of the C-terminally fused His10-tag and mCherry was accomplished by proteolytic cleavage using TEV protease and separation of these components resulted in highly pure mature properdin. Size exclusion chromatograms of properdin prior and after tag removal showed three overlapping elution peaks, which most likely corresponded to reported oligomeric species of properdin. This assumption was confirmed by direct comparison with properdin from human serum Factor P (Figure 4.18).
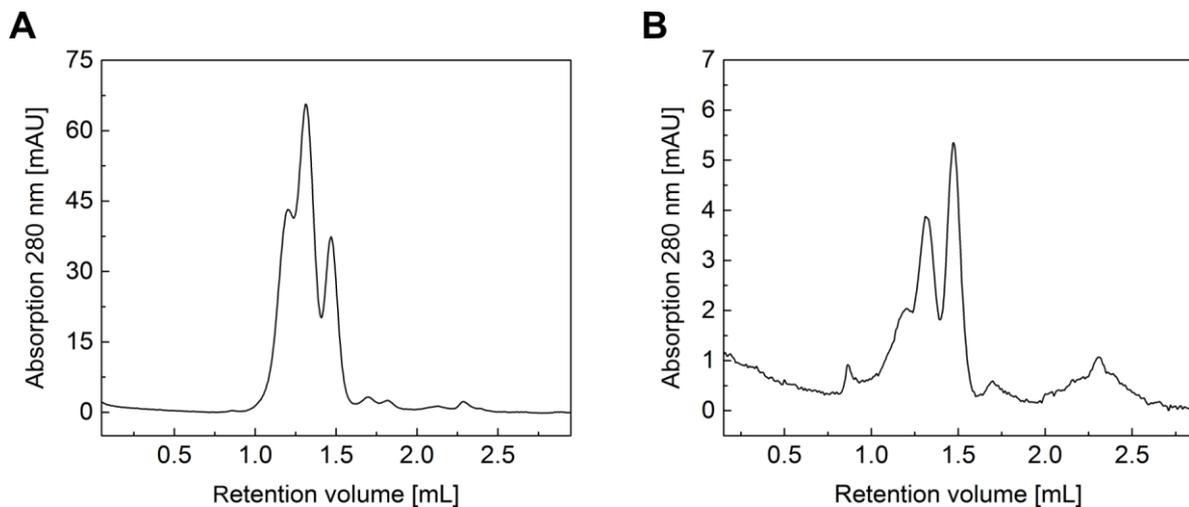
Figure 4.18 Comparison of analytical size exclusion chromatogram of recombinantly produced properdin after tag removal and Factor P purified from human serum. Factor P was purchased from Complement Technologies. On a Superose 6 PC 3.2/300 column three elution peaks corresponding to three different oligomerization states of properdin, namely tetramers, trimers and dimers, were observed.

Despite the presence of three elution peaks at identical retention volumes, the relative proportion of the oligomeric species differed between serum-purified and recombinantly produced properdin. Oligomers of properdin were reported to be present as mixture of tetramers, trimers and dimers in a ratio of 26:54:20, whereas for recombinantly produced properdin an approximate ratio of 29:47:24 was estimated based on the absorption elution peak maxima at 280 nm. However, this ratio was neither observed for cleaved properdin nor for Factor P. Redistribution of the oligomerization profile of properdin has been observed before after denaturation-renaturation cycles, exposure to low pH or to guanidine [59]. The fact that serum-purified properdin did not oligomerize in the reported ratio suggests that other ratios are possible and dependent on environmental parameters such as concentration, storage and buffer components, all of which are parameters well known to impact protein oligomerization. No void volume peak was observed for serum-purified nor recombinant properdin, excluding the formation of higher-order oligomers that were reported to occur [84].

Expression and purification of properdin fused to c-myc and His_{10}-tag showed similar elution profiles and purification levels as purified properdin fused to mCherry, indicating that a larger tag like mCherry at the C-terminus does not influence the structural and biochemical properties of recombinantly produced mature properdin. In contrast to protein recombinantly produced in *E. coli*, no aggregates or solubility issues were observed for both proteins produced in mammalian cells. Structural analysis by negative-staining TEM of unfractionated properdin samples, including properdin fused to mCherry or c-myc and cleaved properdin, showed discrete cyclic structures with variable sizes similar to those previously described by Smith *et al.* and Farries *et al.* [60], [84]. Besides these ring-shaped structures, which

corresponded to the different oligomeric species of properdin, non-physiological aggregates were also observed in the electron micrographs. Since no aggregates were observed via analytical SEC, it can be assumed that these large amorphous aggregates might have resulted from the harsh negative-staining treatment. Storage of protein samples fixed to the carbon-grids followed by negative-staining showed a higher amount of aggregates compared to samples that were prepared and immediately analyzed, which supported the assumption that properdin aggregates were an artifact of the sample treatment. Although the magnification used was not high enough to clearly visualize monomeric mCherry in the electron micrographs of properdin fused to mCherry, the similar morphology observed for all samples as well as the almost identical size exclusion chromatograms suggested that mCherry does not disturb the formation of properdin oligomers. Based on these results, the use of a fluorescent reporter like mCherry for direct detection of secreted protein in the cell culture media since is highly advantageous, as previously demonstrated for other proteins [267]. Measuring fluorescence signals of cell culture supernatant containing secreted protein has several advantages including measurement time, ease-of-use and reduced costs when compared to typical western blot and ELISA analysis. Thus, the expression construct of properdin C-terminally fused to mCherry was selected to further express protein for biochemical characterization and interactions studies.

Deglycosylation of properdin was achieved by incubation with the endoglycosidase PNGase F. Properdin contains a single *N*-linked glycosylation site in TSR6 and previous studies have demonstrated that deglycosylated recombinant properdin was about 5 kDa smaller than plasma properdin [66]. In contrast to Endo Hf, which only resulted in partial deglycosylation of properdin, a shift of about 5 kDa was also observed by SDS-PAGE after cleavage with PNGase F indicating that properdin expressed in Freestyle 293F cells showed a degree of *N*-glycosylation similar to that of plasma properdin. Up to date the role of this glycosylation site it is not clear and properdin lacking *N*-glycosylation has been shown to form the normal distribution of oligomers and to have unaltered activity in the AP hemolytic assay [66]. However, *N*-linked glycosylation has been shown to be essential for correct folding of eukaryotic proteins and might play a role in properdin production. Since freeze-thawing cycles have been shown to convert native properdin into "activated" properdin, which consists of large amorphous aggregates that might form non-specific ionic interactions due to their high positive charge, storage conditions of properdin were analyzed [84]. Size exclusion chromatograms of properdin stored at 4°C for one week or properdin after one cycle of freeze-thawing revealed an identical elution profile to properdin directly after purification, excluding the presence of aggregates. Recombinantly produced properdin after tag removal showed a thermal unfolding temperature of about 57.4°C and an untypical CD spectrum lacking the common characteristics of secondary structural elements.

A similar spectrum has been previously reported with variable ellipticity signals in the far-UV and a constant maximum at approximately 230 nm, which was most likely due to the high content of disulfide bonds in the TSR domains [60]. Although such spectrum suggested the absence of significant classical secondary structure, TSR domains have been shown to have an antiparallel three-stranded fold [63]. It is well possible that the TSR fold does not result in CD spectra comparable to other types of β-stranded proteins.

Up to date, determination of a three-dimensional structure of properdin at an atomic resolution was not possible due to its heterogeneous oligomerization state. However, alternative attempts using EM and small angle X-ray scattering (SAXS) were carried out, and most recently, crystals of properdin monomer variants were reported [56], [62], [64]. In 2017, properdin with the single point mutation E244K was found in a patient with type II properdin deficiency and discovered to form monomeric properdin due to is compact conformation analyzed by SAXS [71]. Therefore, site-directed mutagenesis was used to introduce this single mutation in the expression construct of properdin fused mCherry. Contrarily to published observations, recombinantly produced E244K properdin mutant was not monomeric and showed an oligomerization distribution similar to that of wild-type properdin. Due to this unexpected outcome and due to the fact that it is difficult to obtain homogeneous fractionated properdin oligomers in large amounts, crystallization of properdin for structure determination was not pursued.

To evaluate the biologic activity of recombinant properdin and to investigate the different variants produced in the course of this thesis, a complement activity hemolytic assay was performed using sheep RBCs and compared to Factor P obtained from normal human serum. Similarly to human RBCs, sheep RBCs present sialic acid on their surface and are protected from complement mediated lysis when exposed to human serum in the absence of specific antibodies [288]. Upon treatment with neuraminidase, sheep RBCs activate the human AP in the presence of normal human serum with Mg-EGTA, which silences the CP and MBL pathways of complement. Complement activation, which results in lysis of RBCs, was measured by UV/Vis absorption due to an intensity increase at 415 nm of the samples as a result of released hemoglobin [100]. Addition of purified properdin variants, including properdin fused to mCherry or c-myc, cleaved properdin or E244K mutant to properdin-depleted serum in concentrations within the range of properdin present in plasma stimulated the lysis of desialylated sheep RBCs in a concentration-dependent manner. The highest level of sheep RBCs lysis was observed in the presence of cleaved properdin, whereas the opposite trend was determined in the presence of Factor P. Resurrection of cell lysis by addition of purified properdin to properdin-depleted serum was previously published,

and experiments performed with properdin and Factor H suggested that P activity on RBCs was regulated by both complement regulatory proteins [101].

Although evidence has accumulated that properdin may act as pattern-recognition molecule and initiate AP activity besides the stabilization of AP convertases, its role in glycan recognition remains unclear. In particular, its glycan binding specificity remains unknown. A glycan array screening was performed using a polyclonal anti-properdin antibody for detection to elucidate the glycan specificity of properdin and identify potential glycan PAMPs that may activate the complement AP via direct recruitment of properdin. However, the only fluorescence signals observed in the glycan array screening of purified properdin were also visible in the control array performed with the antibody used for detection in the absence of properdin. Although these results suggested that no binding of properdin to glycans occurs, previous studies have shown that properdin interacts with glycans, such as heparin, HS and sulfatide [90], [91]. NGLs derived from such GAGs as well as sulfatide were present in the arrays used, excluding thus the possibility that no properdin-specific glycans were not part of the micro array slides analyzed. Therefore, these negative results suggested that either properdin does not binding to glycans under the chosen conditions, that the glycan-binding affinity is beyond the detection limit or that properdin oligomerization could result in a scenario where the binding epitopes of anti-properdin antibody are not accessible.

Further glycan-interactions studies were performed in solution using STD-NMR. Difference STD-NMR spectra were recorded for both properdin fused to mCherry and properdin after tag removal in the presence of Arixtra, a synthetic heparin pentasaccharide, and besides showing that interaction occurred, these spectra also highlighted that mCherry does not interfere with binding. Therefore, properdin fused to mCherry could be potentially used to further investigate binding of properdin to apoptotic and necrotic cells, which was previously reported using fluorescence microscopy [81]. The fact that complement activation only occurs after neuramidase treatment of sheep RBCs raised the possibility that enzymatic removal of sialic acid of mammalian RBCs unmasks an AP-activating glycan epitope that is directly recognized by properdin. Since such an epitope would resemble or could be identical to the TF antigen formed by the Galβ1-3GalNAcα1 disaccharide, binding to properdin fused to mCherry was analyzed by STD-NMR but not signals were observed in the difference STD-NMR spectrum. Besides the highly charged Arixtra, binding of additional GAGs such as size-fractionated heparin dp4, dp6 and sequence-defined GAG oligosaccharides (HepMers) was also investigated and resonance peaks were observed in all difference STD-NMR spectra recorded. Although negative charge seems to have an effect on glycan binding to properdin, no resonances were observed in the difference STD-NMR spectrum of properdin with the charged polysialic acid dp5, indicating that a specific recognition process rather than

overall negative charge underlie binding of properdin to glycans. Further STD-NMR experiments of properdin with non-charged lactose and asialo-GM1 glycans showed no binding, demonstrating that charge is essential (but not sufficient) for glycan-properdin interactions. Overall, STD-NMR experiments showed that a specific set of glycans bind to properdin in solution and that factors like glycan charge, length and most likely sulfation are critical for the interaction. Nevertheless, caution needs to be taken when investigating charged glycans since properdin is highly positively charged and non-specific interactions with polyanions may occur.

Although properdin and FH play distinct and opposite roles in the regulation of the alternative complement pathway, both proteins have been shown to bind renal tubular epithelial cells via the same HS polysaccharides chains. Through recognition of different and non-overlapping HS epitopes, properdin can activate the AP on tubular cells via HS while FH can inhibit the AP on these same cells also via HS [289]. However, these two regulatory proteins do not always recognize the same glycans, as shown for sialic acid which interacts with FH but not with properdin [101]. In 2005, Yu *et al.* demonstrated via SPR experiments that these two regulatory proteins also bind to heparin [89], but the binding epitopes have not been identified. Analysis of the relative intensities of different STD-NMR signals can be used for epitope mapping, giving thus information about the relative position of the ligand with respect to the binding site on the target protein [290]. Difference STD-NMR spectra recorded for properdin and FH in combination with Arixtra were compared and the binding epitopes are highlighted (Figure 4.19).
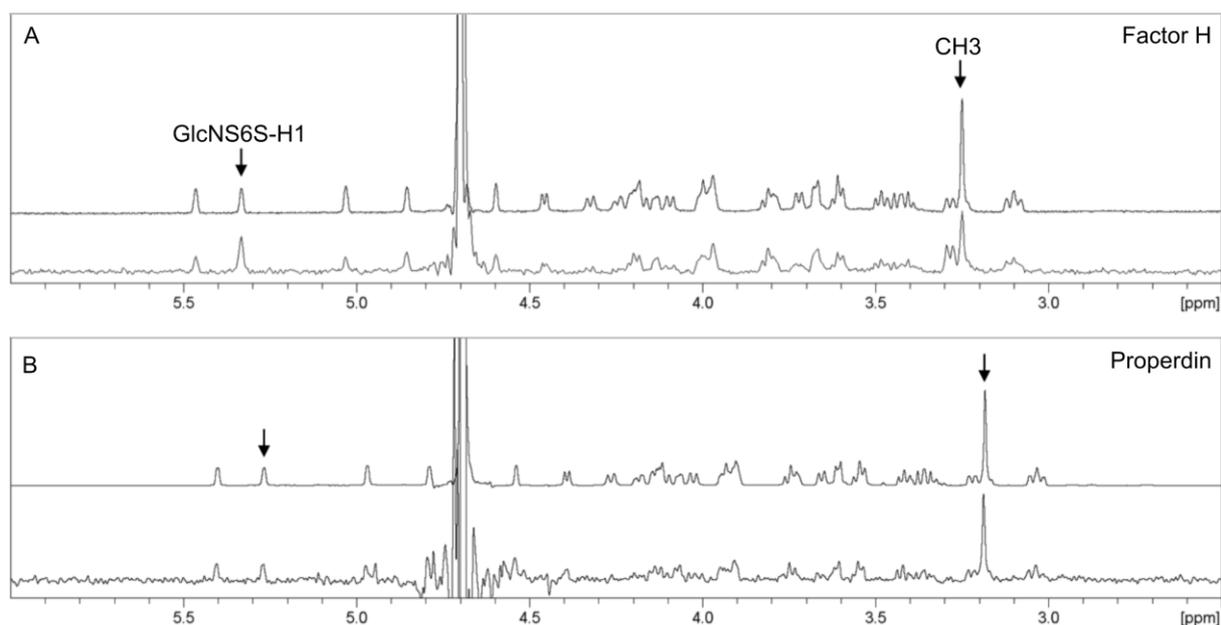
Figure 4.19 Comparison of STD-NMR spectra of purified properdin and Factor H with the highly positively charged heparin pentasaccharide, Arixtra. The truncated peak at 4.7 ppm corresponds to the truncated HDO signal. A) ^{1}H NMR reference spectrum of Arixtra, SD-NMR difference spectrum of 15.7 μM cleaved properdin with 1 mM Arixtra. B) ^{1}H NMR reference spectrum of Arixtra, STD-NMR difference spectrum of 7 μM Factor H with 1 mM Arixtra.

Differences in the relative intensities of STD-NMR signals suggest that binding of properdin and FH to heparin occurs via different epitopes, as previously observed for HS. For instance, the methyl group of Arixtra appears to be a key component of the binding of this pentasaccharide to properdin, while it is less prominent in the respective spectrum with FH. Likewise, while the GlcNS6S-H1 proton resonance is overrepresented in the FH STD-NMR difference spectrum when compared to the glycan's reference ^{1}H spectrum, suggesting an important contribution to the recognition, this is not the case for the respective spectrum obtained with properdin. These spectra confirm, at the atomic level, the assumption that while both regulators can bind to a set of related or even identical glycans they do so via recognition of different epitopes, and therefore likely also with different affinities. Although further studies are required to evaluate the effect of glycans tested on complement AP activation, STD-NMR experiments performed suggest that GAGs are potential properdin PAMPs and understanding their different structural characteristics that preferentially recruit properdin or Factor H could reveal novel ways to alter the level of AP activity for instance in the context of renal disease.

5 Project II

Production of recombinant heparinase I from *Pedobacter heparinus* for structure-based protein engineering

Contributions to this work

Ronja Pogan and Dr. Charlotte Uetrecht from the Heinrich Pette Institute, Hamburg, performed native and top-down mass spectrometry measurements and determined the molecular mass of commercial heparinase I and HepIv1. Most of the generated data has been published in the Master thesis of Ronja Pogan [291]. Based on the obtained results, a new heparinase construct, HepIv2, was designed. Sophie Stotz from the group of Dr. Hubert Kalbacher, University of Tübingen, analyzed the SDS-PAGE bands by matrix-assisted laser desorption/ionization mass spectrometry.

5.1 Results

5.1.1 Expression and purification of HepIv1 construct

To produce recombinant heparinase I from *P. heparinus* for structural determination by X-ray crystallography and GAG cleavage via enzymatic depolymerization, an expression construct was designed based on the expression constructs of heparinase I from *B. thetaiotaomicron*, which resulted in crystal structures [174]. This construct, designated HepIv1, lacked the native leader sequence (amino acids 1 to 21, UNIPROT entry Q05819) but comprised a His$_6$-tag followed by a thrombin cleavage site at the N-terminus to facilitate protein purification (the amino acid sequence is shown in section 9.1).

After initial expression tests in *E. coli* with varying IPTG concentrations and incubation temperatures, the best expressions conditions were selected. Expression of HepIv1 in *E. coli* BL21 (DE3) cells, induced with 1 mM IPTG and incubated overnight at 20°C, yielded about 3 g of cells per liter of LB medium. Harvested cells were disrupted via sonication, and although part of the expressed heparinase was found in the form of inclusion bodies, soluble protein was also obtained and used for purification.

Ni^{2+} affinity chromatography was used as a first purification step and His-tagged HepIv1 was purified from the crude lysate by applying a step gradient elution with increasing imidazole concentrations (Figure 5.1A). A wash step was performed with 50 mM imidazole to remove weakly bound contaminants and the majority of target protein eluted at 250 mM imidazole. On a SDS-PAGE, fractions from the main elution peak showed a similar level of purity with one major band at approximately 45 kDa (Figure 5.1B). This band corresponded to His-tagged HepIv1, for which a molecular weight of 43.6 kDa was calculated based on its primary sequence (Table 9.1). Fractions containing His-tagged HepIv1 were pooled, 1 unit of thrombin protease was added per 0.5 mg of target protein and the reaction mixture was dialyzed overnight against cleavage buffer containing 50 mM Tris-HCl pH 8, 150 mM NaCl, 2.5 mM CaCl$_2$ and 1 mM DTT. Before proceeding with reverse Ni^{2+} affinity chromatography, cleavage efficiency was analyzed by SDS-PAGE, which revealed the presence of a single band at a lower molecular weight (approximately 41.8 kDa) correspondent to digested HepIv1 (Figure 5.1C). The cleaved protein was separated from the His-tag and thrombin protease by collecting the flow-through of the Ni^{2+} affinity column connected to a HiTrap Benzamidine column. Using preparative size exclusion chromatography, a final purification step was performed to separate potentially aggregated species and to exchange the protein's buffer to a buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl and 1 mM TCEP (Figure 5.1D).
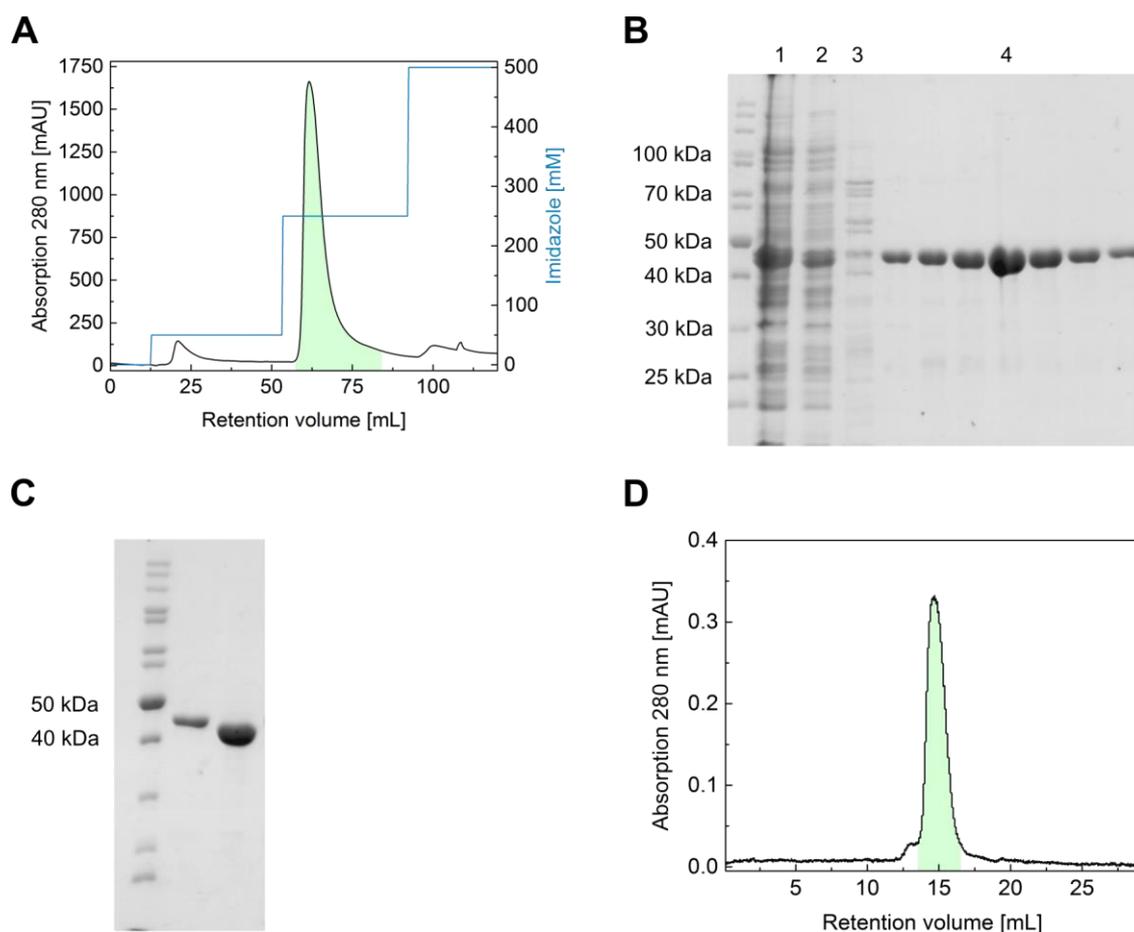
Figure 5.1 Purification of HepIv1 using Ni^{2+} affinity and size exclusion chromatography. A) Purification of His-tagged HepIv1 using a HisTrap 5 mL column. The target protein eluted at 250 mM imidazole. Pooled fractions are highlighted in green. B) SDS-PAGE analysis of (1) crude lysate, (2) flow-through collected during the loading phase, (3) fraction from the first peak obtained with 50 mM imidazole and (4) fractions from the elution peak obtained with 250 mM imidazole, containing His-tagged HepIv1. The band at approximately 45 kDa corresponds to monomeric His-tagged HepIv1. C) SDS-PAGE analysis of HepIv1 prior and after tag cleavage with thrombin. D) Preparative SEC of cleaved HepIv1 after removal of His-tag and thrombin. A Superdex 200 Increase 10/300 GL column was used and one major peak corresponding to monomeric HepIv1 was observed. Pooled fractions are shown in green.

SEC results showed one major elution peak at 14.7 mL, which corresponded to monomeric HepIv1, and a small peak at approximately 13 mL. This peak was found to correspond to dimeric HepIv1 by SEC when establishing its purification strategy. Consequently, a reducing agent, such as TCEP, was added to the final buffer to prevent potential non-native disulfide bond formation. Since size-homogeneous samples are an important parameter for successful protein crystallization, only fractions of the peak corresponding to monomeric protein were pooled, yielding a final amount of 26 mg of HepIv1.

Purified HepIv1 was concentrated and vapor diffusion ('sitting drop') crystallization drops were set using a crystallization robot and commercially available crystallization screens.

Despite multiple attempts with different crystallization conditions, incubation temperatures (4°C and 20°C) and starting material (protein with and without His-tag, addition of $CaCl_2$, different HepIv1 concentrations), no protein crystal formation occurred. In all crystallization trials, the crystallization drop remained either clear, precipitation occurred or salt crystals had formed (identified as inorganic salt by X-ray diffraction).

5.1.2 Biochemical and biophysical analysis of HepIv1

To identify possible parameters that could prevent protein crystallization and to obtain more information related to the structure and biophysical characteristics of HepIv1, biochemical and biophysical methods such as CD spectroscopy, analytical SEC, native MS and DSF were employed.

A CD spectrum of purified His-tagged HepIv1 was recorded and a minimum of ellipticity was observed at 215 nm (Figure 5.2A). To estimate the secondary structure content and to elucidate the fold of HepIv1, the algorithm BeStSel was used. By providing the measured mean residue weight ellipticity and the wavelength as input parameters, an experimental CD spectrum was calculated by the BeStSel server [263]. This spectrum was fitted to the linear combinations of pre-calculated and fixed basis spectra sets in order to estimate the secondary structure content of HepIv1. Antiparallel β-sheet (43.7%) was the major structural element and only a minor amount of α-helix was estimated to be present (1.7%).

Although a reducing agent was included during purification, two elution peaks could still be observed when His-tagged HepIv1 was analyzed on an analytical Superdex 200 increase column (Figure 5.2B). The two elution peaks eluted at 1.39 mL and 1.56 mL and showed that His-tagged HepIv1 was not homogenous, with two species, monomer and dimer, present in solution. Peak fractions that were collected during elution were analyzed again on the same column, and the resulting chromatogram obtained with the monomeric sample was identical to the first chromatogram of His-tagged HepIv1. In contrast, the chromatogram obtained with dimeric protein showed two identical elution peaks corresponding to both dimers and monomers of His-tagged HepIv1.

A

B
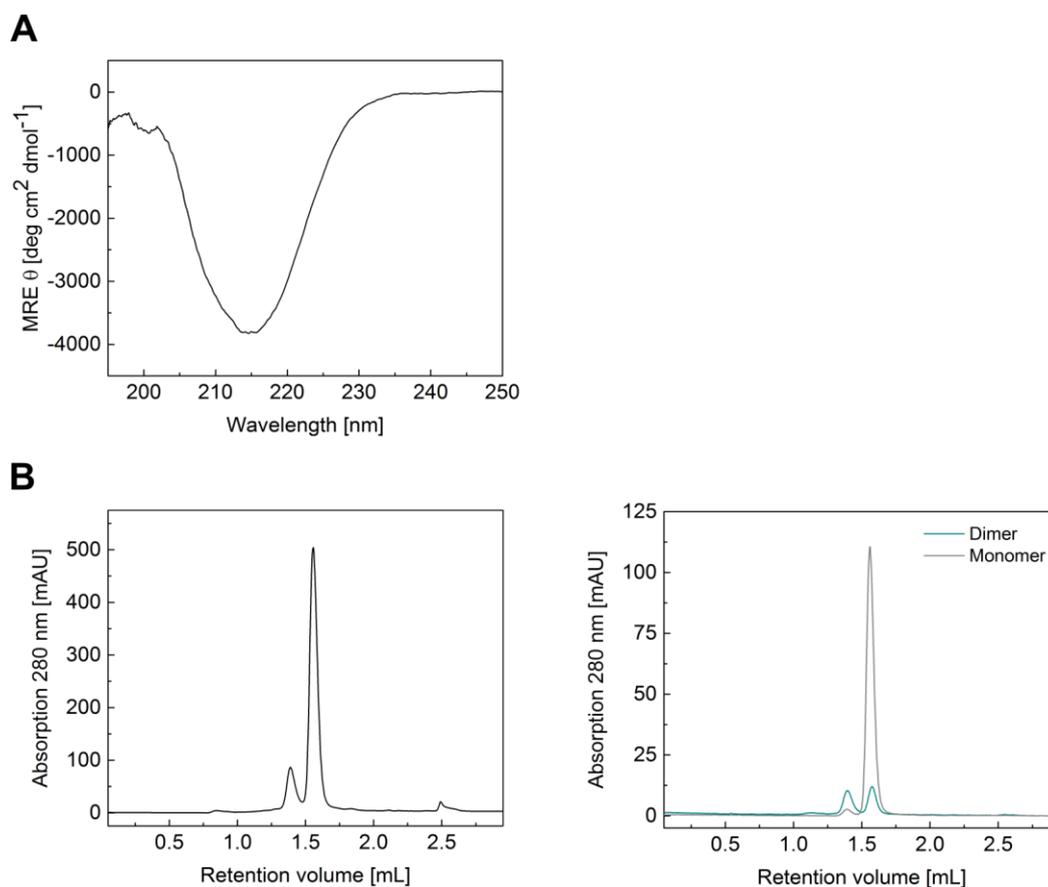
Figure 5.2 Biophysical and biochemical analysis of His-tagged HepIv1. A) CD spectrum of His-tagged HepIv1 (0.4 mg/mL) in buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl, 1 mM TCEP. The mean residue weight ellipticity is plotted as function of the wavelength with a minimum at 215 nm. B) Analysis of His-tagged HepIv1 by analytical SEC using a Superdex 200 increase PC 3.2/300 column. Panel on the left shows the chromatogram of His-tagged HepIv1 with two elution peaks at 1.39 mL and 1.56 mL corresponding to dimers and monomers species, respectively. Panel on the right shows the chromatograms obtained from isolated dimer and monomer elution peak fractions (depicted in teal and gray, respectively).

Since nucleation and subsequent crystal growth is time-dependent and often slow, the thermal stability of HepIv1 was analyzed over time. Aliquots of purified HepIv1 without His-tag were incubated at different temperatures (-80°C, 4°C, 20°C and 30°C) for one week and the samples were then analyzed via analytic SEC and SDS-PAGE (Figure 5.3).
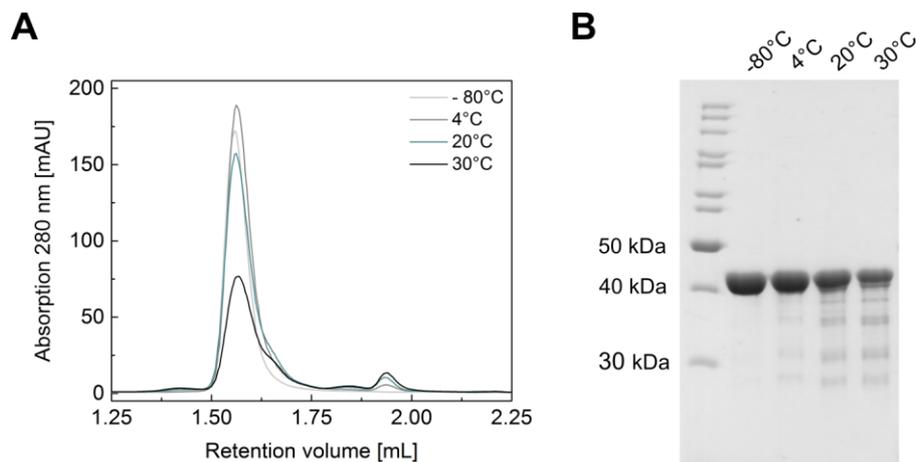
Figure 5.3 Thermal stability analysis of HepIv1. Samples of purified protein were incubated at -80°C, 4°C, 20°C and 30°C and analyzed after one week. A) Analytical SEC performed with a Superdex 200 increase PC 3.2/300 column. B) SDS-PAGE analysis. Protein bands visible at about 42 kDa correspond to intact HepIv1, whereas bands observed at smaller MWs most likely correspond to degradation fragments of HepIv1.

The single band observed for sample incubated at -80°C on the SDS-PAGE confirmed the purity of HepIv1 initially achieved during purification. However, elution peaks with lower intensities observed by SEC and the presence of bands at smaller molecular weights on the SDS-PAGE indicated that HepIv1 degradation occurred over time and with an increase of temperature. Since degradation of HepIv1 could have been caused by proteases that were not removed during purification, a protease inhibitor mix (cOmplete by Roche) as well as EDTA were added to purified HepIv1 to test this hypothesis. However, no change on the degradation effect was observed indicating that degradation of HepIv1 was caused by its intrinsically low thermal stability.

Next, native MS measurements were performed to further investigate the low stability of HepIv1. This technique preserves the quaternary protein structure and it has the ability to simultaneously measure and determine the molecular mass of several species present in a mixture [292]. Besides purified HepIv1, a commercial heparinase I purchased from Iduron was also analyzed in order to determine the native N-terminus of the protein isolated from *P. heparinus* after secretion as a possible help for improved construct design. For protein analysis by native MS, a Q-TOF 2 (Quadrupole time of flight) mass spectrometer was used and the buffers of both enzymes were exchanged to a volatile buffer containing 150 mM ammonium acetate pH 7.5. Multiple ion signals observed in the MS spectra originated from multiple charged species of heparinase I and HepIv1 and were plotted as function of mass-to-charge ratio (m/z values) of detected ions. Native MS measurements of heparinase I from Iduron showed signals correspondent to two different species with molecular masses of 42512.40 Da and 67106.87 Da (Figure 5.4A). Based on theoretical molecular weights, these species were identified to be monomeric heparinase I (mature protein has a MW of 41.3 Da)

and bovine serum albumin (BSA). Heparinase I from Iduron is supplied with 0.2% BSA for stabilization. Initial MS measurements of HepIv1, for which a theoretical MW of 41.75 kDa after His-tag cleavage has been determined (Table 9.1), were performed using a protein concentration of 2.5 µM. However, after 2 weeks of protein freezing and storage, a higher concentration of about 21 µM was necessary to obtain a good signal. For the native MS spectrum of HepIv1 prior to storage, signals corresponding to three different species were observed (Figure 5.4B). Molecular masses of 41750.77 Da, 41753.77 Da and 83811.6 Da were determined and identified as denatured HepIv1 and HepIv1 in the monomeric and dimeric states, respectively. Unfolded proteins display a larger surface area and therefore gain more charges, whereas less charges are clustered on proteins with more compact conformations [292]. Native MS spectrum of HepIv1 after storage showed a double monomer peak as well as an increase of Hepv1 denaturation and dimerization (Figure 5.4C). The molecular masses of both peaks corresponded to monomeric HepIv1 (41878.6 Da and 42130.50 Da), with one peak most likely resulting from a covalent attachment of a TCEP molecule to HepIv1 (the difference was about 250 Da, which corresponds to the molecular mass of TCEP). TCEP was included in the SEC buffer and it was therefore present in the HepIv1 solution prior to buffer exchange. The denaturation and dimerization of HepIv1 upon freezing and storage demonstrated once more the low stability of this enzyme. However, these species were not observed in native MS measurements of heparinase I from Iduron, indicating that this enzyme has a higher stability, which could either result from its native N-terminus sequence, the expression system or the presence of BSA.
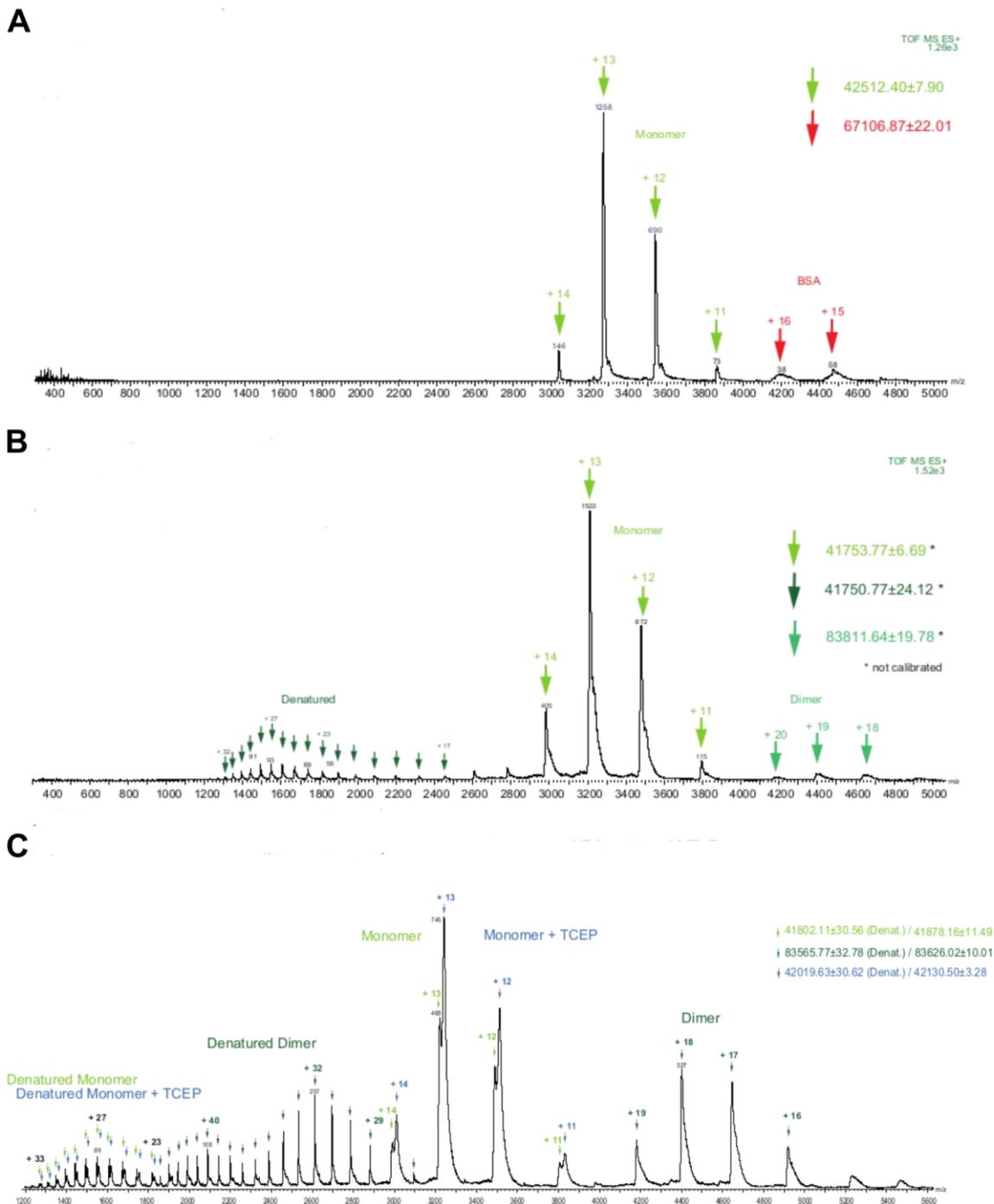
Figure 5.4 Native MS spectra of heparinase I purchased from Iduron and purified HepIv1. MS measurements were performed with protein in buffer containing 150 mM ammonium acetate at pH 7.5. Signal intensities are plotted as a function of m/z values of detected ions. The determined MW for each protein is indicated on the right side. Values highlighted with asterisks were determined without calibration of the system. A) Native MS spectrum of heparinase I (Iduron) at 3 µM. Assigned charge states are depicted in green for heparinase I and in red for BSA. B) Native MS spectrum of HepIv1 a 2.5 µM. Assigned charge states are depicted in dark geen for denaured HepIv1, light green for monomers of HepIv1 and intermediate green for dimers. C) Native MS spectrum of HepIv1 after storage in 150 mM ammonium acetate at pH 7.5 for two weeks at -20°C (21 µM). Assigned charged states are depicted in light green for denatured and folded HepIv1 monomer, in dark green for denatured and folded HepIv1 dimer and in blue for HepIv1 associated with TCEP. Spectra provided by Ronja Ponja and Dr. Charlotte Uetrecht, Heinrich Pette Institute, Hamburg.

DSF was used to determine the thermal stability, reflected by the melting temperatures (T_m values), of heparinase I and to investigate the difference in stability observed between the two enzymes. In addition to HepIv1 and heparinase I from Iduron, an additional commercial heparinase I purchased from R&D systems was investigated. Similar to HepIv1, this enzyme derived from *P. heparinus* was recombinantly produced in *E. coli* without the first 21 amino acids and possessed a His$_6$-tag at the N-terminus. Thermal unfolding curves of the three enzymes were determined by measuring the fluorescence signal of SYPRO ORANGE dye upon temperature increase (Figure 5.5A). Determination of T_m values for each protein was based on the first derivative analysis of the fluorescence curves (Figure 5.5B and Figure 5.5C).

A

C

Samples	T_m [°C]
HepIv1	31.60 ± 0.10 / 45.34 ± 0.16
HepI from Iduron	43.13 ± 0.04
HepI from R&D	32.38 ± 0.10

B
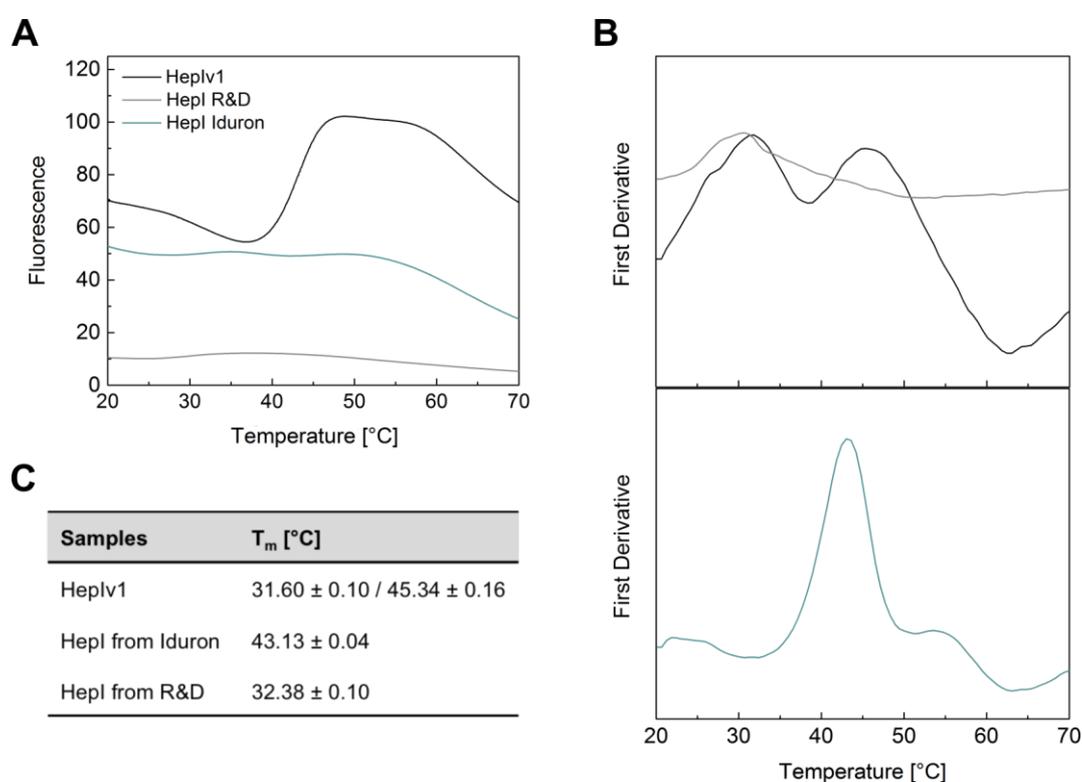
Figure 5.5 Thermal stability analysis of HepIv1 and commercially available heparinase I by DSF. Fluorescence intensities and first derivatives are plotted as a function of temperature. Results for HepIv1 in SEC buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl and 1 mM TCEP are depicted in black. For heparinase I from Iduron and R&D systems, results are shown in teal and gray teal, respectively. A) Thermal unfolding curves of HepIv1 and purchased enzymes. B) First derivative analysis obtained from the fluorescence profile. Due to differences in fluorescence intensity, results are shown in different panels for clarity. Results obtained for HepIv1 and heparinase I from R&D systems are shown at the top panel and results for heparinase I from Iduron are shown at the bottom panel. C) T_m values determined by DSF. Data shown represents mean values ± standard deviation of three measurements.

Thermal unfolding curves and their corresponding first derivative demonstrated that recombinantly heparinases produced (HepIv1 and heparinase I from R&D systems) had a lower thermal stability in comparison to heparinase I from Iduron isolated from *P. heparinus*.

The high initial fluorescence observed for HepIv1 suggested that unfolded protein was already present in solution, as shown by native MS. Two T_m values were determined for this enzyme, 31.60°C and 45.34°C, and they are most likely due to the presence of species with distinct thermal stability in solution. A T_m of 32.38°C was determined for heparinase I from R&D systems, whereas a significantly higher T_m of 43.13°C was determined for heparinase I from Iduron. Since buffer formulations, as well as stabilizing additives such as BSA, have a significant impact on how proteins behave in solution and can strongly influence their stability, a 96-well commercial buffer screen was used in combination with DSF to analyze the effect of different buffer systems, salt concentrations and pH values on the thermal stability of HepIv1. Purified and concentrated HepIv1 prepared in SEC buffer was diluted in solutions from the buffer screen and the unfolding curves were recorded (Figure 5.6A). By determining T_m values for each condition from the first derivative analysis, the effect of the formulation components was evaluated. Results showing a significant increase of thermal stability are displayed in Figure 5.6.

Buffer condition	T_m [°C]
0.25 M NaCl, pH 4	42.1
0.5 M NaCl, pH 4	42.8
1 M NaCl, pH 4	48.3
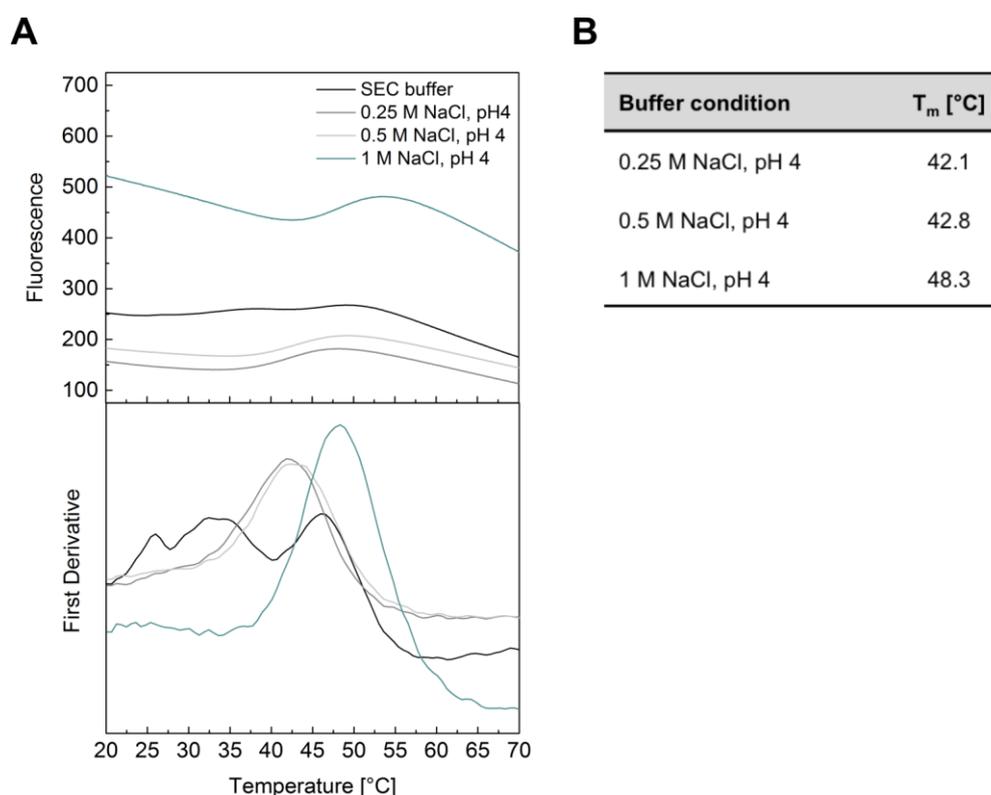
Figure 5.6 Effect of formulation components on the thermal stability of HepIv1. A) Unfolding curves of HepIv1 determined by DSF. HepIv1 was diluted in a Super Buffer composed of a citric acid:HEPES:CHES mixture with varying salt concentrations at pH 4 from a commercial screen (JBScreen Thermofluor FUNDAMENT). Conditions that resulted in a higher T_m are shown and compared to HepIv1 stored in SEC buffer (depicted in black). The bottom panel shows the first derivative analysis obtained from the fluorescence profile. Both panels are plotted as a function of temperature. B) Calculated T_m values for HepIv1 in buffers that result in higher thermal stability.

Buffer formulations containing low pH and high salt concentrations exhibited a single inflection point, as well as a significant increase of the thermal unfolding temperature. The best results were observed when using 100 mM of a Super Buffer composed of citric acid:HEPES:CHES in molar ratios of 2:3:4 in combination with 0.5-1 M NaCl at pH 4. A difference of about 10°C was observed for the T_m of HepIv1 in buffers containing 0.25 M and 0.5 M NaCl (42.1°C and 42.8°C, respectively) in comparison to HepIv1 in SEC buffer. Thermal stability of HepIv1 was further influenced by an increase of the salt concentration to 1 M NaCl, which resulted in a T_m of 48.3°C. Based on these findings, a modified SEC buffer was used for purification of HepIv1. However, high salt concentrations are sometimes not favorable for protein crystallization as they can promote the formation of salt crystals. Therefore, since lower salt concentrations (0.25 M to 0.5 M) showed similar thermal unfolding compared to heparinase I from Iduron, a buffer containing 20 mM of Super Buffer at pH 4, 250 mM NaCl and 1 mM TCEP was used. In SEC-based experiments it was observed that presence of a reducing agent increased the content of the monomeric species. A T_m of 41.85°C was determined by DSF for HepIv1 purified with the new buffer and the thermal unfolding curve was comparable to the previously reported results obtained with the commercial buffer screen.

A CD spectrum of HepIv1 after His-tag cleavage was recorded in the new buffer and a minimum of ellipticity was observed at about 215 nm (Figure 5.7A). This spectrum was identical to the spectrum reported for His-tagged HepIv1 in SEC buffer at pH 7.5, indicating that the buffer components did not have any measurable effect on the secondary structure content and folding of HepIv1. Additionally, analytic SEC was performed with two different concentrations of HepIv1 to analyze its influence on the molecular species present in solution (Figure 5.7B). Chromatograms obtained with 1 mg/mL and 2 mg/mL of HepIv1 showed a concentration-dependent effect on the oligomerization state of HepIv1. In contrast to the previously shown chromatograms, the major elution peak was observed at 1.42 mL, which corresponded to HepIv1 in the dimeric state. A peak at 1.59 mL corresponding to monomeric HepIv1 was also observed in both cases, but for HepIv1 at 2 mg/mL a third peak at lower retention volume was also present, indicating the formation of larger oligomers.
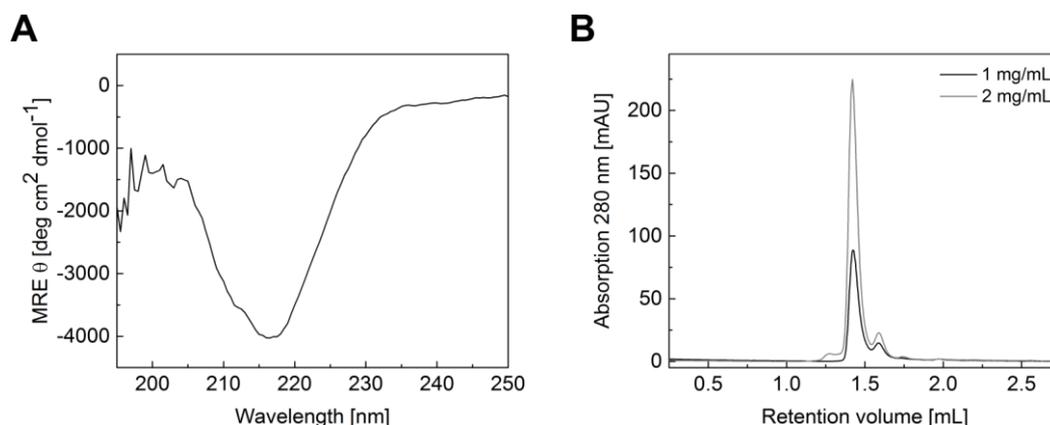
Figure 5.7 Biophysical and biochemical analysis of HepIv1 purified with the modified SEC buffer. A) CD spectrum of HepIv1 (0.4 mg/mL) in buffer containing 20 mM Super buffer, 250 mM NaCl and 1 mM TCEP at pH 4. The mean residue weight ellipticity is plotted as function of the wavelength. B) Analytical size exclusion chromatograms of HepIv1 obtained with protein concentrations of 1 mg/mL and 2 mg/mL (shown in black and in gray, respectively). On a Superdex 200 increase PC 3.2/300 column three elution peaks were observed at 1.27 mL, 1.42 mL and 1.59 mL with the major peak corresponding to a dimer of HepIv1.

Despite the increase of the T_m value of HepIv1 with the modified SEC buffer, protein degradation could still be observed by SDS-PAGE and size heterogeneity was still a challenge for crystallization. Nevertheless, this protein was used to set up new crystallization drops using different conditions, but no protein crystals were observed.

5.1.3 Design of HepIv2 construct

The stability difference observed between HepIv1 and heparinase I purchased from Iduron suggested that their N-terminal sequence might differ and influence their biochemical properties. Studies have been performed to precisely identify the native leader sequence of heparinase I from *P. heparinus*, however it is still not clear where cleavage of the secretion signal occurs [164]. To determine the accurate mass of commercially available heparinase I different MS techniques were employed. Heparinase I from Iduron possessing the native secretion N-terminus sequence, was investigated by top-down analysis both in native and denatured conditions. Prior to top-down analysis measurements, the protein's buffer was exchanged to 150 mM or 250 mM ammonium acetate at pH 7.4. A molecular mass of 42501.22±1.3 Da was determined for the major protein species detected using an Orbitap Fusion mass spectrometer. Measurements performed with a Q-TOF 2 mass spectrometer showed an almost identical molecular mass of 42504.9±1.3 Da. Considering that the theoretical molecular weight of heparinase I lacking the first 21 amino acids is 41337.63 Da, the difference in molecular masses could be derived from 10 additional amino acids at the N-terminus sequence of mature heparinase I isolated from *P. heparinus*. The theoretical MW of heparinase I containing these additional 10 amino acids located at the N-terminus

(QQLFLCSAYA) was determined to be 42462.94 Da. The remaining molecular mass difference between could result from an additional modification, such as the substitution of histidine to valine (38 Da) or acetylation (42 D). Based on these findings from Ronja Ponja and Dr. Charlotte Uetrecht, a new construct of heparinase I containing the additional 10 amino acids at the N-terminus was designed to study their impact on the enzyme's stability. No further potential modifications were included.

5.1.4 Expression and purification of HepIv2 construct

The new heparinase construct, named HepIv2, was expressed using the pET28a expression vector with the extended N-terminus sequence containing a part of the putative leader sequence of native heparinase I from *P. heparinus* (residues 12-21, UNIPROT entry Q05819). As for the HepIv1 construct this construct also encoded a N-terminal His$_6$-tag followed by a cleavage sequence specific for the HRV 3C protease (the amino acid sequence is shown in section 9.1).

During initial expression and purifications attempts using the same experimental procedures as described for HepIv1 two major challenges were encountered. First, HepIv2 exhibited a very low solubility, which resulted in very small amounts of soluble protein after cell lysis and a high propensity to precipitate during purification. Second, a high level of contaminants was observed after the first purification step using Ni^{2+} affinity chromatography. Although additional purification steps were performed some impurities could not be removed. To improve solubility of HepIv2 and increase the amount of soluble protein for purification, different expression conditions were tested, including different *E. coli* strains, cell culture media and starting material (glycerol stock or freshly transformed cells). Best results (amount of target protein vs. amount of impurities) were obtained using *E. coli* BL21 (DE3) cells, induced with 0.1 mM IPTG overnight at 20°C. In addition, lysis buffers with varying salts (150 mM and 300 mM NaCl and KCl), pH values (7.0, 7.5 and 8.0) and additives that influence protein solubility and reduce aggregation, such as Triton X-100, Urea and L-Arginine, were tested [293]. From this buffer screening approach, a buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 5 mM β-mercaptoethanol, 125 mM L-Arginine-HCl and 0.01% (w/v) Triton X-100 was selected as lysis buffer. Cell disruption was also modified by using cell homogenization at a pressure of 700-800 bar instead of sonication. This new strategy contributed to a minor increase of soluble HepIv2, which was used for purification.

Ni^{2+} affinity chromatography was selected as a first purification step using a step elution gradient with increasing concentrations of imidazole (50mM, 100 mM, 250 mM and 500 mM) to potentially increase the purity of HepIv2 (Figure 5.8A). Prior to elution, a washing step was

performed with 1 M NaCl and 20% (v/v) glycerol to disrupt non-specific interactions formed between contaminants and HepIv2. Collected fractions were analyzed by SDS-PAGE and in contrast to previous purifications the amount of impurities was significantly reduced (Figure 5.8B). One major band was observed at approximately 45 kDa, which corresponded to the MW of His-tagged HepIv2 (Table 9.1). Fractions from elution steps performed with 250 mM and 500 mM imidazole were pooled and further purified. Since neither removal of the His-tag neither hydrophobic interaction nor size exclusion chromatography improved purity, cation exchange chromatography was selected as a second purification step. HepIv2 has a very high theoretical pI of 9.2 and it is thus expected to be positively charged at a pH of 8 (Table 9.1). Prior to purification, the buffer was exchanged with to reduce the salt concentration to 20 mM NaCl. Two washing steps were performed using 50 mM and 100 mM NaCl followed by a linear elution gradient up to 1 M NaCl to elute bound HepIv2 (Figure 5.8C). Resulting chromatogram showed that no significant separation occurred, which was confirmed by SDS-PAGE (data not shown). Thus, most of the elution peak fractions were pooled for biochemical analysis and crystallization. Analytic SEC performed using a buffer containing 50 mM HEPES pH 8, 150 mM NaCl, 5 mM β-mercaptoethanol showed that on a Superdex 75 PC 3.2/300 column, HepIv2 eluted at 1.20 mL (Figure 5.8.D) This elution volume corresponded to a molecular weight of 37.5 kDa, which indicated the presence of monomeric His-tagged HepIv2.
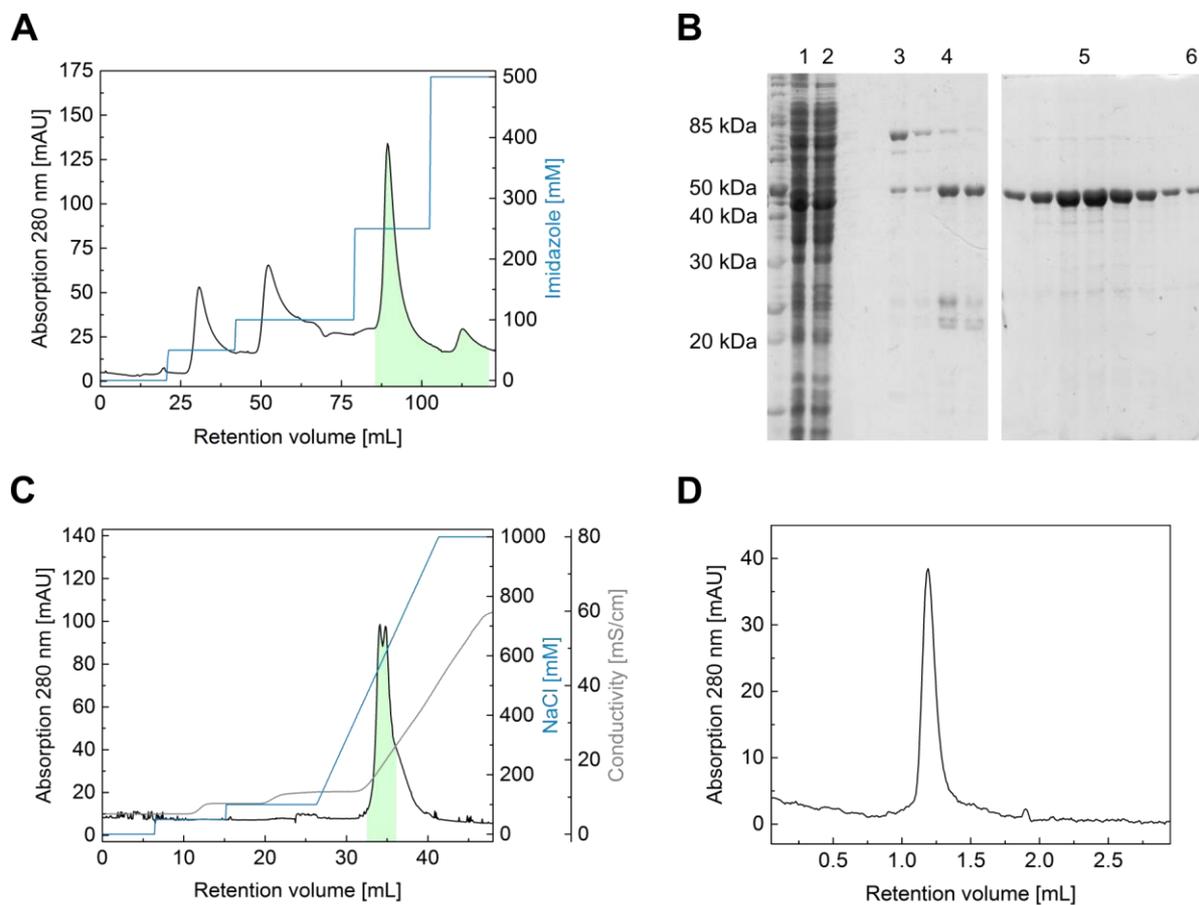
Figure 5.8 Purification of HepIv2 using Ni^{2+} affinity and ion exchange chromatography. A) Chromatogram of His-tagged HepIv2 purification using a 5 mL HisTrap column. A major elution peak was observed with 250 mM Imidazole and pooled fractions are highlighted in green. B) SDS-PAGE analysis of (1) crude lysate, (2) flow-through collected during loading phase, (3) elution with 50 mM imidazole, (4) 100 mM imidazole, (5) 250 mM and (6) 500 mM Imidazole. The major band at approximately 45 kDa corresponds to His-tagged HepIv2. C) Cation exchange chromatography performed with a 1 mL HiTrap SP Sepharose column. A linear elution gradient was applied from 100 mM to 1 M NaCl to elute HepIv2. Fractions that were pooled are shown in green. D) Size exclusion chromatogram obtained using a Superdex 75 PC 3.2/300 column. A single peak was observed at 1.20 mL corresponding to monomeric His-tagged HepIv2.

The HepIv2 purification was thus optimized to yield pure and monomeric proteins, however only small amounts of pure HepIv2 were obtained due to its poor solubility. On average, from a cell pellet of 10 g, around 2 mg of protein were obtained after the first Ni^{2+} affinity purification step and less than 1 mg of purified His-tagged HepIv2 was obtained after the second, ion exchange chromatography (IEX), purification step. During HepIv2 concentration attempts using centrifugal concentrators, protein precipitation as well as a tendency of His-tagged HepIv2 to adhere to the membrane of the concentrator were observed, further decreasing the protein yield.

5.1.5 Analysis of HepIv2 thermal stability

The impact of the extended N-terminus sequence on the thermal stability was analyzed by DSF and SDS-PAGE using purified His-tagged HepIv2. Initial thermal unfolding curves were determined for HepIv2 samples obtained after each purification step by monitoring the fluorescence of the Thermo Fisher Protein Thermal Shift dye upon steady temperature increase (Figure 5.9A). For each sample a T_m (mean values ± standard deviation of three measurements) was determined using the first derivative analysis. Similar thermal unfolding curves were observed for both samples (after Ni^{2+} affinity and IEX) and no increase in stability occurred in comparison to HepIv1. A T_m of 27.2±0.7°C was determined for pooled fractions containing HepIv2 after Ni^{2+} affinity chromatography, while a T_m of 31.7±0.4°C was determined for pooled fractions after cation exchange chromatography containing a higher amount of NaCl and no imidazole. These results showed once more the effect of buffer components on a protein's stability. Although a stability screen from Hampton Research was used to further increase the thermal stability of HepIv2, T_m values determined for the different conditions did not show a significant increase compared to T_m determined for HepIv2 after cation exchange chromatography. Further analysis as well as crystallization attempts were performed using purified His-tagged HepIv2 that was dialyzed into a buffer containing 50 mM HEPES, 150 mM NaCl, 2 mM DTT at pH 7.5. Aliquots of HepIv2 were incubated for one week at different temperatures (-80°C, 4°C and 20°C) and subsequently analyzed by SDS-PAGE (Figure 5.9B). In contrast to previous results, only faint bands were observed at lower molecular weights but reduced intensity of the band correspondent to HepIv2 at 20°C indicated that protein degradation nevertheless occurred.
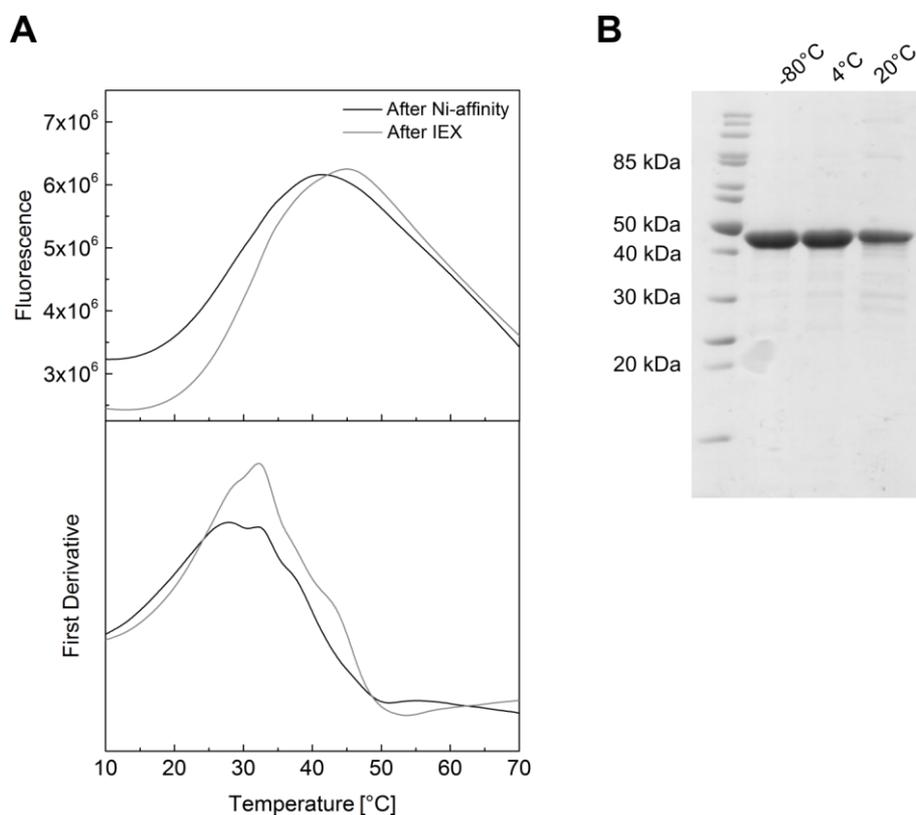
Figure 5.9 Thermal stability analysis of His-tagged HepIv2. A) Thermal unfolding curves of HepIv2 determined by DSF. First derivative analysis of the fluorescence profile is shown in the bottom panel. All curves are plotted as function of temperature. Samples of His-tagged HepIv2 were analyzed after first and second purification steps (Ni-affinity and IEX depicted in black and in gray, respectively). B) SDS-PAGE analysis of HepIv2 samples incubated for one week at -80°C, 4°C and 20°C. The major band at approximately 45 kDa corresponds to intact HepIv2, whereas faint bands with smaller MWs likely resulted from protein degradation.

The temperature effect on degradation of heparinase I was more severe for HepIv1 than for HepIv2, however similarly to HepIv1, crystallization attempts of His-tagged HepIv2 did not result in protein crystal formation.

5.1.6 Expression and purification of HepIv3 construct

A third expression construct, designated HepIv3, was designed based on a publication where the crystal structure of recombinant heparinase III from *P. heparinus* was reported [160]. The pET21b expression vector was used and besides lacking the putative leader sequence (i.e. with the N-terminus chosen as for HepIv1), this construct was designed with a His$_6$-tag at the C-terminus without a cleavage site between the protein sequence and the tag (the amino acid sequence is shown in section 9.1).

A strategy similar to that used for the production of HepIv2 was followed to express and purify HepIv3. A main difference consisted in the addition of 5 mM CaCl$_2$ to all buffers since

this additive was reported to aid in the stabilization of heparinase I [294]. During expression in *E. coli* BL21 (DE3) cells, samples were taken at different time points for SDS-PAGE analysis, as well as after cell harvest and lysis (Figure 5.10A). Besides the band at 42 kDa, possibly corresponding to expressed HepIv3 (Table 9.1), additional bands were observed at lower molecular weights. The intensity of these bands increased simultaneously with increasing expression of the target protein, which could be a result of co-expression, protein degradation or partial truncation of HepIv3 during protein expression. Ni^{2+} affinity chromatography was selected to purify the crude cell lysate containing soluble HepIv3 and verify if the potential contaminants or degradation fragments with smaller molecular weights could be separated (Figure 5.10B). Two washing steps with 25 mM and 50 mM imidazole were performed followed by a linear elution gradient up to 500 mM imidazole. A single elution peak was observed on the chromatogram and fractions from this peak were pooled for further purification. The protein buffer was exchanged via a gravity-flow PD-10 desalting column to decrease the salt concentration to 20 mM NaCl before proceeding with cation exchange chromatography. In this second purification, a linear elution gradient up to 1 M NaCl was applied and, as in the previous step, a single elution peak was observed (Figure 5.10C). SDS-PAGE analysis of pooled fractions from both steps revealed that separation of bands at lower molecular weights could not be achieved and that the purity of HepIv3 was not improved by cation exchange chromatography (Figure 5.10D). A western blot was also performed with the same samples to find out if the observed bands corresponded to co-expressing and co-purifying contaminant proteins or if they were a result of protein degradation (Figure 5.10E). Using an anti-His antibody for detection, a similar set of bands was observed, which indicated the presence of full-length HepIv3 as well as partially cleaved HepIv3 with an intact C-terminal sequence.
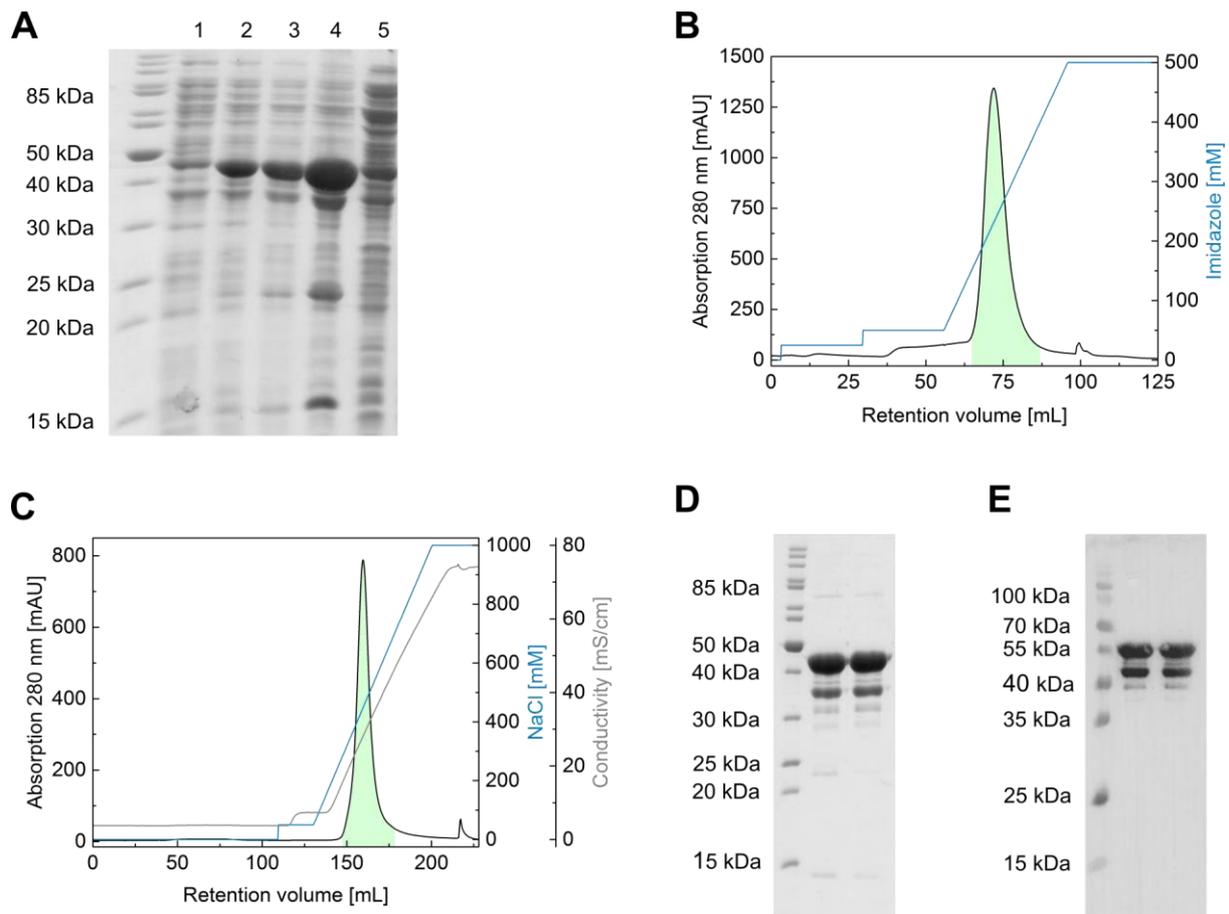
Figure 5.10 Expression and purification of HepIv3 using Ni^{2+}-affinity and cation exchange chromatography. A) SDS-PAGE analysis of HepIv3 expression and cell lysis. Bands visible at approximately 42.5 kDa correspond to the molecular weight of HepIv3. (1) Before induction with IPTG, (2) 3.5 h after induction, (3) overnight growth, (4) solubilized cell pellet after lysis and (5) soluble fraction. B) Purification of His-tagged Hep1v3 using a HisTrap 5 mL column. After an initial washing step with low concentrations of imidazole, a linear elution gradient from 50 mM to 500 mM imidazole was applied to elute HepIv3. Pooled fractions are highlighted in green. C) Cation exchange chromatography used as a second purification step. A linear elution gradient was applied up to 1 M NaCl to elute HepIv3. Pooled fractions are highlighted in green. D) SDS-PAGE and E) Western blot analysis of pooled fractions containing HepIv3 from first (1) and second (2) purification steps. A composite figure was generated by overlapping results from both chemiluminescent and colorimetric detections with the Image Lab Touch Software (Bio-Rad).

To identify HepIv3, the main bands observed by SDS-PAGE were analyzed by MALDI-MS. The results obtained showed that the protein band observed at an MW of approximately 42.5 kDa corresponded to the full-length protein. For this band, a sequence coverage of 79.8% was determined and both N-terminal and C-terminal sequence stretches were identified (Figure 5.10A). In contrast, the protein band observed around 36 kDa with sequence coverage of 77.2% showed that the N-terminus sequence of HepIv3 was missing (Figure 5.10B).

A Sequence data:

HepI

| Intensity Coverage: | 71.3 % (577133 cnts) | Sequence Coverage MS: | 79.8% |
| Sequence Coverage MS/MS: | 0.0% | pI (isoelectric point): | 9.7 |

```
        10         20         30         40         50         60         70         80         90        100
MQQKKSGNIP YRVNVQADSA KQKAIIDNKW VAVGINKPYA LQYDDKLRFN GKPSYRFELK AEDNSLEGYA AGETKGRTEL SYSYATTNDF KKFPPSVYQN

       110        120        130        140        150        160        170        180        190        200
AQKLKTVYHY GKGICEQGSS RSYTFSVYIP SSFPDNATTI FAQWHGAPSR TLVATPEGEI KTLSIEEFLA LYDRMIFKKN IAHDKVEKKD KDGKITYVAG

       210        220        230        240        250        260        270        280        290        300
KPNGWKVEQG GYPTLAFGFS KGYFYIKANS DRQWLTDKAD RNNANPENSE VMKPYSSEYK TSTIAYKMPF AQFPKDCWIT FDVAIDWTKY GKEANTILKP

       310        320        330        340        350        360        370        380
GKLDVMMTYT KNKKPQKAHI VNQQEILIGR NDDDGYYFKF GIYRVGNSTV PVTYNLSGYS ETARLEHHHH HH
```

B Sequence data:

HepI

| Intensity Coverage: | 71.4 % (222953 cnts) | Sequence Coverage MS: | 77.2% |
| Sequence Coverage MS/MS: | 0.0% | pI (isoelectric point): | 9.7 |

```
        10         20         30         40         50         60         70         80         90        100
MQQKKSGNIP YRVNVQADSA KQKAIIDNKW VAVGINKPYA LQYDDKLRFN GKPSYRFELK AEDNSLEGYA AGETKGRTEL SYSYATTNDF KKFPPSVYQN

       110        120        130        140        150        160        170        180        190        200
AQKLKTVYHY GKGICEQGSS RSYTFSVYIP SSFPDNATTI FAQWHGAPSR TLVATPEGEI KTLSIEEFLA LYDRMIFKKN IAHDKVEKKD KDGKITYVAG

       210        220        230        240        250        260        270        280        290        300
KPNGWKVEQG GYPTLAFGFS KGYFYIKANS DRQWLTDKAD RNNANPENSE VMKPYSSEYK TSTIAYKMPF AQFPKDCWIT FDVAIDWTKY GKEANTILKP

       310        320        330        340        350        360        370        380
GKLDVMMTYT KNKKPQKAHI VNQQEILIGR NDDDGYYFKF GIYRVGNSTV PVTYNLSGYS ETARLEHHHH HH
```

Figure 5.11 MALDI-MS analysis of HepIv3 bands observed by SDS-PAGE. Gray bars and red residues denote areas covered by analysis of the tryptic digest. A) Analysis of protein band with the highest intensity observed at approximately 42.5 kDa. B) Analysis of protein band at around 36 kDa. Results were provided by Sophie Stotz, University of Tübingen.

Taken together, both western blot and MALDI-MS results confirmed the presence of full-length and partial fragments of HepIv3. Furthermore, these results showed that expression of HepIv3 construct resulted in partial N-terminal truncation of heparinase I. Expression of HepIv3 was repeated without CaCl$_2$ but no increase in stability was observed.

5.1.7 Activity assay of heparinase I and size-fractionation of oligosaccharides

An enzymatic assay was performed to determine the catalytic activity of recombinantly produced HepIv1. This enzyme can be used to obtained small oligosaccharides from heparin via an enzymatic depolymerization reaction that proceeds via β-elimination, producing a carboxy-conjugated C-C double bond that can be detected at 232 nm [141]. Therefore, mixtures containing HepIv1 and sodium heparin as substrate were prepared and incubated at 35°C. Formation of the C4-C5 double bond at the non-reducing end of heparin-derived oligosaccharide products was measured by UV absorption at 232 nm, with heparinase I from R&D systems serving as a positive control. HepIv1 used in this assay was produced according to the experimental procedure developed for HepIv2 and stored at 4°C in buffer containing 50 mM HEPES pH 7.5, 150 mM NaCl and 2 mM DTT. For the assay a buffer

containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 mM CaCl$_2$ was used, except when studying the effect of Ca^{2+} on HepIv1's activity.

The enzymatic depolymerization rate of heparin by HepIv1 was first investigated with varying concentrations of CaCl$_2$ (2 mM, 5 mM and 10 mM) since Ca^{2+} is known to enhance activity of heparinase I (Figure 5.12A) [145]. Increasing absorption at 232 nm suggested that HepIv1 was active and oligosaccharides products derived from heparin were produced. From the three concentrations of CaCl$_2$ tested, the highest activity was observed for reactions prepared with 10 mM CaCl$_2$. Thus, an assay buffer containing this concentration of CaCl$_2$ was selected for further experiments, in which the activity of HepIv1 was compared to the activity of commercial heparinase I from R&D systems (Figure 5.12B). Since the concentration of heparinase I from R&D systems was not known, two concentrations of HepIv1 were tested (2 µg/mL and 20 µg/mL) and a control reaction containing substrate and buffer was prepared to determine background absorption. UV absorption recorded at 232 nm showed that the concentration of heparinase I had a significant influence the rate of enzymatic depolymerization of heparin. Whereas reaction incubated with 2 µg/mL of HepIv1 showed a very low rate of heparin degradation, an increase of UV absorption was observed when using HepIv1 at 20 µg/mL. However, while a continuous linear increase of the UV absorption was observed for heparinase I from R&D, a stagnation in the activity of the higher concentrated HepIv1 sample occurred after a few minutes of reaction time despite the presence of a large excess of substrate, possibly as a consequence of the thermal instability of HepIv1.

Heparin-derived oligosaccharides obtained with both reactions were size-fractionated on a Superdex Peptide PC 3.2/300 column using assay buffer containing 10 mM CaCl$_2$ (Figure 5.12C). Separation was monitored by UV absorption at 232 nm and the observed elution peaks corresponded to oligosaccharides with different sizes/degrees of polymerization (dp) [4]. Both reactions yielded the main disaccharide (dp2) product (peak at approximately 1.55 mL). However, only heparin degradation by heparinase I from R&D produced longer oligosaccharides, such as dp4 and dp6, while a large amount of non-cleaved substrate was observed for the HepIv1 reaction in the void volume peak.
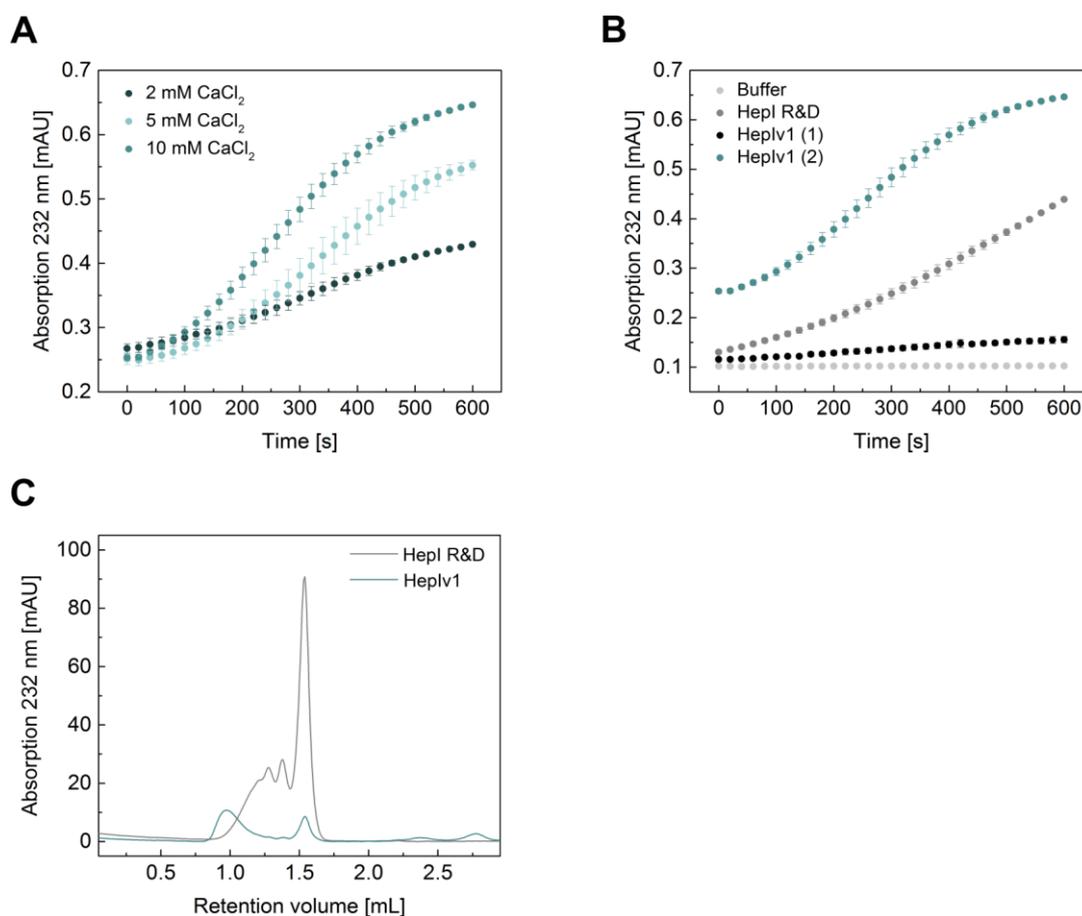
Figure 5.12 Enzymatic depolymerization of heparin by heparinase I monitored by UV absorption at 232 nm. Heparinase was incubated with sodium heparin at 0.75 mg/mL and incubated at 35° for 10 min. A) Effect of Ca^{2+} concentration on the catalytic activity of purified HepIv1 at a protein concentration of 20 µg/mL. Concentrations of 2 mM (dark teal), 5 mM (light teal) and 10 mM (intermediate teal) of $CaCl_2$ were tested. B) Comparison of catalytic activity of heparinase I from R&D systems and HepIv1. A substrate blank with buffer and substrate was included to determine the background absorbance (light gray). Reaction of commercial heparinase I and HepIv1 at concentrations of 2 µg/mL (1) and 20 µg/mL (2) are shown in dark gray, black and teal, respectively. C) Separation of heparin products released by both enzymes on a Superdex Peptide PC 3.2/300 column after a reaction incubation period of 1 h. The assay buffer containing 50 mM Tris-HCl pH 7.5, 100 mM NaCl and 10 mM $CaCl_2$ was used for elution.

Although a quantitative comparison of enzymatic cleavage of heparin by these two enzymes cannot be performed since the concentration of the commercial product was unknown, the obtained results suggested that recombinantly purified HepIv1 was less active than commercial heparinase I from R&D systems.

5.2 Discussion

Glycosaminoglycans such as heparin and HS have been implicated in diverse biological activities through interaction with proteins. However, chemically defined and pure GAGs that can be used for functional and structural studies are difficult to obtain. A common strategy involves the enzymatic depolymerization of heparin by heparinase I from *P. heparinus* [4], [104]. Up to date, although the soil bacterium *P. heparinus* is the major source of heparinases, the only available three-dimensional structure of heparinase I is from *B. thetaiotaomicron* [174]. Despite sharing a sequence identity of 66.3%, heparinase I from *B. thetaiotaomicron* differs from heparinase I from *P. heparinus* by the absence of an apparent signal peptide at the N-terminus and an insertion of 8 amino acids within the heparin binding domain and calcium coordinating motif (Figure 5.13) [295]. This insertion is part of the tip of the so-called thumb domain in the *B. thetaiotaomicron* heparinase I and it has been shown to influence the enzyme's activity. Distinct amino acids were identified to be critical for the catalytic activity of both enzymes, with only a catalytic base histidine residue conserved among them [170], [174].
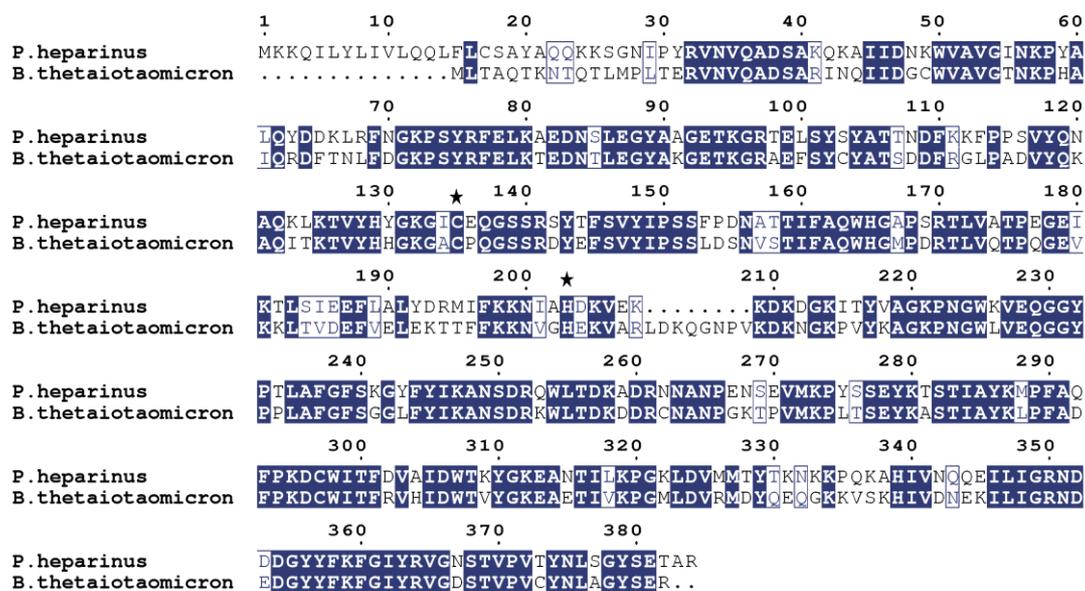
Figure 5.13 Sequence alignment of native heparinase I from *P. heparinus* and heparinase I from *B. thetaiotaomicron*. Only heparinase I from *P. heparinus* is predicted to contain a putative leader sequence. An insertion of 8 amino acids observed in the sequence of heparinase I from *B.* thetaiotaomicron, which corresponds to the tip of the thumb domain. Conserved residues are highlighted in blue and the catalytic critical amino acids C135 and H203 are marked with asterisks. Sequence alignment was carried out using Clustal Omega and displayed with ESPript [296], [297].

Although recombinant heparinase I expressed in *E. coli* cells is prone to aggregation and formation of insoluble inclusion bodies, expression conditions used in this study yielded soluble heparinase I protein that could be used for purification [298]. For the purification of recombinant HepIv1 lacking the signal peptide, a strategy comprising two Ni^{2+} affinity

chromatography steps, tag removal and a final polishing step using SEC resulted in highly pure mature HepIv1. Size exclusion chromatograms revealed, however, the presence of two species in solution, which corresponded to monomeric and dimeric HepIv1.

Since dimerization of HepIv1 could be a result of non-native disulfide bonds formation between free cysteines in the protein, the sequence of heparinase I was analyzed and two cysteine residues (C135 and C297) were found in the predicted mature protein. Sequence alignment of heparinases I from *P. heparinus* and *B. thetaiotaomicron* showed that these two cysteine residues were conserved in the sequence of the homolog, which has in total six cysteines (Figure 5.13). To estimate the position and orientation of these cysteines, a homology model of heparinase I from *P. heparinus* was calculated via the SWISS Model server using the three-dimensional crystal structure of its homologue, which folds into a β-jelly roll with extended loops and an inserted partially helical domain, as a template (Figure 5.14) [299].
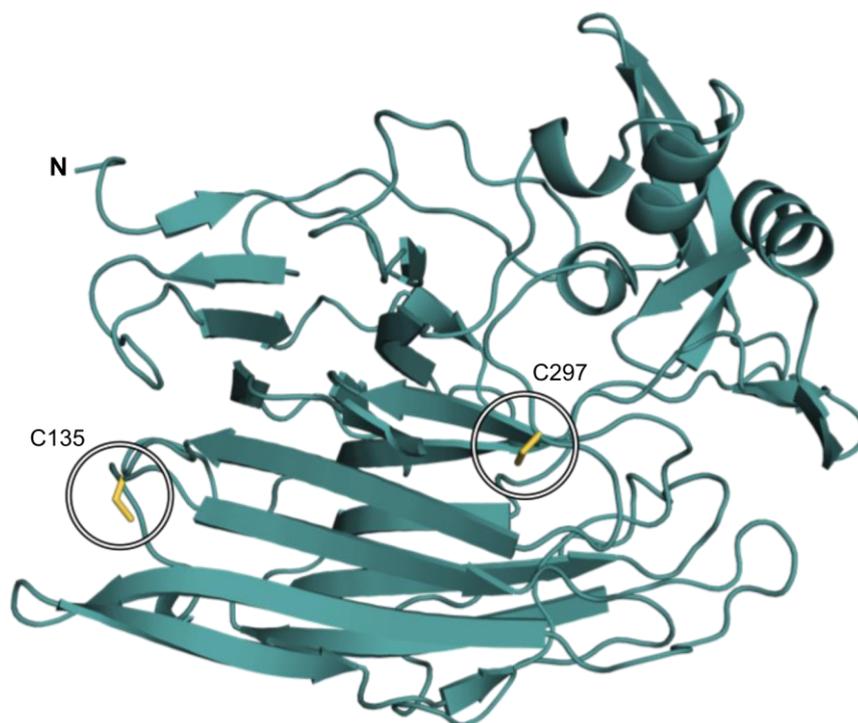
Figure 5.14 Homology model of mature heparinase I from *P. heparinus*. The three-dimensional structure of heparinase I from B. *thetaiotaomicron* (PDB code 3IKW) was used as a template to predict the homology model with the SWISS Model server [299]. The N-terminus is indicated and C135 and C297 are highlighted in yellow.

In the predicted structure, C135 and C297 were not located in close proximity, which makes the possibility of intramolecular disulfide bond formation unlikely. However, both residues were predicted to be exposed at the surface of the protein and might moderate intermolecular disulfide bonds between HepIv1 molecules. Reducing agents such as DTT

and TCEP were thus included in the end of the purification to prevent non-native disulfide bond formation and to reduce the sample heterogeneity for crystallization. However, size exclusion chromatograms still showed the presence of both species in solution, with monomeric HepIv1 comprising the major species. Since conditions that facilitate protein crystallization cannot be predicted, an automated system for setting crystallization drops in combination with a variety of commercially available crystallization screens was used to obtain initial hints that could lead to the crystal structure of heparinase I. Although multiple parameters were tested, no protein crystals were observed. A substitution mutant where the non-catalytic active cysteine, C297, was replaced by serine (C297S) was also produced for protein crystallization experiments. However, this mutation did not prevent formation of oligomeric species and the protein showed a very low stability.

To verify the structural integrity of HepIv1 and possibly understand the absence of crystal formation, different biophysical techniques were used. CD studies demonstrated that HepIv1 was folded with the expected antiparallel β-sheet (43.7%) as its major secondary structural element. Based on these results, a β-sandwich fold was assumed for HepIv1, in line with the β-jellyroll fold observed in the homologue heparinase I from *B. thetaiotaomicron*. CD studies of heparinase I previously performed in the presence and absence of Ca^{2+} ions revealed no conformational change, which suggested that this ion is essential for catalytic activity rather than structural stability [168]. Besides, analysis of the oligomerization state of HepIv1 by analytical SEC demonstrated that separation of the monomeric and dimeric species could not be easily achieved since a dynamic equilibrium appeared to exist between both species. The most probable reason for the lack of protein crystallization was found when HepIv1 was exposed to a wide range of temperatures over a defined storage period. Severe protein degradation was observed and the effect was stronger for samples incubated at higher temperatures (20°C and 30°C), indicating that stability of HepIv1 was strongly temperature-dependent. In the past, Lohse and Linhardt have also described the low thermal stability of heparinase I from *P. heparinus*. Upon a cycle of freeze-thawing and storage at 4°C, they observed that the initial catalytic activity of heparinase I was reduced to about 50% [145]. This reduction of activity was most likely due to a loss of structural integrity of heparinase I, which resulted in its inactivation. Similar to previous observations, a mixture of different HepIv1 species was also detected by native MS measurements, this time including denatured protein. The fact that a higher protein concentration was required to obtain a good signal-to-noise ratio and that an increase of dimeric and denatures species occurred after storage of HepIv1 at -20°C for two weeks, demonstrated once more the low thermal stability of produced HepIv1. In agreement with previous observations, dimerization of HepIv1 was most likely unspecific and a result of concentration-dependent molecular interactions. Native MS measurements of heparinase I from Iduron containing the predicted mature N-terminus

suggested that this enzyme had a higher stability than HepIv1 since, besides BSA, which is commonly used to enhance storage stability of enzymes, only monomeric species were present in solution [145].

Thermal unfolding curves determined by DSF also showed that heparinase I from Iduron was more thermally stable than purified Hepv1. Recombinant heparinase from R&D systems behaved similarly to purified HepIv1, showing a very low thermal stability with a T_m of about 30°C. This difference of about 10°C observed between mature heparinase I isolated from *P. heparinus* (Iduron) and enzymes recombinantly produced in *E. coli* lacking the first 21 amino acids (R&D systems and HepIv1) suggested that the N-terminus sequence might influence heparinase stability. Investigation of HepIv1 samples diluted in different buffer conditions from a commercial screen by DSF demonstrated that low pH values and higher salt concentrations resulted in increased thermal stability, with the best buffer showing a T_m increase greater than 10°C. This stability increase was probably derived from an electrostatic effect created by the counter ions from the salt since heparinase I has a theoretical pI greater than 9 and it is therefore strongly positively charged at pH 4. Based on these findings, a modified SEC buffer containing a higher salt concentration of 250 mM and a lower pH value of 4 was used for further experiments. However, protein degradation was still observed as well as the presence of monomers and dimers of HepIv1 in solution. The absence of protein crystals was thus most likely due to the low thermal and conformational stability of HepIv1 in combination with the presence of different oligomerization states in solution, including denatured protein. During crystallization, the long time span necessary for nucleation to occur increases the risk of protein degradation and the presence of denatured protein can induce excessive nucleation and lead to precipitation, lowering the soluble protein concentration below the critical range required for crystallization [187].

Although the first 21 amino acids were predicted to comprise the signal peptide sequence of heparinase I from *P. heparinus*, the actual start site of mature heparinase I after cleavage of the signal peptide has not been experimentally determined. Absence of these amino acids in recombinantly produced heparinase I was shown to not influence its catalytic activity, but it could still have an impact on protein stability and folding [164]. Top-down MS measurements of commercially available heparinase I from *P. heparinus* comprising the mature protein sequence indicated a difference of almost 1 kDa between the MW determined for heparinase I from *P. heparinus* (about 42505 Da) and the theoretical MW of recombinant heparinase I produced in *E. coli* (41750 Da, including four non-native amino acids at the N-terminus, which result from the cleavage site sequence of thrombin). Based on these results, construct HepIv2 was designed containing 10 additional amino acids at the N-terminus (QQLFLCAYA), which resulted in a protein with an identical MW to mature

heparinase I from P. *heparinus*. In contrast to HepIv1, this protein showed a very low solubility, which significantly reduced the amount of soluble protein available for purification. Besides, most likely due to a higher hydrophobicity of this protein introduced by the extended N-terminus sequence, the purification strategy used for HepIv1 did not result in highly pure HepIv2. To improve levels of soluble HepIv2 and optimize its purification strategy, buffers with varying conditions were tested as well as different chromatographic techniques. Although purification of HepIv2 was successfully achieved, only very low amounts of protein, with low solubility, could be obtained. Folding and oligomerization of HepIv2 as determined by CD spectroscopy and SEC were identical compared to HepIv1. In contrast to previous assumptions regarding the N-terminus sequence of mature heparinase I, HepIv2 also showed a low thermal stability with a T_m of about 30°C and protein degradation was still observed when the protein was incubated at higher temperatures. Attempts to crystallize HepIv2 did also not result in protein crystal formation.

According to results obtained with HepIv2, the difference of about 1 kDa observed between mature heparinase I from *P. heparinus* and recombinant heparinase I lacking amino acids 1 to 21 did not seem to be explained by an extended N-terminus sequence of the mature protein. Another possible reason for the observed mass difference could be post-translational modifications of heparinase I by *P. heparinus.* In fact, one publication reported that upon cleavage of the leader sequence native heparinase I becomes *O*-glycosylated at S39 and that the N-terminal Q22 residues was further converted to pyroglutamate. The MW of the bacterial carbohydrate structure (about 1.1 kDa) and the absence of this as well as the N-terminal PTM in *E. coli* could be a plausible explanation for the observed discrepancy in molecular masses [148], [298]. Although glycosylation is more common for proteins produced in mammalian organisms, a similar glycosylation has also been reported for enzymes from *F. meningosepticum* [300]. Considering the previous observations, the higher thermal stability determined for heparinase I from Iduron could also be due to the addition of BSA in the Iduron preparation rather than a result of its possibly differing N-terminal sequence. Addition of BSA has been suggested to reduce molecular collision and formation of intermolecular associations, thus generally preventing dimerization and denaturation [301]. Addition of a protein such as BSA is a common strategy to preserve catalytic activity of enzymes for commercial use, however such a stabilization measure cannot be applied in protein crystallization trials.

Investigation of a third expression construct, HepIv3, comprising the sequence of predicted mature heparinase I followed by a non-cleavable His$_6$-tag at the C-terminus was terminated at a very early stage since N-terminal truncations were observed already during protein expression. A similar construct design was previously reported for recombinantly produced

heparinase III from *P. heparinus* where such an effect was not observed and its expression resulted in the determination of a three-dimensional crystal structure [160]. Although it is unknown why N-terminal truncation occurred it is possible that the His$_6$-tag followed by a cleavage site at the N-terminus protected the N-terminus from proteolytic degradation at the N-terminus in HepIv1 and HepIv2.

A preliminary activity assay performed with HepIv1 showed that the enzymatic depolymerization of heparin is dependent on the concentration of calcium ions. Ca^{2+} can enhance the activity of this enzyme and the highest activity, determined by monitoring the absorption at 232 nm, was indeed observed with a concentration of 10 mM CaCl$_2$. Compared to commercial heparinase I from R&D systems, HepIv1 showed a lower activity resulting in a lower level of heparin-derived oligosaccharides, which was probably due to its thermal denaturation and subsequent inactivation.

Up to date, research done on heparinase I from *P. heparinus* has been based on its predicted secondary structure and models developed using site-directed mutagenesis. Based on the obtained results, it is possible that the lack of structural studies on this enzyme is due to its very low thermal and conformational stability. This seems to be an intrinsic property of this enzyme and it could be that heparinase I produced by *P. heparinus* has a short half-life thus avoiding complete depolymerization of heparin to disaccharides. To select a stabilization strategy for heparinase I, understanding its thermal inactivation mechanism would be essential to identify possible enzyme-stabilizing agents that are compatible with crystallization, in order to enhance rigidity and prevent inactivation of the enzyme [301]. Protein engineering by random mutagenesis (error-prone PCR) is an alternative approach to obtain enzymes with higher stability. However, large libraries and a few mutagenesis cycles are required to find improved variants [302]. Structure-based engineering of heparinase I could not be accomplished since no protein crystals of recombinantly produced heparinase I were obtained. However, considering that crystal structures of heparinase I from *B. thetaiotaomicron* are available, this approach could be used in the future to modify this enzyme instead of heparinase I from *P. heparinus*. Modification of amino acids located at the heparin binding site could potentially switch the mechanism of action of heparinase I, resulting in a non-processive enzyme that would release medium-sized fragments of heparin instead of subsequently cleaving them until mostly disaccharides are obtained.

6   Project III

Structural-based optimization of selective type I c-Jun N-terminal kinase 3 inhibitors

Contributions to this work

Dr. Francesco Ansideri and Prof. Dr. Pierre Koch from the Institute of Pharmaceutical Sciences, University of Tübingen, designed and synthesized inhibitors used for structural analysis in this work and determined their biological activity. The work presented in this chapter has been published in collaboration.

Publication

6.1 Results

6.1.1 Development of selective JNK3 inhibitors 38 and 44

To obtain selective inhibitors for JNK3, a compound with a pyridinylimidazole scaffold with dual activity for both JNK3 and p38α MAPK was selected as the starting point for subsequent modifications (compound **1a**, Table 6.1). This compound was previously reported as a precursor for the synthesis of a fluorescent probe used in fluorescence polarization-based binding assay for JNKs [303], [304]. After several attempts to optimize this ATP-competitive lead compound, two type I inhibitors displaying high selectivity towards JNK3 were obtained and used for crystallization (compounds **38** and **44**, Table 6.1). In detail, the pyridinylimidazole scaffold of compound **1a** was first modified by substitution of the pyridine C2-position with a 4-morpholinoaniline group. To shift the preference of the inhibitor towards JNK3, the *p*-fluorophenyl group at the imidazole-C4 position of compound **1a** was replaced by a methyl substituent at the imidazole-C4 position resulting in compound **44**. This 4,5-disubstituted imidazole was further modified by removal of the *S*-methyl group at the imidazole-C2 position yielding compound **38**. The inhibitors were evaluated by ELISA to determine their ability to inhibit JNK3 and p38α MAPK [305]. Both inhibitors displayed no significant inhibition of p38α MAPK at tested concentrations (concentration producing 50% of inhibition (IC_{50}) > 10 μM), while still being able to inhibit JNK3 with IC_{50} values in the sub-micromolar range. Additionally, presence of the 2-methylsulfanyl moiety in compound **44** resulted in a two-fold increase in the inhibitory potency on JNK3 (IC_{50} = 363 nM).

Table 6.1 Structure and biological activity of type I inhibitors **38** and **44**. IC$_{50}$ values represent the mean value of three independent experiments. Values between parentheses correspond to inhibition at 10 µM concentration in percent. Results were obtained by Dr. Francesco Ansideri and Prof. Dr. Pierre Koch, Institute of Pharmaceutical Sciences, University of Tübingen [8].

Compound	Structure	IC$_{50}$ ± SD [nM] JNK3	IC$_{50}$ ± SD [nM] p38α MAPK
1a		24	17
38		833±139	>10 000 (41%)
44		363±34	>10 000 (48%)
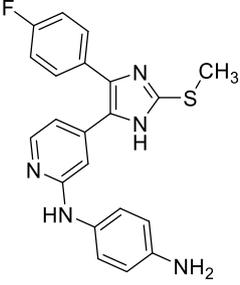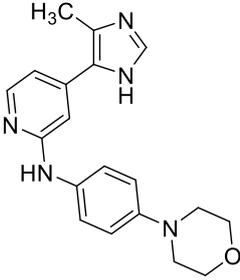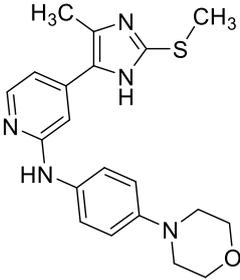
To elucidate the binding mode of 4-methyl-5-(pyridine-4-yl) imidazole derivatives and gain insight into the role of the *S*-methyl group introduced in compound **44**, structures of JNK3 in complex with compounds **38** and **44** were determined by X-ray crystallography. The human JNK3 protein used for crystallization was truncated (the amino acid sequence is shown in section 9.1) and recombinantly produced in *E. coli*. A JNK3 construct lacking 39 residues at the N-terminus and 62 residues at the C-terminus was previously used to grow large, well-ordered crystals that diffracted to 2.3 Å resolution [239]. Contrarily to the phosphorylated protein used for inhibition studies, recombinant JNK3 was not phosphorylated and was, therefore, inactive. To date no structure was published of phosphorylated JNK3 and all available structures have several regions missing [244].

6.1.2 Crystallization of JNK3-inhibitor complexes

Crystal growth is highly dependent on protein quality used for crystallization, and the protein should ideally be pure and homogenous. To verify the quality of the JNK3 sample provided before proceeding with crystallization experiments, analytical size exclusion chromatography and SDS-PAGE were used and sample purity and homogeneity was confirmed (Figure 6.1). On a Superdex 75 PC 3.2/300 column, JNK3 eluted at 1.18 mL, which corresponded to a MW of 41.9 kDa, indicating the presence of a monomeric protein. One major band at approximately 42 kDa was also visible in SDS-PAGE analysis, corresponding to the MW of the truncated protein as well (Table 9.1).
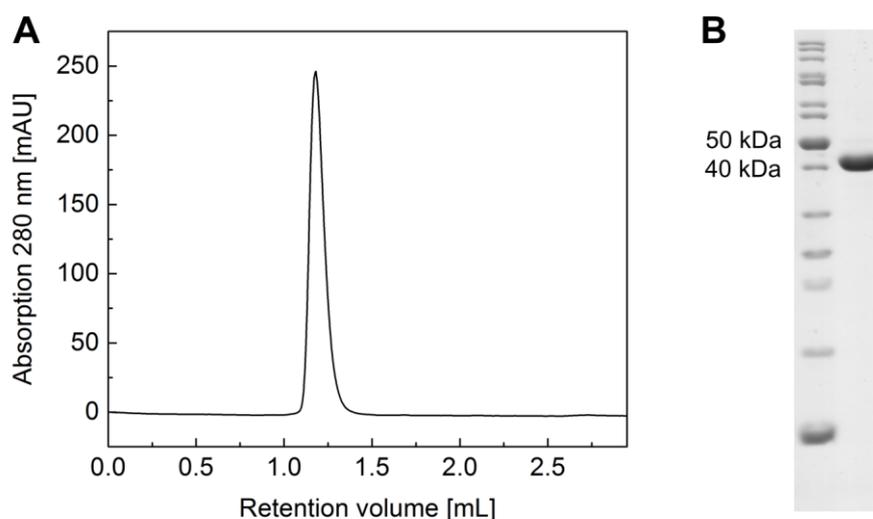
Figure 6.1 Analysis of JNK3 protein sample using analytical SEC and SDS-PAGE. A) Size exclusion chromatogram obtained using a Superdex 75 PC 3.2/300 column. A single peak was observed at 1.18 mL corresponding to monomeric JNK3. B) SDS-PAGE analysis of purified JNK3. The band at approximately 42 kDa corresponds to the molecular weight of monomeric truncated protein.

Initial JNK3 crystals were obtained using a similar experimental procedure as described by Lange *et al.* [272]. Before setting up crystallization drops, JNK3 protein was incubated on ice for 30 min with the non-hydrolyzable ATP analogue AMP-PCP. The sitting drop vapor diffusion method was used for crystallization, and the amount of PEG 3350 was varied between 28-31% (v/v). Similar crystals were observed for all conditions tested (Figure 6.2). Crystal growth was observed over time and multiple lattices were present in the X-ray diffraction patterns. The obtained crystals only showed X-ray diffraction up to 3 Å resolution.
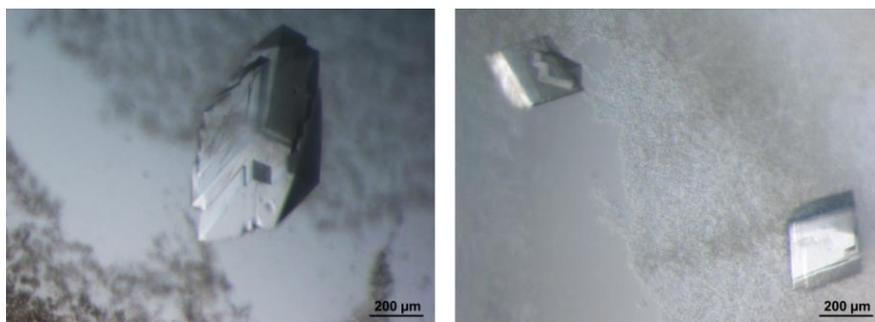
Figure 6.2 Initial JNK3 crystals containing AMP-PCP. Large crystals were obtained containing multiple diffraction lattices and showing limited diffraction up to 3 Å.

To improve crystal quality and achieve higher resolution, the protein concentration was reduced from 10 mg/mL to 5 and 2.5 mg/mL, and microseeding was used. Smaller crystals containing a single lattice and diffracting up to 1.76 Å (Figure 6.3) were obtained. Subsequently, crystals were incubated with 10 mM of inhibitors **38** and **44** for 36 h to obtain the complex structures.
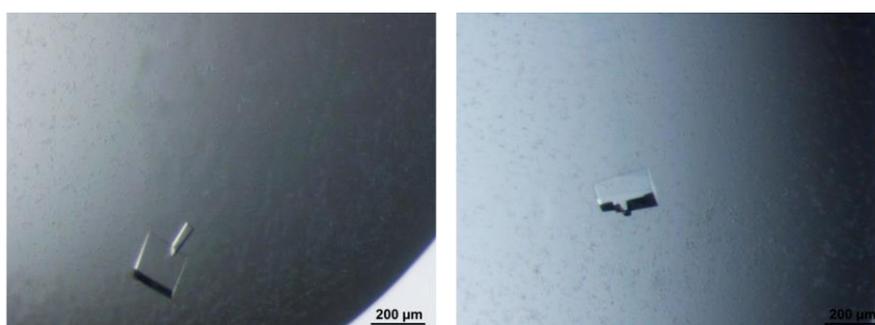
Figure 6.3 Optimized JNK3 crystals containing AMP-PCP used for data collection after incubation with inhibitors. Crystals were improved by reducing the protein concentration from 10 mg/mL to 5 or 2.5 mg/mL and microseeding.

Data collection of crystals containing inhibitors **38** and **44** and crystals containing AMP-PCP was performed at the X06DA beamline (Swiss Light Source, Villigen, Switzerland). Three data sets with diffraction limits ranging from 2.10 Å to 1.76 Å were obtained, and crystal structures were determined (Table 6.2). Although several JNK3-ATP analogue structures are available, JNK3-AMP-PCP was also determined for structural comparison purposes. Three-dimensional structures were solved in different orthorhombic space groups using molecular replacement. R-factors of the JNK3-**38** structure (R_{work}/R_{free}: 21.81/26.85%) are slightly higher than those for the **44** and AMP-PCP complexes, most likely as a consequence of pseudo-translational symmetry as estimated using Xtriage [277].

6.1.3 Crystal structures of JNK3 complexes

Table 6.2 Data collection and refinement statistics of JNK3 complex structures. Values between parentheses correspond to the respective highest resolution shell.

	JNK3+AMP-PCP	JNK3+38	JNK3+44
PDB ID	6EQ9	6EMH	6EKD
Data collection			
Space group	$P2_12_12$	$P2_12_12_1$	$C222_1$
Cell dimensions			
a, b, c (Å)	156.73,110.49, 43.95	88.56,114.26, 157.80	81.51,124.82, 68.89
α, β, γ (°)	90, 90, 90	90, 90, 90	90, 90, 90
JNK3 monomer/ASU	2	4	1
Resolution [Å]	47.23-1.83 (1.94-1.83)	48.01-1.76 (1.81-1.76)	48.49-2.10 (2.15-2.10)
Measured reflections	784898 (108776)	4164005 (296660)	540357 (39002)
Unique reflections	68590 (10906)	157963 (11494)	20913 (1538)
Completeness [%]	98.8 (95.9)	99.9 (98.9)	100 (99.9)
Redundancy	11.4 (10.0)	26.4 (25.8)	25.8 (25.4)
CC1/2 [%]	99.9 (66.2)	100 (63.2)	100 (55.6)
I/σ(I)	18.1 (1.6)	23.0 (1.4)	26.4 (1.6)
Wilson *B*-factor [Å^2]	29.75	30.84	43.8
Refinement			
Resolution [Å]	47.23-1.83	48.01-1.76	48.49-2.10
R_{work}/R_{free}	20.62/25.18	21.81/26.85*	20.56/25.90
Number of atoms			
Protein chain a/b/c/d	2812/2606	2826/2826/2700/2665	2558
Water	391	861	104
Ligand**	62	100	27
B-factors [Å^2]			
Protein chain a/b/c/d	34.7/39.6	38.5/40.7/44.2/46.4	52.6
Water	41.3	43.6	48.5
Ligand**	47.9	37.7	46.9
R.m.s. deviations			
Bond lengths [Å]	0.015	0.014	0.015
Bond angles [°]	1.44	1.52	1.53

* Pseudo translational symmetry is present.

** AMP-PCP and compounds **38** and **44**.

During model building, all ligands could be placed in the difference electron density map of JNK3 chains present in the asymmetric units of different crystals. Simulated annealing omit difference electron density maps were calculated with a radius of 5 Å around the ligand for all data sets (Figure 6.4). These maps confirmed the presence and highlighted the features of all ligands bound at the ATP binding pocket of JNK3.
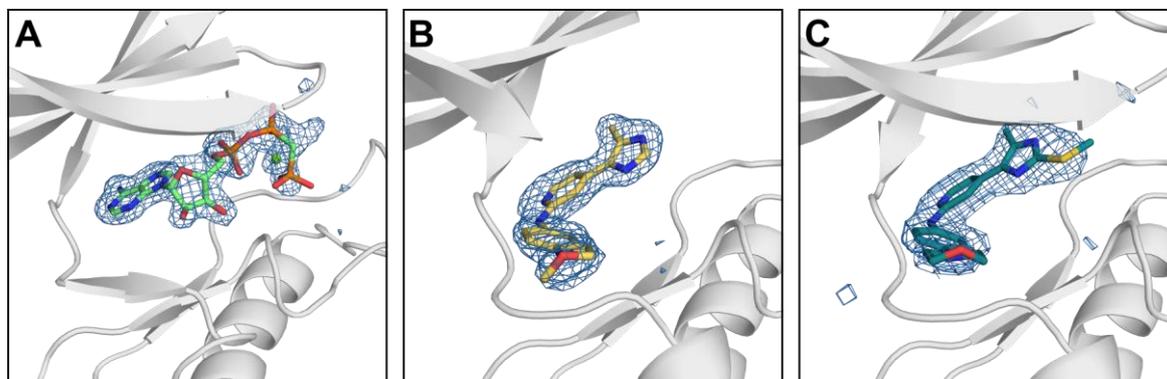
Figure 6.4 Binding of non-hydrolyzable ATP analogue AMP-PCP and inhibitors **38** and **44** to JNK3. Simulated annealing omit F_{obs}-F_{calc} electron density maps for the ligands bound to JNK3 were contoured at 3.0 σ and are displayed with a radius of 5 Å around the ligand. A) AMP-PCP bound to chain A of the crystal asymmetric unit. B) Compound 38 bound to chain A of the crystal asymmetric unit. C) Compound 44 bound to the only chain present in the asymmetric unit.

For JNK3 bound to compound **38** and AMP-PCP, more than one protein-ligand chain was present in the crystal asymmetric unit. Due to the lower B-factors and the high structural similarity to the other copies, only chain A from the symmetric unit of the crystal was used for further structural analysis. The four copies of JNK3-**38** were compared by superposition using the align function of PyMOL (Figure 6.5). A root mean square deviation (RMSD) of 0.2 Å was determined when chains B, C and D where aligned to chain A, indicating an overall very good alignment. Some displacement was caused by crystal contacts, however the orientation and placement of all compounds was almost identical. The same was observed for the JNK3 structure containing AMP-PCP, which comprised two copies in the asymmetric unit.
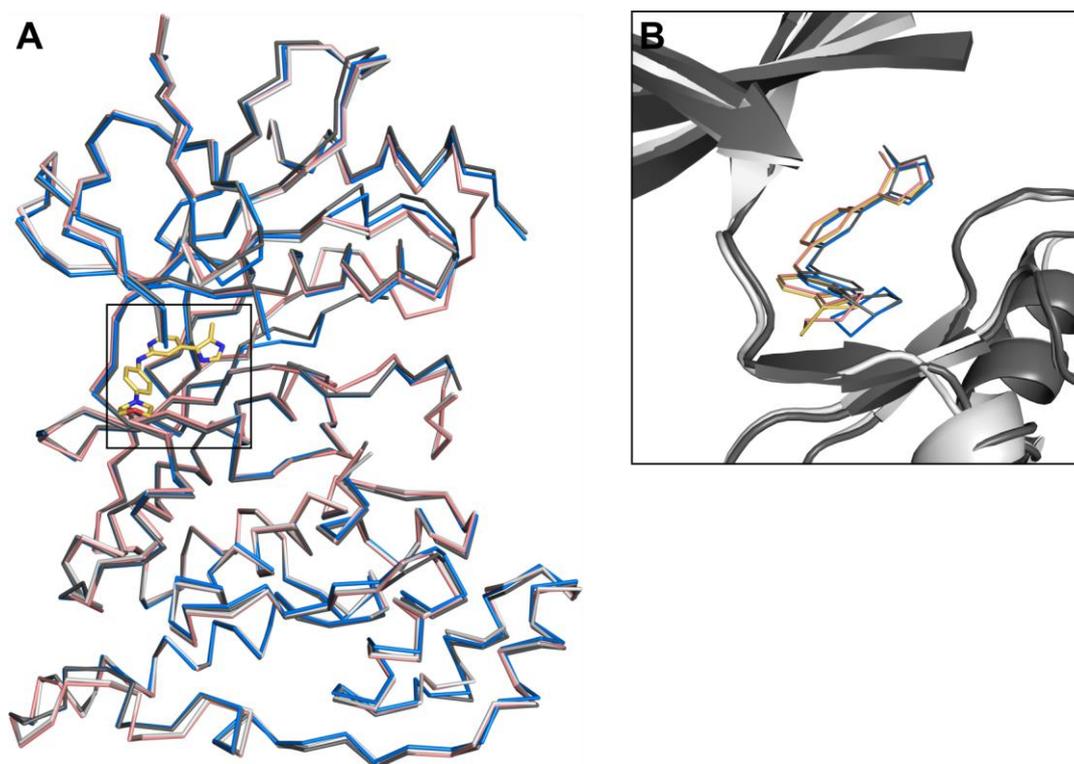
Figure 6.5 Overlay of JNK3-**38** copies present in the asymmetric unit. Superposition of the four copies was performed using the align function in PyMOL. A) Overview of the four protein chains and the ligand binding. The protein backbone is displayed in ribbon (chains A, B, C and D are highlighted in light gray, pink, dark gray and blue, respectively) and compound **38** bound to chain A is depicted as sticks. B) Close-up view of the binding site of chains A (light gray) and C (dark gray) and superposition of compounds bound to each copy of the asymmetric unit. Compounds are shown in in light gray, pink, dark gray and blue, for chains A, B, C and D, respectively.

The JNK3 protein comprises two domains, the N-terminal domain mostly formed by β-strands and the C-terminal domain that is predominantly α-helical. A cleft formed between the two domains comprises the ATP-binding site where all ligands, inhibitors **38** and **44** and AMP-PCP had bound (Figure 6.6). During crystal incubation, these ATP-competitive inhibitors replaced the already bound AMP-PCP molecule in the binding pocket. Binding of compounds **38** and **44** occurred without causing major conformational changes in the protein backbone (only shown for the JNK3-**44** structure). Both inhibitors interacted with the hinge region of the kinase via two hydrogen bonds involving the main chain carbonyl and the backbone amide of M149. In addition, the imidazole-N atom distal from the pyridine ring was part of a network of water-mediated hydrogen bonds, involving the side chain of K193 and the main chain of L206. The imidazole-N atom proximal to the pyridine ring participated in a water-mediated hydrogen bond with the side chain of N152. Further water-mediated hydrogen bonds included the backbone of G76 and the side chain of D207 in the JNK3-**44** structure and the side chain of N194 in the JNK3-**38** structure. Multiple hydrophobic interactions involving the side chains of I70, V78, M146, V196 and L206 also contributed to inhibitor binding. The methyl group present in both inhibitors was oriented towards the HR1,

and the orientation of the side of the gatekeeper residue M146 was identical for both complex structures. The 4-morpholinoaniline moiety occupied the solvent-exposed HR2 and no directed interactions with JNK3 were observed. A major structural difference between the two complex-inhibitor structures was observed for the glycine-rich loop. In the JNK3-**44** structure, the electron density in this region was well defined and all residues could be built in the model, while in the JNK3-**38** crystal structure no electron density was visible for residues G71-G76. This region was clearly defined in the JNK3-AMP-PCP crystal structure and residues I70-V78 formed nonpolar interactions with the purine ring of AMP-PCP. Additionally, the adenine group of AMP-PCP interacted with JNK3 via two hydrogen bonds formed with the carbonyl group of E147 and the backbone amide of M149. The ribose hydroxyls formed a hydrogen-bonding network to the side chain of N152 and the carbonyl group of S193. The triphosphate group formed direct hydrogen bonds involving the side chains of residues G76, Q75, K93, K191. Other indirect hydrogen bonds were formed between the protein chain and the phosphate groups via water molecules and one Mg^{2+} ion. The interfaces formed between the ligands and the JNK3 protein were analyzed using the PISA server by measuring surface area that becomes inaccessible to solvent upon complex formation [306]. Buried surface areas of 455.4 Å, 385.1 Å and 538.8 Å were determined for complex structures containing compounds **44**, **38** and AMP-PCP, respectively.
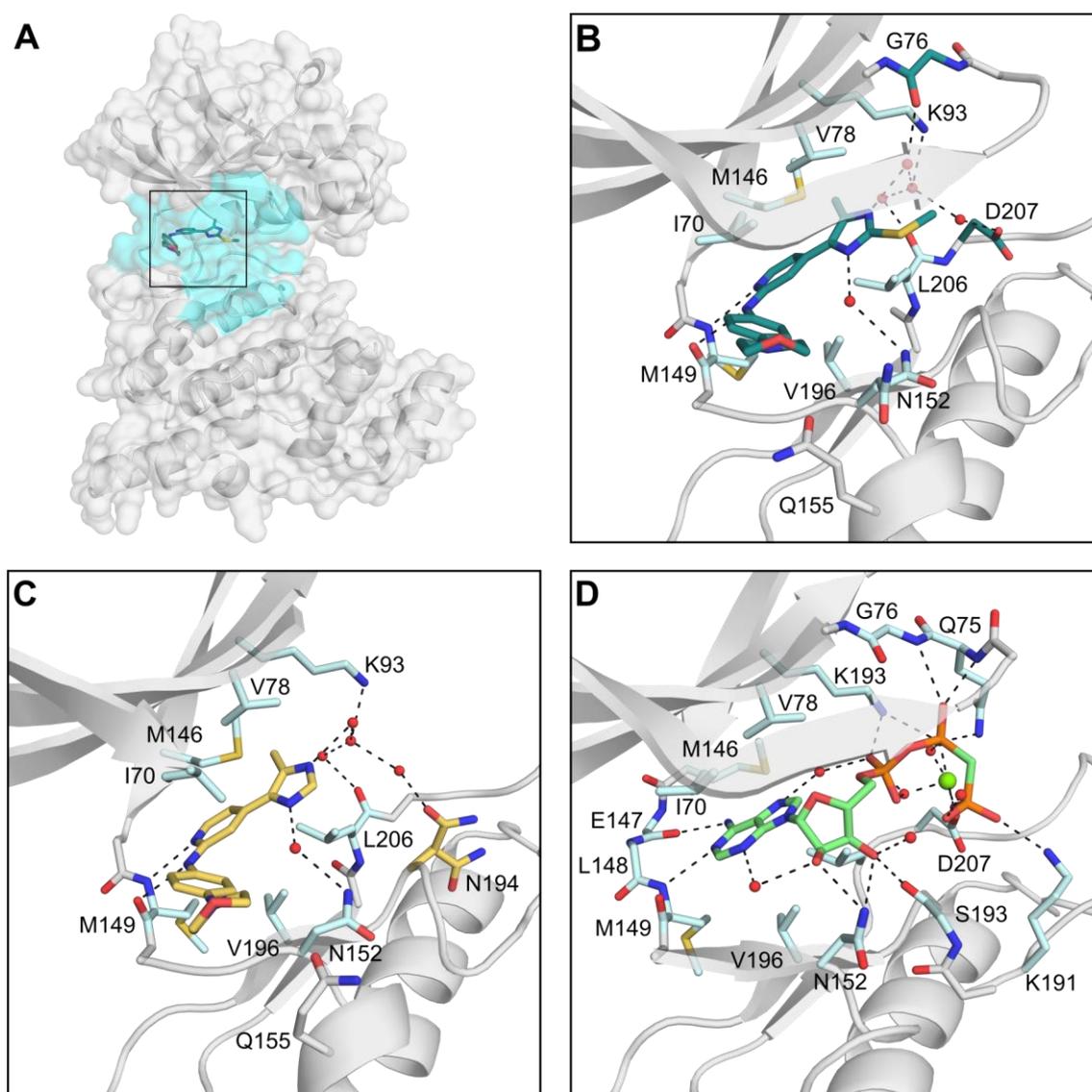
Figure 6.6 Crystal structure of JNK3 in complex with inhibitors featuring a pyridinylimidazole scaffold and AMP-PCP. A) Surface representation of JNK3 structure in complex with compound **44**. Protein backbone is displayed as cartoon and compound **44** is shown as sticks. The ATP binding pocket is highlighted in cyan. JNK3 active site from JNK3-**44** (B), JNK3-38 (C) and JNK3-AMP-PCP (D) complex structures are shown in the same orientation. Ligands and key residues for the interaction are highlighted in stick display. Water molecules are represented as red spheres and hydrogen bonds are depicted as black dashed line. Residues with common orientations and interactions are shown in light blue for the complex-inhibitor structures, whereas residues differing in their orientations are shown in the same color as the respective compound. Side chains of N152 (B) and N194 (C) displayed multiple orientations. In (D), the adenine and ribose groups of AMP-PCP are highlighted in green and the three phosphate groups in orange. A Mg^{2+} ion interacting with the phosphate groups of AMP-PCP is depicted as a green sphere.

6.1.4 Determination of melting temperatures for unliganded and liganded JNK3

An additional characterization of the ATP analogue and the two compounds included the determination of JNK3 protein melting temperature (T_m) in presence and absence of these ligands by nanoDSF. This methodology assesses the effect of a binding event on the target

protein's thermal stabilization and can be employed as a screening tool for potential binders. Changes of the intrinsic fluorescence of tryptophan and tyrosine residues, as a result of temperature induced protein unfolding, were monitored for T_m determination. Measured melting curves and their correspondent first derivative analysis exhibited an increase in the thermal stability of JNK3 upon ligand binding, which can be seen by a shift in the inflection point of the melting curves towards higher temperatures (Figure 6.7A). First derivative analysis was used to determine the melting temperatures (Figure 6.7B). T_m of free JNK3 was determined to be 46.28°C. T_m increased to 48.12°C upon AMP-PCP binding and to 53.87°C and 54.83°C upon on binding of compounds **38** and **44**, respectively.

Samples	T_m [°C]
JNK3	46.28 ± 0.58
JNK3+AMP-PCP	48.12 ± 0.15
JNK3+38	53.87 ± 0.04
JNK3+44	54.83 ± 0.04
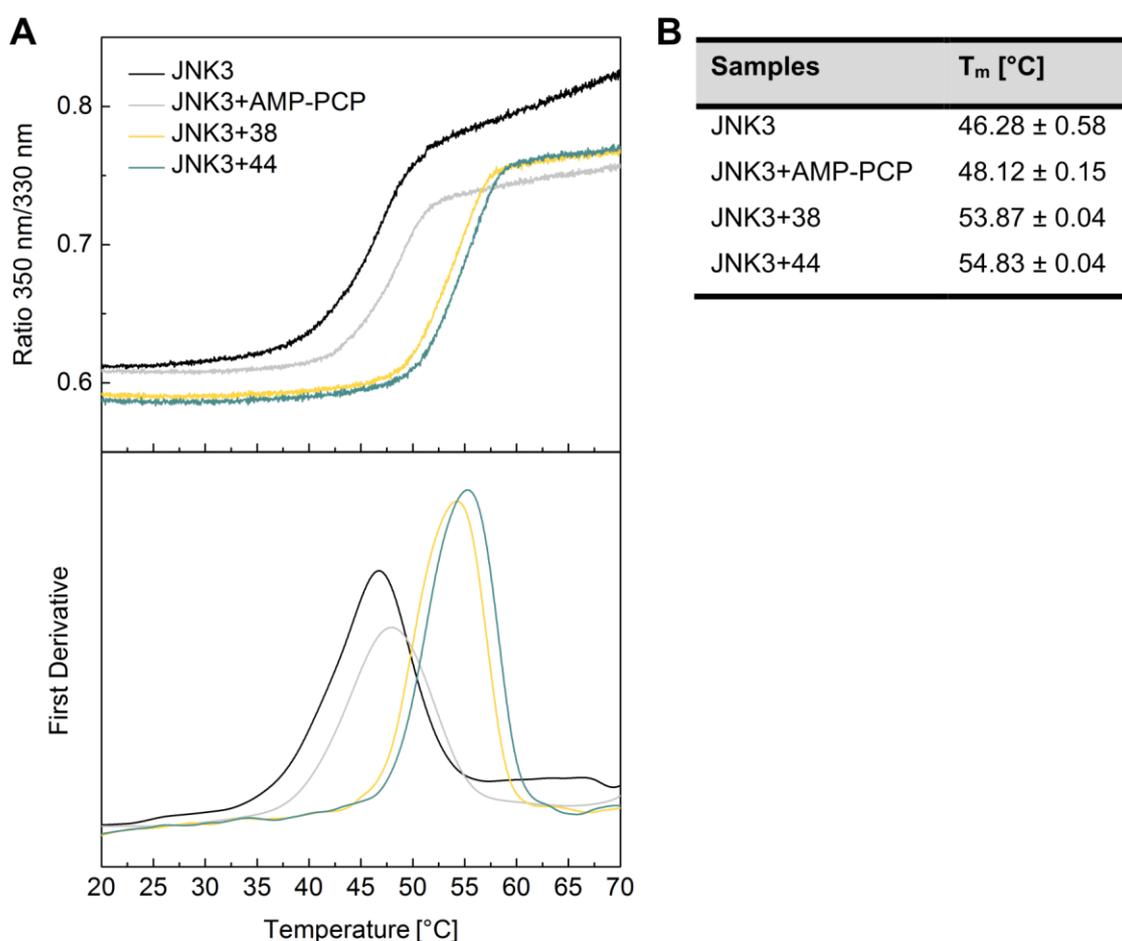
Figure 6.7 Effect of ligand binding on JNK3 thermal stability determined by nanoDSF. Melting curve corresponding to free JNK3 is shown in black. Melting curves of JNK3 complexes containing AMP-PCP and compounds **38** and **44**, curves are displayed in light gray, yellow and teal, respectively. A) Intrinsic fluorescence intensity ratio of tryptophans and tyrosines (350/330 nm) as well as the first derivative are plotted as a function of temperature. B) Melting temperatures determined by nanoDSF. The data shown represents mean values ± standard deviation of three experiments.

6.2 Discussion

Despite the efforts and the high number of reports concerning JNK3 inhibitors, there are still no approved drugs targeting this enzyme. Due to a high conservation of the ATP binding site within the kinome and the JNK family, achieving selectivity when investigating ATP-competitive inhibitors remains a significant challenge. Modification of the substitution pattern of a dual JNK3/p38α MAPK compound with a pyridinylimidazole scaffold (**1a**) proved to be effective in shifting the selectivity towards JNK3. With this approach, two new JNK3 selective inhibitors (**38** and **44**) that exhibited a 2-fold difference in inhibitory potency were obtained. Both inhibitors share a high similarity, with the only difference being an additional *S*-methyl group at the imidazole-C2 position in compound **44**. This substituent seems to be effective in enhancing the inhibitory potency towards JNK3 without losing selectivity. The crystal structures of JNK3 in complex with inhibitors **38** and **44** revealed the binding mode of these ATP-competitive inhibitors and shed light on the structural influence of *S*-methyl group in binding affinity. JNK3 crystals were obtained using a truncated form of JNK3 since some regions of the native protein are highly flexible and are known to interfere with lattice formation during crystallization [239].

A structure of non-phosphorylated JNK3 containing the non-hydrolyzable ATP analogue AMP-PCP was initially solved to compare the structural differences upon inhibitor binding. This structure is highly similar to the first non-phosphorylated JNK3 structure in complex with an ATP analogue (AMP-PNP) reported in 1998 [239]. Superposition of the N-terminal (residues 56-149 and 382-400) and C-terminal (residues 150-208, 226-315 and 329-369) yielded a protein backbone RMSD of 0.853 Å. The main difference observed between the two structures was related to Mg^{2+} binding. In contrast to the first JNK3 structure, where two Mg^{2+} ions were mediating ATP binding, only one Mg^{2+} ion was defined in the electron density of the JNK3-AMP-PCP complex structure. This ion is essential for catalysis of protein kinases and is considered the primary metal since it is visible in the crystal structure at low concentrations of Mg^{2+} [243], [245]. In the JNK3-AMP-PCP structure, the β- and γ- phosphoryl groups of ATP and residue D207 chelated the Mg^{2+} ion.

Incubation of JNK3 crystals containing AMP-PCP with compounds **38** and **44** proved to be an effective method to obtain the complex JNK3-inhibitors-crystal structures. AMP-PCP bound at the ATP binding pocket was replaced and a similar binding mode was observed for both inhibitors. As previously seen for this class of inhibitors, the pyridinylimdazole scaffold interacted with the hinge region of JNK3 via a bidentate hydrogen bond. These interactions are known to mimic the ATP binding to JNK3 and to highly contribute to an increase of the binding affinity of inhibitors [307]. Additionally, the methyl substituent at the imidazole-C4 position was accommodated in the HR1 region, and the 4-morpholinoaniline group occupied

the solvent-exposed HR2 region. In the case of compound **44**, the 2-methylsulfanyl moiety was positioned in the phosphate region. Docking studies reported by Dr. Francesco Ansideri, supported by a publication from Wityak *et al.*, predicted an alternative binding mode for compound **38** that could explain the 2-fold difference in binding affinity [308], [309]. According to this hypothesis, a 180° flip of the imidazole ring with respect to the pyridine ring of compound **38** could occur in the absence of the *S*-methyl substituent. Due to this rotation, the methyl group would not occupy the hydrophobic selectivity pocket thus reducing its inhibitory potency on JNK3. This hypothesis could not be confirmed with the solved crystal structures since both inhibitors assumed the classical conformation with the 4-methyl-substituent positioned towards the HR1 of JNK3, and the imidazole-N atom interacting with the conserved K93 via a water-mediated network of hydrogen bonds. The 4-morpholinoaniline displayed high flexibility and due to the lack of direct interactions with JNK3, this group barely contributes to the inhibitor binding affinity. This substituent was solely introduced to replace the free terminal aniline group present in compound **1a**, which was recognized by the ZINC 15 pattern tool to be potentially responsible for compound aggregation and might therefore cause assay interference that could result in misleading biological activities [310].

Complex structures of JNK3-containing inhibitors **38** and **44**, AMP-PCP and a dual JNK3/ p38α inhibitor from Scapin *et al.* (PDB code 1PMN) were compared to gain further insights into the structural basis for the observed selectivity of compounds **38** and **44** [311]. Superposition of these structures highlights the position of the gatekeeper residue M146 and the Gly-rich loop (Figure 6.8). As expected, in both complex structures containing the 4-methyl substituted inhibitors, no movement of the gatekeeper residue occurred, and the orientation of its side chain is similar to the one observed in the presence of AMP-PCP. In contrast, when a bulky substituent such as the dichlorophenyl moiety of the dual JNK3/p38α inhibitor is present in the HR1 region, an induced-fit of movement of the hydrophobic side chain of M146 occurs. The shift of M146 creates a larger selectivity pocket that resembles the cavity of p38α MAPK, where the gatekeeper residue is T106. The small methyl substituent of compounds 38 and 44 was able to fit into the narrow selectivity pocket of JNK3 and interact with the gatekeeper residue without altering its orientation. A variety of substituents, including aromatic and cycloalkyl moieties and branched aliphatic groups were tested to replace the *p*-fluorophenyl group at the imidazole-C4 position of the dual inhibitor **1a**. The effect of small alkyl groups was also evaluated with 4-unsubstituted and 4-ethyl derivatives. As confirmed by the high-resolution crystal structures, the methyl substituent possessed the optimal length to target the selectivity pocket of JNK3. This moiety was therefore decisive in shifting the selectivity towards JNK3. Furthermore, the multiple hydrophobic interactions involving residues I70, V78, M146, V196 and L206 also seemed to

promote JNK3 selectivity since small and flat inhibitors that cannot reach into binding pockets are unable to form these interactions in the wider ATP binding site of p38α MAPK. These interactions have also been previously described by Scapin *et al.* [311].

The most significant structural difference between the complex structures containing the 4-methyl-5-(pyridine-4-yl) imidazole derivatives was observed for the G-rich loop. In the JNK3-**38** complex structure, the lack of electron density in this area indicates high local flexibility, which has also been reported for other JNK3-inhibitor crystal structures [312]-[314]. However, in the JNK3-**44** complex structure, the electron density for this loop was well defined, indicating a structural stabilization of this region upon interaction with the 2-methylsulfanyl moiety. To confirm that the stabilization of the G-rich loop was driven by the presence of inhibitor **44**, the crystal packing was analyzed and no strong crystal contacts were observed within this region. The hydrophobic and polar interactions formed between the *S*-methyl group and the G-rich loop appear therefore to be responsible for the downward positioning of this flexible loop, which adopted a similar position to the G-rich loop of JNK3-AMP-PCP (Figure 6.8). The compression of this loop was previously reported for a JNK3 complex crystal structure by Kamenecka *et al.* [315] and compared to the structure of Scapin *et al.* containing the dual JNK3/p38α MAPK inhibitor. The ability of the G-rich loop of kinases to "collapse" onto the ligand, thus creating a smaller, less solvent-exposed cavity with higher binding affinity for the inhibitor, has been also considered one element for selectivity [241]. The effect of the *S*-methyl group on the positioning of the G-rich loop is most likely the structural reason for the 2-fold gain in affinity of inhibitor **44** in comparison to inhibitor **38**.
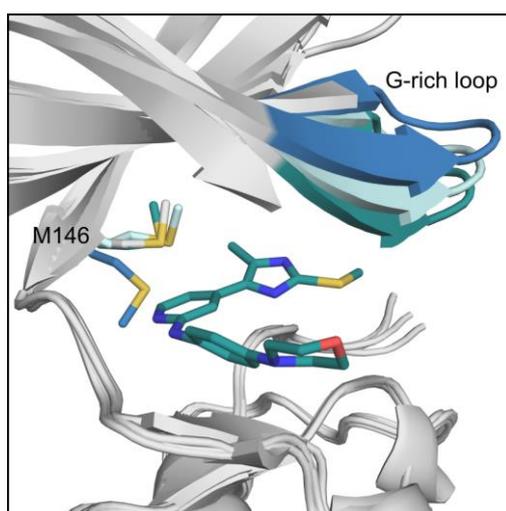
Figure 6.8 Comparison of ligand-bound JNK3 structures with focus on the gatekeeper residue M146 and the G-rich loop positioning upon inhibitor binding. Superposition of JNK3-**44** (teal), JNK3-38 (gray), JNK3-AMP-PCP (light blue) and dual JNK3/p38α MAPK inhibitor (dark blue) complex structures. Structures were superposed using the align function in PyMOL. The dual inhibitor structure was reported by Scapin *et al.* [311] (PDB code 1PMN). Side chains of gatekeeper residue and G-rich loop are depicted using the color code mentioned above. Only inhibitor **44** is shown for sake of clarity.

Binding of pyridinylimidazole-based inhibitors **38** and **44** to JNK3 was also assessed using a thermal shift assay. Tight binding of inhibitors stabilize the protein in its folded conformation, resulting in significantly higher melting temperatures [256]. Comparison of unfolding profiles and the determined T_m values of JNK3 upon ligand binding clearly showed that inhibitors **38** and **44** have a significant influence on the thermal stability of the protein. The difference in observed T_m for both inhibitors correlates well with the results concerning inhibitory potency and stability of the G-rich loop. The determined T_m for inhibitor **44** supports the stabilization of G-rich loop by the 2-methylsulfanyl substituent observed in the crystal structure, contributing to its higher inhibitory potency on JNK3.

The determined crystal structures allow for further rational design of analogues of inhibitor **44** in order to increase its inhibitory potency on JNK3 without losing selectivity. One potential modification is the replacement of the 4-morpholinoaniline moiety accommodated in the HR2 region, since no direct interactions were observed between this group and JNK3. This solvent-exposed area can be exploited to increase inhibitory potency by introducing substituents that can form interactions with the non-conserved amino acids of JNK3. Targeting residue Q155 located in the HR2 is a potential strategy to gain affinity and simultaneously selectivity on JNK3. This residue can act as both acceptor and donor of hydrogen bonds and is replaced by a shorter asparagine residue in p38α MAPK (Figure 6.9). In the JNK3-**44** crystal structure, residue Q155 is located about 4 Å away from the 4-morpholinoaniline moiety but it cannot be reached due to the rigidity of this substituent. Besides, this moiety is only able to accept a single hydrogen bond. A plausible strategy could consist in the replacement of this group by a *trans*-4-aminocyclohexanol substituent, which has been previously shown to form a hydrogen bond with Q155 in a crystal structure reported by Krenitsky *et al.* [316]. This approach has already been tested, and although the selectivity over p38α MAPK was preserved, no gain in affinity was obtained (see publication indicated at the beginning of this chapter [8]). This outcome suggested that the introduced moiety was not able to form the desired interaction with Q155 side chain.

Gly-rich loop									N-lobe		
I70	G71	S72	G73	A74	Q75	G76	I77	V78	A91	I92	K93
V30	G31	S32	G33	A34	**Y35**	G36	**S37**	V38	A51	**V52**	K53

αC-Helix				N-lobe				Hinge region						
R107	E111	L112	M115	I124	L126	L144	V145	M146	E147	L148	M149	D150	A151	N152
R67	E71	L72	**L75**	I84	L86	L104	V105	**T106**	**H107**	L108	M109	**G110**	A111	**D112**

C-lobe								DFG-motif			
C154	Q155	D189	K191	P192	S193	N194	I195	V196	L206	D207	F208	G209
N114	N115	D150	K152	P153	S154	N155	**L156**	**A157**	L167	D168	F169	G170

Figure 6.9 Sequence alignment of residues surrounding the ATP binding site of JNK3 and p38α MAPK. Differing amino acids are depicted in bold in the p38α MAPK sequence. Sequence alignment was generated using Clustal Omega from European Molecular Biology Laboratory (EMBL-EBI) [319]. Figure was adapted from [6].

The N152 residue of JNK3 also represents a potential target to increase binding affinity while preserving selectivity, since this residue is not conserved in p38α MAPK (Figure 6.9). In crystal structures containing inhibitors **38** and **44** a water-mediated hydrogen bond was observed between this residue and the imidazole-N atom proximal to the pyridine ring of both inhibitors. Previous attempts to substitute this position with aliphatic groups resulted in a detrimental effect on inhibitory activity, which is mostly like due to the disruption of the water-mediated hydrogen bond from with N152 observed in the crystal structures. An alternative strategy to target N152 and simultaneously gain binding affinity would be the introduction of a polar group that could directly form a direct hydrogen bond with N152. Substituting the imidazole-N atom distal from the pyridine ring by a polar group could also contribute to a gain in affinity by formation of a direct hydrogen bond with K93 instead of a water-mediated hydrogen bond. Direct hydrogen bonds formed at this position have already been observed for p38 MAPK [317], [318]. However, achieving selectivity at this position might be difficult since the same residue is present in p38α MAPK (Figure 6.9). In addition to selectivity against p38α MAPK, a crucial aspect that needs to be considered when investigating this type of inhibitors is the intra-JNK selectivity. Compound **44** was tested on the three existing JNK isoforms and a similar inhibitory potency was determined for all of them. Nevertheless, a moderate preference (approximately 2.5- to 3-fold difference in binding affinity) was observed for JNK1 and JNK3 over JNK2 (see publication indicated at the beginning of this chapter [8]).

7   Conclusions

The following conclusions that can be drawn from this thesis:

1. The role of properdin in the activation of the alternative complement pathway

- Folded full-length properdin fused to c-myc or mCherry tags followed by a His_{10}-tag was successfully produced in the mammalian FreeStyle 293F cells.

- Purified protein fused to mCherry or after mCherry and His-tag removal was shown to form the previously reported cyclic structures. Size exclusion chromatograms revealed, however, a ratio of tetramers, trimers and dimers of properdin that differed from the previously reported ratios, which most likely resulted from the use of different experimental parameters.

- In the fusion protein, mCherry did not interfere neither with properdin oligomerization nor with its functional activity, demonstrating that this fluorescence reporter could be used to follow properdin expression and secretion as well as to study the binding of properdin to apoptotic and necrotic cells by fluorescence microscopy.

- Contrarily to published observations, recombinantly produced E244K properdin mutant was not monomeric and showed a oligomerization distribution similar to that of wild-type properdin. Therefore, crystallization of properdin for three-dimensional structural determination by X-ray crystallography was not pursued.

- Although no binding of properdin to glycans was observed in glycan microarrays, STD-NMR experiments showed that a specific set of glycans binds to properdin in solution and that factors like glycan charge, size and sulfation are critical for the interaction.

- Based on the performed experiments, GAGs are potential properdin PAMPs but further studies are required to evaluate their effect on complement activation in a physiological context. STD-NMR studies also showed that binding of both properdin and its antagonist FH to identical glycans occurs via recognition of different epitopes.

2. Production of recombinant heparinase I from *Pedobacter heparinus* for structure-based protein engineering

- The recombinant production of heparinase I from *P. heparinus* was successfully achieved using *E. coli*. CD spectroscopy and SEC analyses revealed that heparinase I was folded and mostly monomeric with an antiparallel β-sheet as the major secondary structural element.

- Oligomerization of HepIv1 was also observed. Formation of oligomers was shown to be dynamic and concentration-dependent. Native MS also demonstrated that denatured species were present in solution and increased upon prolonged storage.

- Protein degradation was visualized over time and was found to increase with increasing temperature. Recombinant heparinase I was shown to have a very low thermal stability.

- Despite different expression constructs comprising varied N- and C-terminus sequences and varying parameters such as protein concentrations, temperatures and buffer components, no proteins crystals were obtained.

- Enzymatic depolymerization of heparin by HepIv1 was monitored via the absorbance at 232 nm and shown to be dependent on the concentration of Ca^{2+}. The low activity of HepIv1 construct and the low amount of heparin-derived oligosaccharides obtained from this protein was most likely due to its thermal denaturation and subsequent inactivation.

- Modification of heparinase I by structure-based engineering could not be accomplished. The persistent thermal instability observed in this work suggests that, instead, the homologue crystal structures of heparinase I from *B. thetaiotaomicron* should be used in the future to alter the product profile of heparinase I.

3. Structural-based optimization of selective type I c-Jun N-terminal kinase 3 inhibitors

- The binding mode of selective type I JNK pyridinylimidazole inhibitors **38** and **44** was confirmed by X-ray crystallography. Both inhibitors revealed a similar binding mode in the ATP binding pocket, which matched the binding of previous pyridinylimidazole-based inhibitors.

- The small methyl group oriented towards the hydrophobic region I was most likely the selectivity driver since it presented the optimal length to target this region of JNK3.

- Contrary to docking studies no flip of the imidazole core was observed in the JNK3-**38** complex structure and the G-rich loop was only visible in the JNK3-**44** structure, pointing out the role of the *S*-methyl group present in **44** in the stabilization of this flexible region. The inward "collapse" of the G-rich loop, concurrent with a conformational stabilization, seemed to be the structural reason for the 2-fold difference in inhibitory potency of JNK3 between the two inhibitors.

- The JNK3-**44** complex showed the highest unfolding temperature when compared to the apo-form, AMP-PNP or JNK3-**38** complex, in agreement with a structural stabilization of the G-rich loop by this inhibitor.

- Targeting residues Q155, N152 or K93 of JNK3 for additional directed interactions are potential future strategies to increase the inhibitory potency of compound **44** while preserving its selectivity over p38α MAPK.

8 References

[1]   L. Pillemer, L. Blum, I. H. Lepow, O. A. Ross, E. W. Todd, and A. C. Wardlaw, The properdin system and immunity: demonstration and isolation of a new serum protein, properdin, and its role in immune phenomena, *Science*, vol. 120, pp. 279–285, 1954

[2]   R. A. Harrison, The properdin pathway: an "alternative activation pathway" or a "critical amplification loop" for C3 and C5 activation?, *Semin. Immunopathol.*, vol. 40, no. 1, pp. 15–35, 2018

[3]   J. Y. Chen, C. Cortes, and V. P. Ferreira, Properdin: A multifaceted molecule involved in inflammation and diseases, *Mol. Immunol.*, vol. 102, pp. 58–72, 2018

[4]   A. K. Korir and C. K. Larive, Advances in the separation, sensitive detection, and characterization of heparin and heparan sulfate, *Anal. Bioanal. Chem.*, vol. 393, no. 1, pp. 155–169, 2009

[5]   M. Suflita, L. Fu, W. He, M. Koffas, and R. J. Linhardt, Heparin and related polysaccharides: Synthesis using recombinant enzymes and metabolic engineering, *Appl. Microbiol. Biotechnol.*, vol. 99, no. 18, pp. 7465–7479, 2015

[6]   P. Koch, M. Gehringer, and S. A. Laufer, Inhibitors of c-Jun N-Terminal Kinases: An Update, *J. Med. Chem.*, pp. 72–95, 2015

[7]   M. Gehringer, F. Muth, P. Koch, and S. A. Laufer, c-Jun N-terminal kinase inhibitors: a patent review (2010 – 2014), *Expert Opin. Ther. Pat.*, vol. 25, no. 8, pp. 849–872, 2015

[8]   F. Ansideri *et al.*, Structural Optimization of a Pyridinylimidazole Scaffold: Shifting the Selectivity from p38α Mitogen-Activated Protein Kinase to c-Jun N-Terminal Kinase 3, *ACS Omega*, vol. 3, no. 7, pp. 7809–7831, 2018

[9]   D. Ricklin, G. Hajishengallis, K. Yang, and J. D. Lambris, Complement: a key system for immune surveillance and homeostasis, *Nat. Immunol.*, vol. 11, no. 9, pp. 785–97, 2010

[10]   P. Lachmann, Complement before molecular biology, *Mol. Immunol.*, vol. 43, no. 6, pp. 496–508, 2006

[11]   N. S. Merle, S. E. Church, V. Fremeaux-Bacchi, and L. T. Roumenina, Complement system part I - molecular mechanisms of activation and regulation, *Front. Immunol.*, vol. 6, no. 262, pp. 1–30, 2015

[12]   M. J. Walport, Complement First of Two Parts, *N Engl J Med*, vol. 344, no. 14, pp. 1058–1066, 2001

[13]   G. E. Ritchie, B. E. Moffatt, R. B. Sim, P. Morgan, R. A. Dwek, and P. M. Rudd, Glycosylation and the complement system, *Chem. Rev.*, vol. 102, pp. 305–319, 2002

[14]   G. Bajic, S. E. Degn, S. Thiel, and G. R. Andersen, Complement activation, regulation, and molecular basis for complement-related diseases, *EMBO J.*, vol. 34, no. 22, pp. 2735–2757, 2015

[15]   J. D. Lambris, D. Ricklin, and B. V Geisbrecht, Complement evasion by human pathogens., *Nat. Rev. Microbiol.*, vol. 6, no. 2, pp. 132–42, 2008

[16]   C. A. Diebolder *et al.*, Complement is activated by IgG hexamers assembled at the cell surface, *Science*, vol. 343, no. 6176, pp. 1260–1263, 2014

[17]   T. Fujita, M. Matsushita, and Y. Endo, The lectin-complement pathway - Its role in innate immunity and evolution, *Immunol. Rev.*, vol. 198, pp. 185–202, 2004

[18]   P. J. Lachmann, The amplification loop of the complement pathways, *Adv. Immunol.*, vol. 104, pp. 115–149, 2009

[19]   D. R. Mathern and P. S. Heeger, Molecules great and small: The complement system, *Clin. J. Am. Soc. Nephrol.*, vol. 10, no. 9, pp. 1636–1650, 2015

[20]   D. Ricklin, E. S. Reis, D. C. Mastellos, P. Gros, and J. D. Lambris, Complement component C3 - The "Swiss Army Knife" of innate immunity and host defense, *Immunol. Rev.*, vol. 274, no. 1, pp. 33–58, 2016

[21]   H. J. Muller-Eberhard, Molecular organization and function of the complement system, *Annu. Rev. Biochem.*, vol. 57, pp. 321–347, 1988

[22]   P. J. Haas, Anaphylatoxins: their role in bacterial infection and inflammation, *Immunol. Res.*, vol. 37, no. 3, pp. 161–175, 2007

[23]   B. J. C. Janssen, A. Christodoulidou, A. McCarthy, J. D. Lambris, and P. Gros, Structure of C3b reveals conformational changes that underlie complement activity, *Nature*, vol. 444, pp. 213–6, 2006

[24]   M. K. Pangburn, V. P. Ferreira, and C. Cortes, Discrimination between host and pathogens by the complement system, *Vaccine*, vol. 26, pp. 1–17, 2008

[25]   J. R. Dunkelberger and W. C. Song, Complement and its role in innate and adaptive immune responses, *Cell Res.*, vol. 20, no. 1, pp. 34–50, 2010

[26]   J. P. Atkinson and T. Farries, Separation of self from non-self in the complement system, *Immunol. Today*, vol. 8, no. 7–8, pp. 212–215, 1987

[27]   M. K. Pangburn, R. D. Schreiber, and H. J. Muller-Eberhard, Formation of the initial C3 convertase of the alternative pathway. Acquisition of C3b-like activities by spontaneous hydrolysis of the putative thioester in native C3., *J. Exp. Med.*, vol. 154, pp. 856–867, 1981

[28]   O. A. Hamad *et al.*, Contact activation of C3 enables tethering between activated platelets and polymorphonuclear leukocytes via CD11b/CD18, *Thromb. Haemost.*, vol. 114, no. 6, pp. 1207–1217, 2015

[29]   B. Nilsson and K. Nilsson Ekdahl, The tick-over theory revisited: Is C3 a contact-activated protein?, *Immunobiology*, vol. 217, no. 11, pp. 1106–1110, 2012

[30]   E. L. G. Pryzdial and D. E. Isenman, Alternative complement pathway activation fragment Ba binds to C3b. Evidence that formation of the factor B-C3b complex involves two discrete points fo contact., *J. Biol. Chem.*, vol. 262, no. 4, pp. 1519–1525, 1987

[31]   B. J. Janssen *et al.*, Insights into complement convertase formation based on the structure of the factor B-cobra venom factor complex, *EMBO J.*, vol. 28, pp. 2469–

2478, 2009

[32]   M. K. Pangburn and H. J. Müller-Eberhard, The C3 convertase of the alternative pathway of human complement, *Biochem. J.*, vol. 235, pp. 723–730, 1986

[33]   M. Martin and A. M. Blom, Complement in removal of the dead – balancing inflammation, *Immunol. Rev.*, vol. 274, no. 1, pp. 218–232, 2016

[34]   V. M. Holers, Complement and Its Receptors: New Insights into Human Disease, *Annu. Rev. Immunol.*, vol. 32, no. 1, pp. 433–459, 2014

[35]   M. K. Pangburn and N. Rawal, Structure and function of complement C5 convertase enzymes, *Biochem. Soc. Trans.*, vol. 30, pp. 1006–1010, 2002

[36]   E. T. M. Berends *et al.*, Molecular insights into the surface-specific arrangement of complement C5 convertase enzymes, *BMC Biol.*, vol. 13, no. 1, pp. 1–13, 2015

[37]   D. T. Fearon and K. F. Austen, Properdin: binding to C3b and stabilization of the C3b-dependent C3 convertase, *J. Exp. Med.*, vol. 142, pp. 856–863, 1975

[38]   D. T. Fearon and K. F. Austen, The alternartive pathway of complement - a system for host resistance to microbial infection, *N. Engl. J. Med.*, vol. 303, no. 5, pp. 259–263, 1980

[39]   Y. van Kooyk and G. A. Rabinovich, Protein-glycan interactions in the control of innate and adaptive immune responses, *Nat. Immunol.*, vol. 9, no. 6, pp. 593–601, 2008

[40]   A. Langford-Smith, A. J. Day, P. N. Bishop, and S. J. Clark, Complementing the sugar code: Role of GAGs and sialic acid in complement regulation, *Front. Immunol.*, vol. 6, pp. 1–7, 2015

[41]   C. Kemper, J. P. Atkinson, and D. E. Hourcade, Properdin: emerging roles of a pattern-recognition molecule, *Annu. Rev. Immunol.*, vol. 28, pp. 131–55, 2010

[42]   A. Z. Blatt, S. Pathan, and V. P. Ferreira, Properdin: a tightly regulated critical inflammatory modulator, *Immunol. Rev.*, vol. 274, no. 1, pp. 172–190, 2016

[43]   M. Noris and G. Remuzzi, Overview of complement activation and regulation, *Semin. Nephrol.*, vol. 33, no. 6, pp. 479–492, 2013

[44]   P. Gros, F. J. Milder, and B. J. C. Janssen, Complement driven by conformational changes, *Nat. Rev. Immunol.*, vol. 8, no. 1, pp. 48–58, 2008

[45]   C. C. Ferreira V. P., Pangburn M. K., Complement control protein factor H: the good, the bad, and the inadequate, *Mol. Immunol.*, vol. 47, no. 13, pp. 2187–2197, 2011

[46]   J. M. Thurman and V. M. Holers, The Central Role of the Alternative Complement Pathway in Human Disease, *J. Immunol.*, vol. 176, no. 3, pp. 1305–1310, 2006

[47]   M. K.Liszewski and J. P. Atkinson, Complement regulators in human disease: Lessons from modern genetics, *J. Intern. Med.*, vol. 277, no. 3, pp. 294–305, 2015

[48]   H. Ojha, H. S. Panwar, R. D. Gorham, D. Morikis, and A. Sahu, Viral regulators of complement activation: Structure, function and evolution, *Mol. Immunol.*, vol. 61, no. 2, pp. 89–99, 2014

[49]   M. K. Pangburn, Host recognition and target differentiation by factor H, a regulator of the alternative pathway of complement, *Immunopharmacology*, vol. 49, no. 1–2, pp. 149–157, 2000

[50] A. Kopp, M. Hebecker, E. Svobodová, and M. Józsi, Factor H: A complement regulator in health and disease, and a mediator of cellular interactions, *Biomolecules*, vol. 2, no. 1, pp. 46–75, 2012

[51] S. Meri and M. K. Pangburn, Discrimination between activators and nonactivators of the alternative pathway of complement: regulation via a sialic acid/polyanion binding site on factor H 438, *Proc Natl Acad Sci U S A*, vol. 87, pp. 3982–3986, 1990

[52] R. B. Sim, A. J. Day, B. E. Moffatt, and M. Fontaine, Complement Factor I and cofactors in control of complement system convertase enzymes, *Methods Enzymol.*, vol. 223, pp. 13–35, 1993

[53] S. Rodríguez de Córdoba, J. Esparza-Gordillo, E. Goicoechea de Jorge, M. Lopez-Trascasa, and P. Sánchez-Corral, The human complement factor H: functional roles, genetic variations and disease associations, *Mol. Immunol.*, vol. 41, no. 4, pp. 355–67, 2004

[54] L. Kouser, M. Abdul-Aziz, A. Nayak, C. M. Stover, R. B. Sim, and U. Kishore, Properdin and factor h: opposing players on the alternative complement pathway "see-saw," *Front. Immunol.*, vol. 4, p. 93, 2013

[55] T. C. Farries, P. J. Lachmann, and R. A. Harrisont, Analysis of the interactions between properdin, the third component of complement (C3), and its physiological activation products, *Biochem. J.*, vol. 252, pp. 47–54, 1988

[56] M. Alcorlo, A. Tortajada, S. Rodriguez de Cordoba, and O. Llorca, Structural basis for the stabilization of the complement alternative pathway C3 convertase by properdin, *Proc. Natl. Acad. Sci.*, vol. 110, no. 33, pp. 13504–13509, 2013

[57] K. F. Nolan, S. Kaluz, K. B. M. Reid, M. P. Dierich, and W. Schwaeble, Molecular cloning of the cDNA coding for properdin, a positive regulator of the alternative pathway of human complement, *Eur. J. Immunol.*, vol. 21, no. 3, pp. 771–776, 1991

[58] C. Cortes, J. a Ohtola, G. Saggu, and V. P. Ferreira, Local release of properdin in the cellular microenvironment: role in pattern recognition and amplification of the alternative pathway of complement, *Front. Immunol.*, vol. 3, p. 412, 2013

[59] M. K. Pangburn, Analysis of the natural polymeric forms of human properdin and their functions in complement activation, *J. Immunol.*, vol. 142, no. 1, pp. 0202–0207, 1989

[60] C. A. Smith, M. K. Pangburn, C. W. Vogel, and H. J. Muller-Eberhard, Molecular architecture of human properdin, a positive regulator of the alternative pathway of complement, *J. Biol. Chem.*, vol. 259, no. 7, pp. 4582–4588, 1984

[61] K. F. Nolan and K. B. M. Reid, Complete primary structure of human properdin - a positive regulator of the alternative pathway of the serum complement system, *Biochem.Soc.Trans.*, vol. 18, pp. 1161–1162, 1990

[62] Z. Sun, K. B. M. Reid, and S. J. Perkins, The dimeric and trimeric solution structures of the multidomain complement protein properdin by X-ray scattering, analytical ultracentrifugation and constrained modelling, *J. Mol. Biol.*, vol. 343, no. 5, pp. 1327–43, 2004

[63] K. Tan *et al.*, Crystal structure of the TSP-1 type 1 repeats: a novel layered fold and its

biological implication, *J. Cell Biol.*, vol. 159, no. 2, pp. 373–82, 2002

[64]   D. V. Pedersen, M. Revel, T. A. F. Gadeberg, and G. R. Andersen, Crystallization and X-ray analysis of monodisperse human properdin, *Acta Crystallogr. Sect. F Struct. Biol. Commun.*, vol. 75, no. 2, pp. 1–5, 2019

[65]   S. Hartmann and J. Hofsteenge, Properdin, the positive regulator of complement, is highly C-mannosylated, *J. Biol. Chem.*, vol. 275, no. 37, pp. 28569–74, 2000

[66]   J. M. Higgins, H. Wiedemann, R. Timpl, and K. B. Reid, Characterization of mutant forms of recombinant human properdin lacking single thrombospondin type I repeats. Identification of modules important for function, *J. Immunol.*, vol. 155, no. 12, pp. 5777–5785, 1995

[67]   L. A. Halim, F. Liu, A. J. R. Heck, Y. Yang, V. Franc, and H. Schellekens, Hybrid mass spectrometry approaches in glycoprotein analysis and their usage in scoring biosimilarity, *Nat. Commun.*, vol. 7, no. 1, pp. 1–10, 2016

[68]   J. F. G. Vliegenthart, The impact of defining glycan structures, *Perspect. Sci.*, vol. 11, pp. 3–10, 2017

[69]   M. V. Perdikoulis, U. Kishore, and K. B. M. Reid, Expression and characterisation of the thrombospondin type I repeats of human properdin, *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.*, vol. 1548, no. 2, pp. 265–277, 2001

[70]   P. Bertram, A. M. Akk, H. Zhou, L. M. Mitchell, C. T. N. Pham, and D. E. Hourcade, Anti-Mouse Properdin TSR 5/6 Monoclonal Antibodies Block Complement Alternative Pathway-dependent Pathogenesis, *Monoclon. Antib. Immunodiagn. Immunother.*, vol. 34, no. 1, pp. 1–6, 2015

[71]   D. V Pedersen *et al.*, Functional and structural insight into properdin control of complement alternative pathway amplification, *EMBO J.*, vol. 36, no. 8, pp. 1084–1099, 2017

[72]   A. Z. Blatt *et al.*, Properdin-Mediated C5a Production Enhances Stable Binding of Platelets to Granulocytes in Human Whole Blood, *J. Immunol.*, vol. 196, no. 11, pp. 4671–4680, 2016

[73]   J. Figueroa, J. Andreoni, and P. Densen, Complement deficiency states and meningococcal disease, *Immunol. Res.*, vol. 12, no. 3, pp. 295–311, 1993

[74]   L. Skattum, M. Van Deuren, T. Van Der Poll, and L. Truedsson, Complement deficiency states and associated infections, *Mol. Immunol.*, vol. 48, no. 14, pp. 1643–1655, 2011

[75]   C. A. P. Fijen *et al.*, Properdin deficiency : molecular basis and disease association, *Mol. Immunol.*, vol. 36, pp. 863–867, 1999

[76]   D. E. Hourcade, The role of properdin in the assembly of the alternative pathway C3 convertases of complement, *J. Biol. Chem.*, vol. 281, no. 4, pp. 2128–2132, 2006

[77]   C. Kemper and D. E. Hourcade, Properdin: New roles in pattern recognition and target clearance, *Mol. Immunol.*, vol. 45, no. 16, pp. 4048–56, 2008

[78]   D. Spitzer, L. M. Mitchell, J. P. Atkinson, and D. E. Hourcade, Properdin Can Initiate Complement Activation by Binding Specific Target Surfaces and Providing a Platform

for De Novo Convertase Assembly, *J. Immunol.*, vol. 179, no. 4, pp. 2600–2608, 2007

[79] Y. Kimura, T. Miwa, L. Zhou, and W. Song, Activator-specific requirement of properdin in the initiation and, *Immunobiology*, vol. 111, no. 2, pp. 732–741, 2008

[80] C. Kemper, L. M. Mitchell, L. Zhang, and D. E. Hourcade, The complement protein properdin binds apoptotic T cells and promotes complement activation and phagocytosis, *Proc. Natl. Acad. Sci.*, vol. 105, no. 26, pp. 9023–9028, 2008

[81] W. Xu *et al.*, Properdin Binds to Late Apoptotic and Necrotic Cells Independently of C3b and Regulates Alternative Pathway Complement Activation, *J. Immunol.*, vol. 180, no. 11, pp. 7613–7621, 2008

[82] C. Cortes, V. P. Ferreira, and M. K. Pangburn, Native properdin binds to Chlamydia pneumoniae and promotes complement activation, *Infect. Immun.*, vol. 79, no. 2, pp. 724–731, 2011

[83] A. Zaferani *et al.*, Identification of tubular heparan sulfate as a docking platform for the alternative complement component properdin in proteinuric renal disease, *J. Biol. Chem.*, vol. 286, no. 7, pp. 5359–5367, 2011

[84] T. C. Farries, J. T. Finch, P. J. Lachmann, and R. A. Harrison, Resolution and analysis of "native" and "activated" properdin, *Biochem J*, vol. 243, no. 2, pp. 507–517, 1987

[85] S. Agarwal, V. P. Ferreira, C. Cortes, M. K. Pangburn, P. A. Rice, and S. Ram, An Evaluation of the Role of Properdin in Alternative Pathway Activation on *Neisseria meningitidis* and *Neisseria gonorrhoeae*, *J. Immunol.*, vol. 185, no. 1, pp. 507–516, 2010

[86] V. P. Ferreira, C. Cortes, and M. K. Pangburn, Native polymeric forms of properdin selectively bind to targets and promote activation of the alternative pathway of complement, *Immunobiology*, vol. 215, no. 11, pp. 932–940, 2010

[87] G. Saggu, C. Cortes, H. N. Emch, G. Ramirez, R. G. Worth, and V. P. Ferreira, Identification of a Novel Mode of Complement Activation on Stimulated Platelets Mediated by Properdin and C3(H_2O), *J. Immunol.*, vol. 190, no. 12, pp. 6457–6467, 2013

[88] V. P. Ferreira, C. Cortes, and M. K. Pangburn, Native polymeric forms of properdin selectively bind to targets and promote activation of the alternative pathway of complement, *Immunobiology*, vol. 215, no. 11, pp. 932–940, 2010

[89] H. Yu, E. V Muñoz, R. E. Edens, and R. J. Linhardt, Kinetic studies on the interactions of heparin and complement proteins using surface plasmon resonance, *Biochim. Biophys. Acta*, vol. 1726, no. 2, pp. 168–176, 2005

[90] A. Zaferani *et al.*, Identification of tubular heparan sulfate as a docking platform for the alternative complement component properdin in proteinuric renal disease, *J. Biol. Chem.*, vol. 286, no. 7, pp. 5359–5367, 2011

[91] G. D. Holt, M. K. Pangburn, and V. Ginsburg, Properdin binds to sulfatide [Gal(3-SO_4)β1-1Cer] and has a sequence homology with other proteins that bind sulfated glycoconjugates, *J. Biol. Chem.*, vol. 265, no. 5, pp. 2852–2855, 1990

[92] M. Harboe *et al.*, Properdin binding to complement activating surfaces depends on

initial C3b deposition, *Proc. Natl. Acad. Sci.*, vol. 114, no. 4, pp. E534–E539, 2017

[93] A. Sahu, B. K. Kay, and J. D. Lambris, Inhibition of human complement by a C3-binding peptide isolated from a phage-displayed random peptide library, *J. Immunol.*, vol. 157, no. 2, pp. 884–91, 1996

[94] M. Harboe *et al.*, The Role of Properdin in Zymosan- and *Escherichia coli*-Induced Complement Activation, *J. Immunol.*, vol. 189, no. 5, pp. 2606–2613, 2012

[95] A. M. Lesher, B. Nilsson, and W. C. Song, Properdin in complement activation and tissue injury, *Mol. Immunol.*, vol. 56, no. 3, pp. 191–198, 2013

[96] J. D. Marth, A unified vision of the building blocks of life, *Nat. Cell Biol.*, vol. 10, no. 9, pp. 1015–1015, 2008

[97] M. Cohen, Notable aspects of glycan-protein interactions, *Biomolecules*, vol. 5, no. 3, pp. 2056–2072, 2015

[98] S. J. Clark, P. N. Bishop, and A. J. Day, The proteoglycan glycomatrix: A sugar microenvironment essential for complement regulation, *Front. Immunol.*, vol. 4, pp. 1–4, 2013

[99] V. S. Mahajan and S. Pillai, Sialic acids and autoimmune disease, *Immunol. Rev.*, vol. 269, no. 1, pp. 145–161, 2016

[100] M. K. Pangburn and H. J. Müller-Eberhard, Complement C3 convertase: cell surface restriction of beta1H control and generation of restriction on neuraminidase-treated cells, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 75, no. 5, pp. 2416–20, 1978

[101] B. S. Blaum, J. P. Hannan, A. P. Herbert, D. Kavanagh, D. Uhrín, and T. Stehle, Structural basis for sialic acid-mediated self-recognition by complement factor H, *Nat. Chem. Biol.*, vol. 11, no. 1, pp. 77–82, 2015

[102] C. Kemper, J. P. Atkinson, and D. E. Hourcade, Properdin: Emerging Roles of a Pattern-Recognition Molecule, *Annu. Rev. Immunol.*, vol. 28, no. 1, pp. 131–155, 2010

[103] J. van den Born *et al.*, Factor H and Properdin Recognize Different Epitopes on Renal Tubular Epithelial Heparan Sulfate, *J. Biol. Chem.*, vol. 287, no. 37, pp. 31471–31481, 2012

[104] C. K. M. Tripathi, J. Banga, and V. Mishra, Microbial heparin/heparan sulphate lyases: potential and applications, *Appl. Microbiol. Biotechnol.*, vol. 94, pp. 307–321, 2012

[105] A. Imberty, H. Lortat-Jacob, and S. Pérez, Structural view of glycosaminoglycan-protein interactions, *Carbohydr. Res.*, vol. 342, no. 3–4, pp. 430–439, 2007

[106] K. R. Taylor and R. L. Gallo, Glycosaminoglycans and their proteoglycans: host-associated molecular patterns for initiation and modulation of inflammation, *FASEB J.*, vol. 20, no. 1, pp. 9–22, 2006

[107] S. C. Hung and M. M. L. Zulueta, Glycochemical Synthesis: Strategies and Applications, *Glycochemical Synth. Strateg. Appl.*, pp. 1–550, 2015

[108] M. Mende *et al.*, Chemical Synthesis of Glycosaminoglycans, *Chem. Rev.*, vol. 116, no. 14, pp. 8193–8255, 2016

[109] S. Tsuchiya *et al.*, Implementation of GlycanBuilder to draw a wide variety of ambiguous glycans, *Carbohydr. Res.*, vol. 445, pp. 104–116, 2017

[110] Z. Shriver and W. K. Street, Heparin and Heparan Sulfate: Analyzing Structure and Microheterogeneity, *Handb. Exp. Pharmacol.*, vol. 207, pp. 159–176, 2012

[111] D. L. Rabenstein, Heparin and heparan sulfate: Structure and function, *Nat. Prod. Rep.*, vol. 19, no. 3, pp. 312–331, 2002

[112] S. I. Lamberg and A. C. Stoolmiller, Glycosaminoglycans. A biochemical and clinical review., *J. Invest. Dermatol.*, vol. 63, no. 6, pp. 433–449, 1974

[113] M. Salmivirta, K. Lidholt, and U. Lindahl, Heparan sulfate: a piece of information, *FASEB J.*, vol. 10, pp. 1270–1279, 1996

[114] D. Soares da Costa, R. L. Reis, and I. Pashkuleva, Sulfation of Glycosaminoglycans and Its Implications in Human Health and Disorders, *Annu. Rev. Biomed. Eng.*, vol. 19, no. 1, pp. 1–26, 2017

[115] R. Sasisekharan and G. Venkataraman, Heparin and heparan sulfate : biosynthesis , structure and function, *Curr. Opin. Chem. Biol.*, vol. 4, pp. 626–631, 2000

[116] R. Jagsi *et al.*, Bioengineered heparins and heparan sulfates, *Adv Drug Deliv Rev*, vol. 97, pp. 237–249, 2016

[117] U. Lindahl, M. Kusche-Gullberg, and L. Kjellen, Regulated diversity of heparan sulfate, *J. Biol. Chem.*, vol. 273, no. 39, pp. 24979–24982, 1998

[118] E. Gray, B. Mulloy, and T. W. Barrowcliffe, Heparin and low-molecular-weight heparin, *Thromb. Haemost.*, vol. 99, no. 5, pp. 807–818, 2008

[119] M. A. Lima *et al.*, Heparan sulfate and heparin: Structure and function, *J. R. Soc. Interface*, vol. 12, no. 110, pp. 1–13, 2015

[120] R. J. Linhardt, 2003 Claude S. Hudson Award Address in Carbohydrate Chemistry. Heparin: Structure and Activity, *J. Med. Chem.*, vol. 46, no. 13, pp. 2551–2564, 2003

[121] M. Petitou, R. N. Pike, R. W. Carrell, R. Skinner, L. Jin, and J. P. Abrahams, The anticoagulant activation of antithrombin by heparin, *Proc. Natl. Acad. Sci.*, vol. 94, no. 26, pp. 14683–14688, 2002

[122] B. Casu, Structure and active domains of heparin, in *Chemistry and biology of heparin and heparan sulfate*, Elvesier, 2005, pp. 1–28

[123] V. Toschi and M. Lettino, Fondaparinux: Pharmacology and Clinical Experience in Cardiovascular Medicine, *Mini-Reviews Med. Chem.*, vol. 7, no. 4, pp. 383–387, 2007

[124] C. T. Tsai, M. M. L. Zulueta, and S. C. Hung, Synthetic heparin and heparan sulfate: probes in defining biological functions, *Curr. Opin. Chem. Biol.*, vol. 40, pp. 152–159, 2017

[125] M. Lima, T. Rudd, and E. Yates, New applications of heparin and other glycosaminoglycans, *Molecules*, vol. 22, no. 5, pp. 1–11, 2017

[126] N. S. Gandhi and R. L. Mancera, The structure of glycosaminoglycans and their interactions with proteins, *Chem. Biol. Drug Des.*, vol. 72, no. 6, pp. 455–482, 2008

[127] R. E. Hileman, A. E. Smith, T. Toida, and R. J. Linhardt, Preparation and structure of heparin lyase-derived heparan sulfate oligosaccharides, *Glycobiology*, vol. 7, no. 2, pp. 231–239, 1997

[128] A. K. Powell, Y. A. Ahmed, E. A. Yates, and J. E. Turnbull, Generating heparan sulfate

saccharide libraries for glycomics applications, *Nat. Protoc.*, vol. 5, no. 5, pp. 829–841, 2010

[129] S. Brown and B. Kuberan, Production of size-defined heparosan, heparan sulfate, and heparin oligosaccharides by enzymatic depolzymerization, *Methods Mol. Biol.*, vol. 1229, pp. 21–29, 2015

[130] T. N. Laremore, F. Zhang, J. S. Dordick, J. Liu, and R. J, Recent progress and applications in glycosaminoglycan and heparin research, *Curr. Opin. Struct. Biol.*, vol. 13, pp. 633–640, 2009

[131] L. Kjellén and U. Lindahl, Specificity of glycosaminoglycan–protein interactions, *Curr. Opin. Struct. Biol.*, vol. 50, pp. 101–108, 2018

[132] P. H. Liang, C. Y. Wu, W. A. Greenberg, and C. H. Wong, Glycan arrays: biological and medical applications, *Curr. Opin. Chem. Biol.*, vol. 12, no. 1, pp. 86–92, 2008

[133] R. J. Linhardt, J. E. Turnbull, H. M. Wang, D. Loganathan, and J. T. Gallagher, Examination of the Substrate Specificity of Heparin and Heparan Sulfate Lyases, *Biochemistry*, vol. 29, no. 10, pp. 2611–2617, 1990

[134] Z. Wei, M. Lyon, and J. T. Gallagher, Distinct substrate specificities of bacterial heparinases against N-unsubstituted glucosamine residues in heparan sulfate, *J. Biol. Chem.*, vol. 280, no. 16, pp. 15742–15748, 2005

[135] M. Silva and C. Dietrich, Structure of heparin. Characterization of the products from heparin by the action of a heparinase and a heparitinase from *Flavobacterium heparinum*, *Biol. Chem.*, vol. 250, no. 10, pp. 6841–6846, 1975

[136] I. Bhushan, A. Alabbas, J. C. Sistla, R. Saraswat, U. R. Desai, and R. B. Gupta, Heparin depolymerization by immobilized heparinase: a review, *Int. J. Biol. Macromol.*, vol. 99, pp. 721–730, 2017

[137] I. Bhushan, A. Alabbas, B. Kuberan, R. B. Gupta, and U. R. Desai, Immobilization alters heparin cleaving properties of heparinase I, *Glycobiology*, vol. 27, no. 11, pp. 994–998, 2017

[138] D. Korn and A. Nail, The degradation of heparin by bacterial enzymes: I. adaption and lyophilized cells, *J. Biol. Chem.*, vol. 223, pp. 853–858, 1956

[139] P. L. Steyn, P. Segers, M. Vancanneyt, P. Sandra, K. Kersters, and J. J. Joubert, Classification of heparinolytic bacteria into a new genus, *Pedobacter*, comprising four species: *Pedobacter heparinus* comb. nov., *Pedobacter piscium* comb. nov., *Pedobacter africanus* sp. nov. and *Pedobacter saltans* sp. nov. Proposal of the family Sphingobac, *Int. J. Syst. Bacteriol.*, vol. 48, no. 1, pp. 165–177, 1998

[140] T. Him *et al.*, Complete genome sequence of *Pedobacter heparinus*, *Stand. Genomic Sci.*, vol. 1, pp. 54–62, 2009

[141] S. Ernst, R. Langer, C. L. Cooney, and R. Sasisekharan, Enzymatic degradation of glycosaminogiycans, *Crit. Rev. Biochem. Mol. Biol.*, vol. 30, no. 5, pp. 387–444, 1995

[142] M. L. Garron and M. Cygler, Uronic polysaccharide degrading enzymes, *Curr. Opin. Struct. Biol.*, vol. 28, no. 1, pp. 87–95, 2014

[143] B. T. Kim, W. S. Kim, Y. S. Kim, R. J. Linhardt, and D. H. Kim, Purification and

characterization of a novel heparinase from *Bacteroides stercoris* HJ-15, *J. Biochem.*, vol. 128, no. 2, pp. 323–328, 2000

[144] A. Cartmell *et al.*, How members of the human gut microbiota overcome the sulfation problem posed by glycosaminoglycans, *Proc. Natl. Acad. Sci.*, vol. 114, no. 27, pp. 7037–7042, 2017

[145] D. L. Lohse and R. J. Linhardtg, Purification and Characterization of Heparin Lyases from *Flavobacterium heparinum*, *J. Biol. Chem.*, vol. 267, no. 34, pp. 24347–24355, 1992

[146] M. L. Garron and M. Cygler, Structural and mechanistic classification of uronic acid-containing polysaccharide lyases, *Glycobiology*, vol. 20, no. 12, pp. 1547–1573, 2010

[147] Z. Joseph J. F., O. Karen, L. Robert, and C. Charles L., The release of heparinase from the periplasmic space of *Flavobacterium heparinum* by three-step osmotic shock, *Appl. Biochem. Biotechnol.*, vol. 30, no. 2, pp. 137–148, 1991

[148] R. Godavarti and R. Sasisekharan, A Comparative Analysis of the Primary Sequences and Characteristics of Heparinases I, II, and III from *Flavobacterium heparinum*, *Biochem. Biophys. Res. Commounications*, vol. 229, no. 1879, pp. 770–777, 1996

[149] U. R. Desai, H. Wang, and R. J. Linhardt, Specificity Studies on the Heparin Lyases from *Flavobacterium heparinum*, *Biochemistry*, vol. 32, pp. 8140–8145, 1993

[150] U. R. Desai, H. M. Wang, and R. J. Linhardt, Substrate Specificity of the Heparin Lyases from *Flavobacterium heparinum*, *Archives of Biochemistry and Biophysics*, vol. 306, no. 2. pp. 461–468, 1993

[151] R. J. Linhardt, P. M. Galliher, and C. L. Cooney, Polysaccharide Lyases, *Appl. Biochem. Biotechnol.*, vol. 12, pp. 135–176, 1986

[152] P. Gacesa, Alginate-modifying enzymes: A proposed unified mechanism of action for the lyases and epimerases, *Science*, vol. 212, no. 2, pp. 199–202, 1987

[153] V. L. Y. Yip and S. G. Withers, Breakdown of oligosaccharides by the process of elimination, *Curr. Opin. Chem. Biol.*, vol. 10, no. 2, pp. 147–155, 2006

[154] G. J. Davies, K. S. Wilson, and B. Henrissat, Nomenclature for sugar-binding subsites in glycosyl hydrolase, *Biochem. J.*, vol. 321, pp. 557–9, 1997

[155] J. F. K. Limtiaco, S. Beni, C. J. Jones, D. J. Langeslay, and C. K. Larive, NMR methods to monitor the enzymatic depolymerization of heparin, *Anal. Bioanal. Chem.*, vol. 399, no. 2, pp. 593–603, 2011

[156] Pervin A, Gallo C, Jandik KA, Han X-J, and L. RJ, Preparation and structural characterization of large heparin-derived oligosaccharides, *Glycobiology*, vol. 5, no. 1, pp. 83–95, 1995

[157] C. Thanawiroon, K. G. Rice, T. Toida, and R. J. Linhardt, LC/MS sequencing approach for highly sulfated heparin-derived oligosaccharides, *J. Biol. Chem.*, vol. 279, no. 4, pp. 2608–2615, 2004

[158] J. Wu, C. Zhang, X. Mei, Y. Li, and X. H. Xing, Controllable production of low molecular weight heparins by combinations of heparinase I/II/III, *Carbohydr. Polym.*, vol. 101, no. 1, pp. 484–492, 2014

[159] D. Shaya *et al.*, Crystal structure of heparinase II from *Pedobacter heparinus* and its complex with a disaccharide product, *J. Biol. Chem.*, vol. 281, no. 22, pp. 15525–15535, 2006

[160] W. Hashimoto, Y. Maruyama, Y. Nakamichi, B. Mikami, and K. Murata, Crystal structure of *Pedobacter heparinus* heparin lyase hep III with the active site in a deep cleft, *Biochemistry*, vol. 53, no. 4, pp. 777–786, 2014

[161] Y. Gu, M. Lu, Z. Wang, X. Wu, and Y. Chen, Expanding the Catalytic Promiscuity of Heparinase III from *Pedobacter heparinus*, *Chem. - A Eur. J.*, vol. 23, no. 11, pp. 2548–2551, 2017

[162] V. C. Yang, R. J. Linhardt, H. Bernsteincvd, and C. L. Cooneya, Purification and characterization of heparinase from *Flavobacterium heparinum*, *J. Biol. Chem.*, vol. 260, no. 3, pp. 1849–1857, 1985

[163] R. S. Godavarti, Protein Engineering of Heparinase I- Elucidation of Structure-Activity Relationships, Massachusetts Institute of Technology, 1996

[164] R. A. M. Sasisekharan, M. Bulmert, K. W. Moremen, C. L. Cooney, and R. Langer, Cloning and expression of heparinase I gene from *Flavobacterium heparinum*, *Proc. Natl. Acad. Sci.*, vol. 90, pp. 3660–3664, 1993

[165] R. Sasisekharan, D. Leckband, Z. J. R. Godavarti, G. Venkataraman, C. L. Cooney, and R. Langerfl, Heparinase I from *Flavobacterium heparinum*: The Role of the Cysteine Residue in Catalysis as Probed by Chemical Modification and Site-Directed Mutagenesis ?, *Biochemistry*, vol. 34, pp. 14441–14448, 1995

[166] R. Godavarti, C. L. Cooney, R. Langer, and R. Sasisekharan, Heparinase I from *Flavobacterium heparinum*. Identification of a critical histidine residue essential for catalysis as probed by chemical modification and site-directed mutagenesis, *Biochemistry*, vol. 35, no. 21, pp. 6846–6852, 1996

[167] R. Godavarti and R. Sasisekharan, Heparinase I from *Flavobacterium heparinum*. Role of positive charge in enzymatic activity, *J. Biol. Chem.*, vol. 273, no. 1, pp. 248–255, 1998

[168] D. Liu, Z. Shriver, R. Godavarti, G. Venkataraman, and R. Sasisekharan, The calcium-binding sites of heparinase I from *Flavobacterium heparinum* are essential for enzymatic activity, *J. Biol. Chem.*, vol. 274, no. 7, pp. 4089–4095, 1999

[169] R. Sasisekharan, G. Venkataraman, S. Ernst, C. L. Cooney, and R. Langer, Heparinase I from *Flavobacterium heparinum*. Mapping and characterization of the heparin binding domain, *J. Biol. Chem.*, vol. 6, no. 9, pp. 3124–3131, 1996

[170] Z. Shriver, D. Liu, Y. Hu, and R. Sasisekharan, Biochemical investigations and mapping of the calcium-binding sites of heparinase I from *Flavobacterium heparinum*, *J. Biol. Chem.*, vol. 274, no. 7, pp. 4082–4088, 1999

[171] S. Ernst, A. J. Rhomberg, K. Biemann, and R. Sasisekharan, Direct evidence for a predominantly exolytic processive mechanism for depolymerization of heparin-like glycosaminoglycans by heparinase I, *Proc. Natl. Acad. Sci.*, vol. 95, no. 8, pp. 4182–4187, 1998

[172] A. J. Rhomberg, Z. Shriver, K. Biemann, and R. Sasisekharan, Mass spectrometric evidence for the enzymatic mechanism of the depolymerization of heparin-like glycosaminoglycans by heparinase II, *Proc. Natl. Acad. Sci.*, vol. 95, no. 21, pp. 12232–12237, 1998

[173] F. Galeotti and N. Volpi, Oligosaccharide mapping of heparinase I-treated heparins by hydrophilic interaction liquid chromatography separation and online fluorescence detection and electrospray ionization-mass spectrometry characterization, *J. Chromatogr. A*, vol. 1445, pp. 68–79, 2016

[174] Y. H. Han *et al.*, Structural snapshots of heparin depolymerization by heparin lyase I, *J. Biol. Chem.*, vol. 284, no. 49, pp. 34019–34027, 2009

[175] M. J. Jedrzejas, Structural and functional comparison of polysaccharide-degrading enzymes, *Crit. Rev. Biochem.*, vol. 35, no. 3, pp. 221–251, 2000

[176] L. C. Macdonald and B. W. Berger, Insight into the role of substrate-binding residues in conferring substrate specificity for the multifunctional polysaccharide lyase Smlt1473, *J. Biol. Chem.*, vol. 289, no. 26, pp. 18022–18032, 2014

[177] P. Michaud, A. Da Costa, B. Courtois, and J. Courtois, Polysaccharide lyases: Recent developments as biotechnological tools, *Crit. Rev. Biotechnol.*, vol. 23, no. 4, pp. 233–266, 2003

[178] Y. Kuang *et al.*, Production of heparin oligosaccharides by fusion protein of MBP-heparinase I and the enzyme thermostability, *J. Mol. Catal. B Enzym.*, vol. 43, no. 1–4, pp. 90–95, 2006

[179] Y. J. Hyun, I. H. Jung, and D. H. Kim, Expression of heparinase I of *Bacteroides stercoris* HJ-15 and its degradation tendency toward heparin-like glycosaminoglycans, *Carbohydr. Res.*, vol. 359, pp. 37–43, 2012

[180] G. Venkataraman, Z. Shriver, R. Raman, and R. Sasisekharan, Sequencing complex polyscccharides, *Science*, vol. 286, pp. 537–542, 1999

[181] N. J. Patel, C. Sharon, S. Baranwal, R. S. Boothello, U. R. Desai, and B. B. Patel, Heparan sulfate hexasaccharide selectively inhibits cancer stem cells self-renewal by activating p38 MAP kinase, *Oncotarget*, vol. 7, no. 51, pp. 84608–84622, 2016

[182] N. V. Sankarayanarayanan *et al.*, A Hexasaccharide Containing Rare 2-O-Sulfate-Glucuronic Acid Residues Selectively Activates Heparin Cofactor II, *Angew. Chemie - Int. Ed.*, vol. 56, no. 9, pp. 2312–2317, 2017

[183] R. Copeland *et al.*, Inhibit the Entry of Herpes Simplex Virus-1, *Biochemistry*, vol. 47, no. 21, pp. 5774–5783, 2009

[184] W. L. Chuang, H. McAllister, and D. L. Rabenstein, Hexasaccharides from the histamine-modified depolymerization of porcine intestinal mucosal heparin, *Carbohydr. Res.*, vol. 337, no. 10, pp. 935–945, 2002

[185] T. L. Blundell, Problems and solutions in protein engineering - towards rational design, *Trends Biotechnol.*, vol. 12, no. 5, pp. 145–148, 1994

[186] M. M. Rahman, I. Sarker, A. H. Elmenoufy, H. S. Bahlol, and S. Zaman, Recent Advances in Structure-Based Protein Engineering, *Sch. J. Appl. Med. Sci.*, vol. 3, no.

September, pp. 1224–1237, 2015

[187] B. Rupp, *Biomolecular Crystallography: Principles, Practice, and Applications to Structural Biology*. New York, 2010

[188] N. Asherie, Protein crystallization and phase diagrams, *Methods*, vol. 34, no. 3, pp. 266–272, 2004

[189] T. Bergfors, Seeds to crystals, vol. 142, pp. 66–76, 2003

[190] D. Friedmann, T. Messick, and R. Marmorstein, Crystallization of Macromolecules, *Curr. Protoc. protein Sci.*, pp. 1–34, 2014

[191] J. W. Pflugrath, Practical macromolecular cryocrystallography, *Acta Crystallogr. Sect. F. Struct. Biol. Cryst. Commun.*, vol. 71, pp. 622–642, 2015

[192] W. H. Bragg and W. L. Bragg, The Reflection of X-rays by Crystals, *Proc. R. Soc. London A*, vol. 88, no. 605, pp. 428–438, 1913

[193] J. Drenth, X-ray Diffraction: Principles, *Encycl. Life Sci.*, pp. 1–8, 2002

[194] R. Gale, *Crystallography Made Crystal Clear*, 3rd ed. Academic Press, 2006

[195] P. H. Zwart, R. W. Grosse-Kunstleve, A. A. Lebedev, G. N. Murshudov, and P. D. Adams, Surprises and pitfalls arising from (pseudo)symmetry, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 64, no. 1, pp. 99–107, 2007

[196] M. W. Parker, Protein Structure from X-Ray Diffraction, *J. Biol. Phys.*, vol. 29, no. 4, pp. 341–362, 2003

[197] I. Usón and G. M. Sheldrick, Advances in direct methods for protein crystallography, *Curr. Opin. Struct. Biol.*, vol. 9, pp. 643–648, 1999

[198] P. Evans and A. McCoy, An introduction to molecular replacement, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 64, no. 1, pp. 1–10, 2007

[199] G. N. Murshudov, A. A. Vagin, and E. J. Dodson, Refinement of macromolecular structures by the maximum-likelihood method, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 53, no. 3, pp. 240–255, 1997

[200] A. T. Brünger, Free R value: a novel statistical quantity for assessing the accuracy of crystal structures, *Nature*, vol. 335, pp. 472–465, 1992

[201] C. Davies and C. Tournier, Exploring the function of the JNK (c-Jun N-terminal kinase) signalling pathway in physiological and pathological processes to design novel therapeutic strategies, *Biochem. Soc. Trans.*, vol. 40, pp. 85–89, 2012

[202] A. M. Manning and R. J. Davis, Targeting JNK for therapeutic benefit: from junk to gold?, *Nature*, vol. 2, pp. 554–565, 2003

[203] M. A. Siddiqui and P. A. Reddy, Small Molecule JNK (c-Jun N-Terminal Kinase) Inhibitors, *J. Med. Chem.*, vol. 53, pp. 3005–3012, 2010

[204] P. Cohen, The origins of protein phosphorylation, *Nat. Cell Biol.*, vol. 4, no. 5, 2002

[205] P. W. Schenk and B. E. Snaar-Jagalska, Signal perception and transduction: The role of protein kinases, *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1449, no. 1, pp. 1–24, 1999

[206] S. K. Grant, Therapeutic Protein Kinase Inhibitors, *Cell. Mol. Life Sci.*, vol. 66, no. 7, pp. 1163–1177, 2009

[207] P. Cohen, Protein kinases - the major drug target of twenty century, vol. 1, no. April, 2002

[208] H. Patterson, R. Nibbs, I. Mcinnes, and S. Siebert, Protein kinase inhibitors in the treatment of inflammatory and autoimmune diseases, *Clin. Exp. Immunol.*, vol. 176, no. 1, pp. 1–10, 2014

[209] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, and S. Sudarsanam, The Protein Kinase Complement of the Human Genome, *Science*, vol. 298, no. December, pp. 1912–1934, 2002

[210] M. E. M. Noble, J. A. Endicott, and L. N. Johnson, Protein Kinase Inhibitors : Insights into Drug Design from Structure, vol. 303, no. March, pp. 1800–1806, 2004

[211] B. Nolen, S. Taylor, and G. Ghosh, Regulation of protein kinases: Controlling activity through activation segment conformation, *Mol. Cell*, vol. 15, no. 5, pp. 661–675, 2004

[212] M. Cherry and D. H. Williams, Recent Kinase and Kinase Inhibitor X-ray Structures: Mechanisms of Inhibition and Selectivity Insights, *Curr. Med. Chem.*, vol. 11, no. 6, pp. 663–673, 2004

[213] M. Strnisková, M. Barancík, and T. Ravingerová, Mitogen-Activated Protein Kinases and Their Role in Regulation of Cellular Processes, *Gen. Physiol. Biophys.*, vol. 21, no. 3, pp. 231–255, 2002

[214] Z. Chen *et al.*, MAP kinases, *Chem. Rev.*, vol. 101, no. 8, pp. 2449–2476, 2001

[215] Y. Keshet and R. Seger, The MAP Kinase Signaling Cascades: A System of Hundreds of Components Regulates a Diverse Array of Physiological Functions, *Methods Mol. Biol.*, vol. 661, pp. 3–38, 2010

[216] D. K. Morrison, MAP Kinase Pathways, *Cold Spring Harb. Lab. Perspect Biol*, vol. 4, pp. 1–6, 2012

[217] S. Lee, T. Zhou, and E. J. Goldsmith, Crystallization of MAP kinases, vol. 40, pp. 224–233, 2006

[218] J. Cui, M. Zhang, Y. Q. Zhang, and Z. H. Xu, JNK pathway: Diseases and therapeutic potential, *Acta Pharmacol. Sin.*, vol. 28, no. 5, pp. 601–608, 2007

[219] G. Pearson *et al.*, Mitogen-Activated Protein (MAP) Kinase Pathways: Regulation and Physiological Functions, vol. 22, no. January, pp. 153–183, 2001

[220] L. Chang and M. Karin, Mammalian MAP kinase signalling cascades, *Nature*, vol. 410, no. 6824, pp. 37–40, 2001

[221] M. Meister, A. Tomasovic, A. Banning, and R. Tikkanen, Mitogen-activated protein (MAP) kinase scaffolding proteins: A recount, *Int. J. Mol. Sci.*, vol. 14, no. 3, pp. 4854–4884, 2013

[222] J.-Q. Engle, K. M.; Mei, T-S.; Wasa, M.; Yu, Protein Scaffolds in MAP Kinase Signalling, *Cell Signal.*, vol. 21, no. 4, pp. 462–469, 2009

[223] J. M. Kyriakis *et al.*, The stress-activated protein kinase subfamily of c-Jun kinases, *Nature*, vol. 369, pp. 156–160, 1994

[224] R. K. Barr and M. A. Bogoyevitch, The c-Jun N-terminal protein kinase family of mitogen-activated protein kinases (JNK MAPKs), vol. 33, pp. 1047–1063, 2001

[225] M. Hibi, A. Lin, T. Smeal, A. Minden, and M. Karin, Identification of an oncoprotein- and UV-responsive protein kinase that binds and potentiates the c-Jun activation domain, *Genes Dev.*, vol. 7, no. 11, pp. 2135–2148, 1993

[226] S. Gupta *et al.*, Selective interaction of JNK protein kinase isoforms with transcription factors, vol. 15, no. 11, pp. 2760–2770, 1996

[227] M. A. Bogoyevitch, The isoform-specific functions of the c-Jun N-terminal Kinases (JNKs): differences revealed by gene targeting, *BioEssays*, no. 28, pp. 923–934, 2006

[228] R. J. Davis, Signal Transduction by the JNK Group of MAP Kinases, vol. 103, pp. 239–252, 2000

[229] X. Wang, A. Destrument, and C. Tournier, Physiological roles of MKK4 and MKK7: Insights from animal models, *Biochim. Biophys. Acta - Mol. Cell Res.*, vol. 1773, no. 8, pp. 1349–1357, 2007

[230] D. Brancho *et al.*, Mechanism of p38 MAP kinase activation *in vivo*, *Genes Dev.*, vol. 17, no. 16, pp. 1969–1978, 2003

[231] D. K. Morrison and R. J. Davis, Regulation of MAP Kinase Signaling Modules by Scaffold Proteins in Mammals, *Annu. Rev. Cell Dev. Biol.*, vol. 19, no. 1, pp. 91–118, 2003

[232] M. A. Bogoyevitch and B. Kobe, Uses for JNK : the Many and Varied Substrates of the c-Jun N-Terminal Kinases, vol. 70, no. 4, pp. 1061–1095, 2006

[233] J. Hess, AP-1 subunits: quarrel and harmony among siblings, *J. Cell Sci.*, vol. 117, no. 25, pp. 5965–5973, 2004

[234] R. Yarza, S. Vela, M. Solas, and M. J. Ramirez, c-Jun N-terminal kinase (JNK) signaling as a therapeutic target for Alzheimer's disease, *Front. Pharmacol.*, vol. 6, pp. 1–12, 2016

[235] J. Liu and A. Lin, Role of JNK activation in apoptosis: A double-edged sword, *Cell Res.*, vol. 15, no. 1, pp. 36–42, 2005

[236] L. Resnick, M. Fennell, and L. Resnick, Targeting JNK3 for the treatment of neurodegenerative disorders, vol. 9, no. 21, 2004

[237] P. P. Graczyk, JNK inhibitors as anti-inflammatory and neuroprotective agents, *Future Med. Chem.*, vol. 5, no. 5, pp. 539–551, 2013

[238] C. R. Weston and R. J. Davis, The JNK signal transduction pathway, vol. 1, pp. 142–149, 2007

[239] X. Xie *et al.*, Crystal structure of JNK3 : a kinase implicated in neuronal apoptosis, *Structure*, vol. 6, no. 8, pp. 983–991, 1998

[240] S. K. Hanks, A. M. Quinn, and T. Hunter, The protein kinase family: Conserved features and deduced phylogeny of the catalytic domains, *Science*, vol. 241, no. 4861, pp. 42–52, 1988

[241] D. Fabbro, S. W. Cowan-Jacob, and H. Moebitz, Ten things you should know about protein kinases: IUPHAR Review 14, *Br. J. Pharmacol.*, vol. 172, no. 11, pp. 2675–2700, 2015

[242] M. Huse and J. Kuriyan, The conformational plasticity of protein kinases, *Cell*, vol.

109, no. 3, pp. 275–282, 2002

[243] D. Bossemeyer, Protein kinases - structure and function, *FEBS Lett.*, vol. 369, pp. 57–61, 1995

[244] P. Mishra and S. Günther, New insights into the structural dynamics of the kinase JNK3, *Sci. Rep.*, vol. 8, no. 1, pp. 1–13, 2018

[245] J. A. Adams, Kinetic and catalytic mechanisms of protein kinases, *Chem. Rev.*, vol. 101, no. 8, pp. 2271–2290, 2001

[246] L. N. Johnson, M. E. M. Noble, and D. J. Owen, Active and inactive protein kinases: Structural basis for regulation, *Cell*, vol. 85, no. 2, pp. 149–158, 1996

[247] A. Messoussi, G. Chevé, K. Bougrin, and A. Yasri, Insight into the selective inhibition of JNK family members through structure-based drug design, *RSC*, vol. 7, no. 4, pp. 686–692, 2016

[248] R. A. Norman, D. Toader, and A. D. Ferguson, Structural approaches to obtain kinase selectivity, *Trends Pharmacol. Sci.*, vol. 33, no. 5, pp. 273–278, 2012

[249] Y. Liu and N. S. Gray, Rational design of inhibitors that bind to inactive kinase conformations, *Nat. Chem. Biol.*, vol. 2, no. 7, pp. 358–364, 2006

[250] P. Traxler and P. Furet, Strategies toward the design of novel and selective protein tyrosine kinase inhibitors, *Pharmacol. Ther.*, vol. 82, no. 2–3, pp. 195–206, 1999

[251] A. Deroide, A. Messoussi, N. Van Hijfte, K. Bougrin, and A. Yasri, Review Recent Progress in the Design, Study, and Development of c-Jun N-Terminal Kinase Inhibitors as Anticancer Agents, pp. 1433–1443, 2014

[252] B. L. Bennett *et al.*, SP600125, an anthrapyrazolone inhibitor of Jun N-terminal kinase, *Proc. Natl. Acad. Sci.*, vol. 98, no. 24, pp. 13681–13686, 2001

[253] S. Tanemura *et al.*, Blockage by SP600125 of Fcε receptor-induced degranulation and cytokine gene expression in mast cells is mediated through inhibition of phosphatidylinositol 3-kinase signalling pathway, *J. Biochem.*, vol. 145, no. 3, pp. 345–354, 2009

[254] M. A. Bogoyevitch and P. G. Arthur, Inhibitors of c-Jun N-terminal kinases - JuNK no more ?, vol. 1784, pp. 76–93, 2008

[255] Z. A. Knight, K. M. Shokat, S. Francisco, and S. Francisco, Features of Selective Kinase Inhibitors Review, *Chem. Biol.*, vol. 12, pp. 621–637, 2005

[256] L. A. Smyth and I. Collins, Measuring and interpreting the selectivity of protein kinase inhibitors, *J. Chem. Biol.*, vol. 2, no. 3, pp. 131–151, 2009

[257] E. K. Kim and E. J. Choi, Pathological roles of MAPK signaling pathways in human diseases, *Biochim. Biophys. Acta - Mol. Basis Dis.*, vol. 1802, no. 4, pp. 396–405, 2010

[258] K. Zheng, S. Iqbal, P. Hernandez, H. Park, P. V. Lograsso, and Y. Feng, Design and synthesis of highly potent and isoform selective JNK3 inhibitors: SAR studies on aminopyrazole derivatives, *J. Med. Chem.*, vol. 57, no. 23, pp. 10013–10030, 2014

[259] E. Gasteiger, A. Gattiker, C. Hoogland, I. Ivanyi, R. D. Appel, and A. Bairoch, ExPASy: The proteomics server for in-depth protein knowledge and analysis, *Nucleic Acids*

Res., vol. 31, no. 13, pp. 3784–3788, 2003

[260] R. W. Woody, Circular Dichroism, *Methods Enzymol.*, vol. 246, pp. 34–71, 1995

[261] N. J. Greenfield, Using circular dichroism spectra to estimate protein secondary structure, *Nat. Protoc.*, vol. 1, no. 6, pp. 2876–2890, 2007

[262] S. M. Kelly, T. J. Jess, and N. C. Price, How to study proteins by circular dichroism, *Biochim. Biophys. Acta - Proteins Proteomics*, vol. 1751, no. 2, pp. 119–139, 2005

[263] A. Micsonai *et al.*, Accurate secondary structure prediction and fold recognition for circular dichroism spectroscopy, *Proc. Natl. Acad. Sci.*, vol. 112, no. 24, pp. E3095–E3103, 2015

[264] F. H. Niesen, H. Berglund, and M. Vedadi, The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability, *Nat. Protoc.*, vol. 2, no. 9, pp. 2212–2221, 2007

[265] A. Simeonov, Recent Developments in the Use of Differential Scanning Fluorometry in Protein and Small Molecule Discovery and Characterization, *Expert Opin. Drug Discov.*, vol. 8, no. 9, pp. 1071–1082, 2013

[266] W. Strober, Trypan Blue Exclusion Test of Cell Viability, *Curr. Protoc. Immunol.*, vol. 111, pp. 1–3, 2015

[267] T. Duellman, J. Burnett, J. Yang, C. Science, and T. Program, Quantitation of secreted proteins using mCherry fusion constructs and a fluorescent microplate reader, *Anal. Biochem.*, vol. 473, pp. 34–40, 2015

[268] P. Sánchez-Corral, C. González-Rubio, S. Rodríguez De Córdoba, and M. López-Trascasa, Functional analysis in serum from atypical Hemolytic Uremic Syndrome patients reveals impaired protection of host cells associated with mutations in factor H, *Mol. Immunol.*, vol. 41, no. 1, pp. 81–84, 2004

[269] R. P. Venkirakrishnan, O. Bernand, M. Max, and J. L. Markley, Use of NMR Saturation Transfer Difference Spectroscopy to Study Ligand Binding to Membrane Proteins, *Methods Mol. Biol.*, vol. 914, pp. 47–63, 2012

[270] M. Mayer and B. Meyer, Characterization of ligand binding by saturation transfer difference NMR spectroscopy, *Angew. Chemie - Int. Ed.*, vol. 38, no. 12, pp. 1784–1788, 1999

[271] A. Viegas, J. Manso, F. L. Nobrega, and E. J. Cabrita, Saturation-transfer difference (STD) NMR: A simple and fast method for ligand screening and characterization of protein binding, *J. Chem. Educ.*, vol. 88, no. 7, pp. 990–994, 2011

[272] A. Lange *et al.*, Targeting the Gatekeeper MET146 of C-Jun N-Terminal Kinase 3 Induces a Bivalent Halogen/Chalcogen Bond, *J. Am. Chem. Soc.*, vol. 137, no. 46, pp. 14640–14652, 2015

[273] W. Kabsch, XDS, *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 66, no. Pt 2, pp. 125–132, 2010

[274] A. Vagin and A. Teplyakov, MOLREP: An Automated Program for Molecular Replacement, *J. Appl. Crystallogr.*, vol. 30, no. 6, pp. 1022–1025, 1997

[275] P. Emsley and K. Cowtan, Coot: Model-building tools for molecular graphics, *Acta*

Crystallogr. Sect. D Biol. Crystallogr., vol. 60, no. 12 I, pp. 2126–2132, 2004

[276] G. N. Murshudov *et al.*, REFMAC5 for the refinement of macromolecular crystal structures, *Acta Crystallogr. Sect. D Biol. Crystallogr.*, vol. 67, no. 4, pp. 355–367, 2011

[277] P. D. Adams *et al.*, PHENIX: a comprehensive Python-based system for macromolecular structure solution, *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 66, pp. 213–221, 2010

[278] V. B. Chen *et al.*, MolProbity: all-atom structure validation for macromolecular crystallography., *Acta Crystallogr. D. Biol. Crystallogr.*, vol. 66, no. Pt 1, pp. 12–21, Jan. 2010

[279] K. F. Nolan, S. Kaluz, J. M. G. Higgins, D. Goundis, and K. B. M. Reid, Characterization of the human properdin gene, *Biochem. J.*, vol. 287, no. 1, pp. 291–297, 1992

[280] S. Hartmann and J. Hofsteenge, Properdin, the positive regulator of complement, is highly C-mannosylated, *J. Biol. Chem.*, vol. 275, no. 37, pp. 28569–28574, 2000

[281] H. H. Freeze and C. Kranz, Endoglycosidase and glycoamidase release of N-linked glycans, *Curr. Protoc. Mol. Biol.*, pp. 1–33, 2010

[282] R. Pain, Determining the CD spectrum of a protein, *Handb. Food Anal. Chem.*, vol. 1–2, pp. 219–243, 2005

[283] D. T. Fearon, Regulation by membrane sialic acid of beta1H-dependent decay - dissociation of amplification C3 convertase of the alternative complement pathway, *Proc. Natl. Acad. Sci. U. S. A.*, vol. 75, no. 4, pp. 1971–1975, 1978

[284] P. a Klenotic, R. C. Page, S. Misra, and R. L. Silverstein, Expression, purification and structural characterization of functionally replete thrombospondin-1 type 1 repeats in a bacterial expression system, *Protein Expr. Purif.*, vol. 80, no. 2, pp. 253–9, Dec. 2011

[285] J. Lobstein, C. A. Emrich, C. Jeans, M. Faulkner, P. Riggs, and M. Berkmen, SHuffle, a novel *Escherichia coli* protein expression strain capable of correctly folding disulfide bonded proteins in its cytoplasm, *Microb. Cell Fact.*, vol. 11, no. 56, pp. 1–16, 2012

[286] N. Portolano *et al.*, Recombinant protein expression for structural biology in HEK 293F suspension cells: a novel and accessible approach., *J. Vis. Exp.*, vol. 92, pp. 1–8, 2014

[287] C. Liu, B. Dalby, W. Chen, J. M. Kilzer, and H. C. Chiou, Transient transfection factors for high-level recombinant protein production in suspension cultured mammalian cells, *Mol. Biotechnol.*, vol. 39, no. 2, pp. 141–153, 2008

[288] S. Meri, Self-nonself discrimination by the complement system, *FEBS Lett.*, vol. 590, pp. 2418–2434, 2016

[289] A. Zaferani *et al.*, Factor H and properdin recognize different epitopes on renal tubular epithelial heparan sulfate, *J. Biol. Chem.*, vol. 287, no. 37, pp. 31471–31481, 2012

[290] S. Walpole, S. Monaco, R. Nepravishta, and J. Angulo, *STD NMR as a Technique for Ligand Screening and Structural Studies*, 1st ed., vol. 615. Elsevier Inc., 2019

[291] R. Pogan, P. J. Kehr, and C. Uetrecht, Establishing native mass spectrometry for

studying virus particle glycan interaction, Heinrich Pete Institute, University of Hamburg, 2017

[292] A. J. R. Heck, Native mass spectrometry: A bridge between interactomics and structural biology, *Nat. Methods*, vol. 5, no. 11, pp. 927–933, 2008

[293] M. Lebendiker and T. Danieli, Production of prone-to-aggregate proteins, *FEBS Lett.*, vol. 588, no. 2, pp. 236–246, 2014

[294] X. Ma, Z. Wang, S. Li, Q. Shen, and Q. Yuan, Effect of CaCl2 as activity stabilizer on purification of heparinase I from *Flavobacterium heparinum*, *J. Chromatogr. B*, vol. 843, no. 2, pp. 209–215, 2006

[295] Y. Luo, X. Huang, and W. L. McKeehan, High yield, purity and activity of soluble recombinant Bacteroides thetaiotaomicron GST-heparinase I from *Escherichia coli*, *Arch. Biochem. Biophys.*, vol. 460, no. 1, pp. 17–24, 2007

[296] F. Sievers *et al.*, Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega, *Mol. Syst. Biol.*, vol. 7, no. 539, 2011

[297] X. Robert and P. Gouet, Deciphering key features in protein structures with the new ENDscript server, *Nucleic Acids Res.*, vol. 42, no. W1, pp. 320–324, 2014

[298] S. Ernst *et al.*, Expression in Escherichia coli, purification and charactezation of heparinase I from *Flavobacterium heparinum*, *Biochem. J.*, vol. 315, pp. 589–597, 1996

[299] A. Waterhouse *et al.*, SWISS-MODEL: Homology modelling of protein structures and complexes, *Nucleic Acids Res.*, vol. 46, no. W1, pp. W296–W303, 2018

[300] T. H. Plummer, A. L. Tarentino, and C. R. Hauer, Novel, specific O-glycosylation of secreted *Flavobacterium meningosepticum* proteins., *Journal of Biological Chemistry*, vol. 270, no. 22. pp. 13192–13196, 2002

[301] S. Chen *et al.*, Biochemical analysis and kinetic modeling of the thermal inactivation of MBP-fused heparinase I: Implications for a comprehensive thermostabilization strategy, *Biotechnol. Bioeng.*, vol. 108, no. 8, pp. 1841–1851, 2011

[302] A. Dror, E. Shemesh, N. Dayan, and A. Fishman, Protein engineering by random mutagenesis and structure-guided consensus of *Geobacillus stearothermophilus* lipase T6 for enhanced stability in methanol, *Appl. Environ. Microbiol.*, vol. 80, no. 4, pp. 1515–1527, 2014

[303] F. Ansideri, A. Lange, A. El-Gokha, F. M. Boeckler, and P. Koch, Fluorescence polarization-based assays for detecting compounds binding to inactive c-Jun N-terminal kinase 3 and p38α mitogen-activated protein kinase, *Anal. Biochem.*, vol. 503, pp. 28–40, 2016

[304] F. Ansideri, M. Dammann, F. M. Boeckler, and P. Koch, Fluorescence polarization-based competition binding assay for c-Jun N-terminal kinases 1 and 2, *Anal. Biochem.*, vol. 532, pp. 26–28, 2017

[305] M. Goettert, S. Luik, R. Graeser, and S. A. Laufer, A direct ELISA assay for quantitative determination of the inhibitory potency of small molecules inhibitors for JNK3, *J. Pharm. Biomed. Anal.*, vol. 55, no. 1, pp. 236–240, 2011

[306] E. Krissinel and K. Henrick, Inference of Macromolecular Assemblies from Crystalline State, *J. Mol. Biol.*, vol. 372, no. 3, pp. 774–797, 2007

[307] S. A. Laufer, D. R. J. Hauser, D. M. Domeyer, K. Kinkel, and A. J. Liedtke, Design, synthesis, and biological evaluation of novel tri- and tetrasubstituted imidazoles as highly potent and specific ATP-mimetic inhibitors of p38 MAP kinase: Focus on optimized interactions with the enzyme's surface-exposed front region, *J. Med. Chem.*, vol. 51, no. 14, pp. 4122–4149, 2008

[308] J. Wityak *et al.*, Lead Optimization toward Proof-of-Concept Tools for Huntington's Disease within a 4-(1 H -Pyrazol-4-yl)pyrimidine Class of Pan-JNK Inhibitors, *J. Med. Chem.*, vol. 58, no. 7, pp. 2967–2987, 2015

[309] F. Ansideri, Multiple strategies targeting c-Jun N-terminal kinases : synthesis of novel inhibitors and development of a new binding assay methodology, Eberhard Karls Universität Tübingen, 2018

[310] T. Sterling and J. J. Irwin, ZINC 15 - Ligand Discovery for Everyone, *J. Chem. Inf. Model.*, vol. 55, no. 11, pp. 2324–2337, 2015

[311] G. Scapin, S. B. Patel, J. Lisnock, J. W. Becker, and P. V Lograsso, The Structure of JNK3 in Complex with Small Molecule Inhibitors : Structural Basis for Potency and Selectivity 3535 General Atomics Court, vol. 10, pp. 705–712, 2003

[312] T. Kamenecka *et al.*, Structure-Activity Relationships and X-ray Structures Describing the Selectivity of Aminopyrazole Inhibitors for c-Jun N-terminal Kinase 3 (JNK3) over p38, *J. Biol. Chem.*, vol. 284, no. 19, pp. 12853–12861, 2009

[313] G. D. Probst *et al.*, Highly selective c-Jun N-terminal kinase (JNK) 3 inhibitors with *in vitro* CNS-like pharmacokinetic properties II. Central core replacement, *Bioorganic Med. Chem. Lett.*, vol. 21, no. 12, pp. 3726–3729, 2011

[314] B. M. Swahn *et al.*, Design and synthesis of 2′-anilino-4,4′-bipyridines as selective inhibitors of c-Jun N-terminal kinase-3, *Bioorganic Med. Chem. Lett.*, vol. 16, no. 5, pp. 1397–1401, 2006

[315] T. Kamenecka *et al.*, Synthesis, biological evaluation, X-ray structure, and pharmacokinetics of aminopyrimidine c-jun-N-terminal kinase (JNK) inhibitors, *J. Med. Chem.*, vol. 53, no. 1, pp. 419–431, 2010

[316] V. P. Krenitsky *et al.*, Bioorganic & Medicinal Chemistry Letters Discovery of CC-930 , an orally active anti-fibrotic JNK inhibitor, *Bioorg. Med. Chem. Lett.*, vol. 22, pp. 1433–1438, 2012

[317] Z. Wang *et al.*, Structural basis of inhibitor selectivity in MAP kinases., *Structure*, vol. 6, no. 9, pp. 1117–1128, 1998

[318] J. Hynes and K. Leftheri, Small molecule p38 inhibitors: novel structural features and advances from 2002-2005, *Curr. Top. Med. Chem.*, vol. 5, no. 10, pp. 967–85, 2005

[319] H. McWilliam *et al.*, Analysis Tool Web Services from the EMBL-EBI., *Nucleic Acids Res.*, vol. 41, no. Web Server issue, pp. 597–600, 2013

9 Appendix

9.1 Amino acid sequences

The native signal peptide of the corresponding protein is colored in blue and tags, cleavage sites as well as cloning artifacts are displayed in gray in the respective construct.

Properdin_mCherry

```
        10         20         30         40         50         60
MITEGAQAPR LLLPPLLLLL TLPATGSDPV LCFTQYEESS GKCKGLLGGG VSVEDCCLNT

        70         80         90        100        110        120
AFAYQKRSGG LCQPCRSPRW SLWSTWAPCS VTCSEGSQLR YRRCVGWNGQ CSGKVAPGTL

       130        140        150        160        170        180
EWQLQACEDQ QCCPEMGGWS GWGPWEPCSV TCSKGTRTRR RACNHPAPKC GGHCPGQAQE

       190        200        210        220        230        240
SEACDTQQVC PTHGAWATWG PWTPCSASCH GGPHEPKETR SRKCSAPEPS QKPPGKPCPG

       250        260        270        280        290        300
LAYEQRRCTG LPPCPVAGGW GPWGPVSPCP VTCGLGQTME QRTCNHPVPQ HGGPFCAGDA

       310        320        330        340        350        360
TRTHICNTAV PCPVDGEWDS WGEWSPCIRR NMKSISCQEI PGQQSRGRTC RGRKFDGHRC

       370        380        390        400        410        420
AGQQQDIRHC YSIQHCPLKG SWSEWSTWGL CMPPCGPNPT RARQRLCTPL LPKYPPTVSM

       430        440        450        460        470        480
VEGQGEKNVT FWGRPLPRCE ELQGQKLVVE EKRPCLHVPA CKDPEEEELE NLYFQGGGSG

       490        500        510        520        530        540
GGSGGVSKGE EDNMAIIKEF MRFKVHMEGS VNGHEFEIEG EGEGRPYEGT QTAKLKVTKG

       550        560        570        580        590        600
GPLPFAWDIL SPQFMYGSKA YVKHPADIPD YLKLSFPEGF KWERVMNFED GGVVTVTQDS

       610        620        630        640        650        660
SLQDGEFIYK VKLRGTNFPS DGPVMQKKTM GWEASSERMY PEDGALKGEI KQRLKLKDGG

       670        680        690        700        710        720
HYDAEVKTTY KAKKPVQLPG AYNVNIKLDI TSHNEDYTIV EQYERAEGRH STGGMDELYK

       730
GGSGHHHHHH HHHH
```

Properdin_c-myc

```
        10         20         30         40         50         60
MITEGAQAPR LLLPPLLLLL TLPATGSDPV LCFTQYEESS GKCKGLLGGG VSVEDCCLNT

        70         80         90        100        110        120
AFAYQKRSGG LCQPCRSPRW SLWSTWAPCS VTCSEGSQLR YRRCVGWNGQ CSGKVAPGTL
```

```
        130        140        150        160        170        180
EWQLQACEDQ QCCPEMGGWS GWGPWEPCSV TCSKGTRTRR RACNHPAPKC GGHCPGQAQE

        190        200        210        220        230        240
SEACDTQQVC PTHGAWATWG PWTPCSASCH GGPHEPKETR SRKCSAPEPS QKPPGKPCPG

        250        260        270        280        290        300
LAYEQRRCTG LPPCPVAGGW GPWGPVSPCP VTCGLGQTME QRTCNHPVPQ HGGPFCAGDA

        310        320        330        340        350        360
TRTHICNTAV PCPVDGEWDS WGEWSPCIRR NMKSISCQEI PGQQSRGRTC RGRKFDGHRC

        370        380        390        400        410        420
AGQQQDIRHC YSIQHCPLKG SWSEWSTWGL CMPPCGPNPT RARQRLCTPL LPKYPPTVSM

        430        440        450        460        470        480
VEGQGEKNVT FWGRPLPRCE ELQGQKLVVE EKRPCLHVPA CKDPEEEELE NLYFQGSSGE

        490        500
QKLISEEDLN SAVDHHHHHH HHHH
```

HepIv1

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLVPRGSH MQQKKSGNIP YRVNVQADSA KQKAIIDNKW VAVGINKPYA

        70         80         90         100        110        120
LQYDDKLRFN GKPSYRFELK AEDNSLEGYA AGETKGRTEL SYSYATTNDF KKFPPSVYQN

        130        140        150        160        170        180
AQKLKTVYHY GKGICEQGSS RSYTFSVYIP SSFPDNATTI FAQWHGAPSR TLVATPEGEI

        190        200        210        220        230        240
KTLSIEEFLA LYDRMIFKKN IAHDKVEKKD KDGKITYVAG KPNGWKVEQG GYPTLAFGFS

        250        260        270        280        290        300
KGYFYIKANS DRQWLTDKAD RNNANPENSE VMKPYSSEYK TSTIAYKMPF AQFPKDCWIT

        310        320        330        340        350        360
FDVAIDWTKY GKEANTILKP GKLDVMMTYT KNKKPQKAHI VNQQEILIGR NDDDGYYFKF

        370        380
GIYRVGNSTV PVTYNLSGYS ETAR
```

HepIv2

```
        10         20         30         40         50         60
MGSSHHHHHH SSGLEVLFQG PQQLFLCSAY AQQKKSGNIP YRVNVQADSA KQKAIIDNKW

        70         80         90         100        110        120
VAVGINKPYA LQYDDKLRFN GKPSYRFELK AEDNSLEGYA AGETKGRTEL SYSYATTNDF

        130        140        150        160        170        180
KKFPPSVYQN AQKLKTVYHY GKGICEQGSS RSYTFSVYIP SSFPDNATTI FAQWHGAPSR

        190        200        210        220        230        240
TLVATPEGEI KTLSIEEFLA LYDRMIFKKN IAHDKVEKKD KDGKITYVAG KPNGWKVEQG
```

```
        250        260        270        280        290        300
GYPTLAFGFS KGYFYIKANS DRQWLTDKAD RNNANPENSE VMKPYSSEYK TSTIAYKMPF

        310        320        330        340        350        360
AQFPKDCWIT FDVAIDWTKY GKEANTILKP GKLDVMMTYT KNKKPQKAHI VNQQEILIGR

        370        380        390
NDDDGYYFKF GIYRVGNSTV PVTYNLSGYS ETAR
```

Heplv3

```
        10         20         30         40         50         60
MQQKKSGNIP YRVNVQADSA KQKAIIDNKW VAVGINKPYA LQYDDKLRFN GKPSYRFELK

        70         80         90         100        110        120
AEDNSLEGYA AGETKGRTEL SYSYATTNDF KKFPPSVYQN AQKLKTVYHY GKGICEQGSS

        130        140        150        160        170        180
RSYTFSVYIP SSFPDNATTI FAQWHGAPSR TLVATPEGEI KTLSIEEFLA LYDRMIFKKN

        190        200        210        220        230        240
IAHDKVEKKD KDGKITYVAG KPNGWKVEQG GYPTLAFGFS KGYFYIKANS DRQWLTDKAD

        250        260        270        280        290        300
RNNANPENSE VMKPYSSEYK TSTIAYKMPF AQFPKDCWIT FDVAIDWTKY GKEANTILKP

        310        320        330        340        350        360
GKLDVMMTYT KNKKPQKAHI VNQQEILIGR NDDDGYYFKF GIYRVGNSTV PVTYNLSGYS

        370
ETARLEHHHH HH
```

JNK3

```
        10         20         30         40         50         60
GGSMSKSKVD NQFYSVEVGD STFTVLKRYQ NLKPIGSGAQ GIVCAAYDAV LDRNVAIKKL

        70         80         90         100        110        120
SRPFQNQTHA KRAYRELVLM KCVNHKNIIS LLNVFTPQKT LEEFQDVYLV MELMDANLCQ

        130        140        150        160        170        180
VIQMELDHER MSYLLYQMLC GIKHLHSAGI IHRDLKPSNI VVKSDCTLKI LDFGLARTAG

        190        200        210        220        230        240
TSFMMTPYVV TRYYRAPEVI LGMGYKENVD IWSVGCIMGE MVRHKILFPG RDYIDQWNKV

        250        260        270        280        290        300
IEQLGTPCPE FMKKLQPTVR NYVENRPKYA GLTFPKLFPD SLFPADSEHN KLKASQARDL

        310        320        330        340        350        360
LSKMLVIDPA KRISVDDALQ HPYINVWYDP AEVEAPPPQI YDKQLDEREH TIEEWKELIY

KEVMNSE
```

9.2 Molecular weight and isoelectric point

Table 9.1 Theoretical MW and pI of the different recombinantly proteins presented in this work. Theoretical values were calculated based on the protein sequence using the ExPASy ProtParam computational tool [259]. Parameters are indicated for the his-tagged protein (after secretion in the case of properdin) and the protein after tag removal (cleaved protein).

Construct	Tagged protein		Cleaved protein	
	MW (kDa)	Theoretical pI	MW (kDa)	Theoretical pI
Properdin_mCherry	78.1	7.63	49.29	8.26
Properdin_c-myc	52.6	7.91	49.29	8.26
HepIv1	43.63	9.36	41.75	9.32
Hep1v2	44.76	9.23	42.62	9.28
HepIv3	42.53	9.27	-	-
JNK3	-	-	42.21	7.52

9.3 Licenses for figures from other publications

Some figures in this dissertation were taken from original publications. All original publications are referenced in the respective figure subtext and are listed here.

Figure 1.3

M. Alcorlo, A. Tortajada, S. Rodriguez de Cordoba, and O. Llorca, "Structural basis for the stabilization of the complement alternative pathway C3 convertase by properdin," *Proc. Natl. Acad. Sci.*, vol. 110, no. 33, pp. 13504–13509, 2013.

Published by PNAS. Permission was granted for use of the material in this dissertation as requested. Because this material was published after 2008, a copyright note is not needed.

Figure 1.6A

Y. H. Han, M. L. Garron, H. Y. Kim, W. S. Kim, Z. Zhang, K. S. Ryu, D. Shaya, Z. Xiao, C. Cheong, Y. S. Kim, R. J. Linhardt, Y. H. Jeon, and M. Cygler, "Structural snapshots of heparin depolymerization by heparin lyase I," *J. Biol. Chem.*, vol. 284, no. 49, pp. 34019-34027, 2009.

Provided by Copyright Clearance Center
Order license ID: 4602111261013
Author/Edithor: American Society for Biochemistry and Molecular Biology
Confirmation number: 11820834
Order date: 04.06.2019

Figure 1.6B

J. F. K. Limtiaco, S. Beni, C. J. Jones, D. J. Langeslay, and C. K. Larive, "NMR methods to monitor the enzymatic depolymerization of heparin," *Anal. Bioanal. Chem.*, vol. 399, no. 2, pp. 593–603, 2011.

Published by Springer Nature. This article is distributed under the terms of the Creative Commons Attribution Non-commercial License which permits any commercial use, distribution, and reproduction in any medium, provided the original author(s) and source are credited. As long as the dissertation is not published commercially, the material can be reused without obtaining permission.

10 Publications and contributions

F. Ansideri, J. T. Macedo, M. Eitel, A. El-Gokha, D. S. Zinad, C. Scarpellini, M. Kudolo, D. Schollmeyer, F. M. Boeckler, B. S. Blaum, S. A. Laufer, and P. Koch, "Structural Optimization of a Pyridinylimidazole Scaffold: Shifting the Selectivity from p38α Mitogen-Activated Protein Kinase to c-Jun N-Terminal Kinase 3," *ACS Omega*, vol. 3, no. 7, pp. 7809–7831, 2018.

Contributions

I have assessed the structural integrity of the JNK3 protein, performed the crystallization and incubation experiments with synthesized inhibitors, tested and improved the crystals, and collected and interpreted diffraction data. Besides, I have also planned and conducted the thermal unfolding experiments using nanoDSF and analyzed the obtained data. Figures concerning the crystal structures and unfolding thermal curves were prepared by me. I contributed to the manuscript writing, especially the sections corresponding to the crystal structures and nanoDSF results.

ACS OMEGA

Article

Structural Optimization of a Pyridinylimidazole Scaffold: Shifting the Selectivity from p38α Mitogen-Activated Protein Kinase to c-Jun N-Terminal Kinase 3

Francesco Ansideri,[†] Joana T. Macedo,[‡] Michael Eitel,[†] Ahmed El-Gokha,[†,§] Dhafer S. Zinad,[†,∥] Camilla Scarpellini,[†] Mark Kudolo,[†] Dieter Schollmeyer,[⊥] Frank M. Boeckler,[†] Bärbel S. Blaum,[‡] Stefan A. Laufer,[†] and Pierre Koch*,[†]

[†]Department of Pharmaceutical and Medicinal Chemistry, Institute of Pharmaceutical Sciences, Eberhard Karls Universität Tübingen, Auf der Morgenstelle 8, 72076 Tübingen, Germany

[‡]Interfaculty Institute of Biochemistry, Eberhard Karls Universität Tübingen, Hoppe-Seyler-Straße 4, 72076 Tübingen, Germany

[⊥]Department of Organic Chemistry, Johannes Gutenberg University Mainz, Duesbergweg 10-14, D-55099 Mainz, Germany

Ⓢ Supporting Information

ABSTRACT: Starting from known p38α mitogen-activated protein kinase (MAPK) inhibitors, a series of inhibitors of the c-Jun N-terminal kinase (JNK) 3 was obtained. Altering the substitution pattern of the pyridinylimidazole scaffold proved to be effective in shifting the inhibitory activity from the original target p38α MAPK to the closely related JNK3. In particular, a significant improvement for JNK3 selectivity could be achieved by addressing the hydrophobic region I with a small methyl group. Furthermore, additional structural modifications permitted to explore structure−activity relationships. The most potent inhibitor 4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine showed an IC_{50} value for the JNK3 in the low triple digit nanomolar range and its binding mode was confirmed by X-ray crystallography.
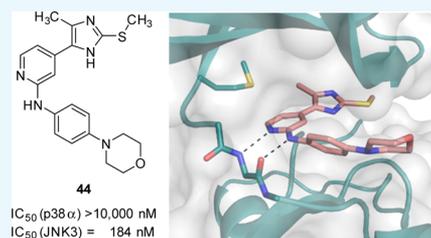
44
IC_{50} (p38α) >10,000 nM
IC_{50} (JNK3) =   184 nM

■ INTRODUCTION

The mitogen-activated protein kinases (MAPKs) represent a family of enzymes involved in several signal transduction pathways, whose activation is part of a phosphorylation cascade triggered by diverse extracellular stimuli. Among the members of this family, the c-Jun N-terminal kinases (JNKs) mostly respond to a variety of stress stimuli such as radiation, osmotic or heat shock, oxidative insult, and proinflammatory cytokines, modulating responses such as cell survival and apoptosis.[1] The JNK subfamily is encoded by the three genes *jnk1*, *jnk2*, and *jnk3*, which in turn give rise to 10 different isoforms through alternative splicing.[2] Despite their structural homology and the partially functional redundancy, these isoforms follow a different tissue distribution pattern, JNK3 being restricted to the central nervous system, heart, and testis oppositely to the ubiquitous expression of JNK1 and 2.[2,3] In addition to this, a different substrate specificity of the JNK1, 2, and 3 suggests the existence of isoform-specific roles of these enzymes, which were partially disclosed through gene knock-out studies.[4] There is well-documented evidence for the critical role of the JNK subfamily members in several neuro-degenerative diseases such as Parkinson's and Alzheimer's disease, as well as in neuronal death derived by stroke and ischemia/reperfusion injury.[3−6] Furthermore, some members of the JNKs are also involved in metabolic and inflammatory diseases, and several studies suggest that these kinases might

contribute to the development and diffusion of some forms of cancer,[7−9] thus emerging as particularly attractive drug targets. Despite the intense endeavor in the research of JNK inhibitors, only a scarce number of candidates have reached clinical trial phases and to date, none of them have been approved.[10−12] Until early 2010s, a major challenge in the development of JNK inhibitors has been the achievement of selectivity over the closely related p38α MAPK,[11] a member of the same family which, analogously to the JNKs, participates in regulating the cellular response to stress stimuli. This protein kinase was also shown to assume a key function in different inflammatory and neurodegenerative diseases[13−15] and the simultaneous inhib-ition of JNK and p38α MAPK is assumed to obtain a synergistic effect in the treatment of some pathological conditions.[16] Nevertheless, obtaining a JNK-selective inhibitor would be beneficial to fully elucidate the effective role of this protein kinase in the aforementioned pathological conditions and thereby assess its therapeutic potential. Furthermore, most of the reported clinical trials on selective p38α MAPK inhibitors have been discontinued because of the insurgence of adverse effects mostly related to liver toxicity,[17] leading to
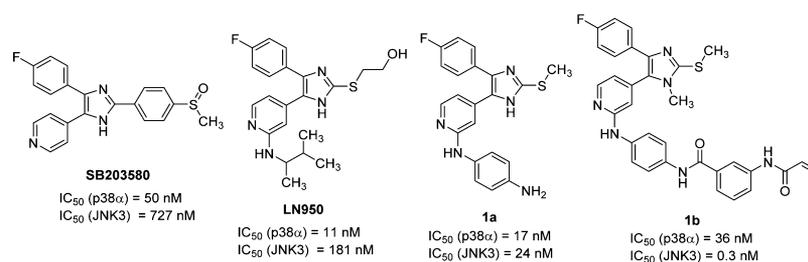
Figure 1. Tri- and tetrasubstituted pyridinylimidazoles. Data are taken from Ansideri et al.[19] and Muth et al.[21]

Scheme 1. Synthesis of Imidazole 5 and Imidazol-2-one 8[a]
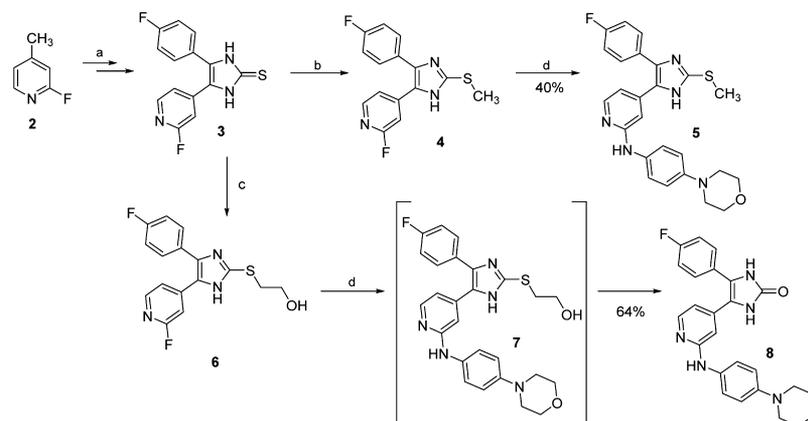
[a]Reagents and conditions: (a) four-step route reported by Laufer and co-workers;[23] (b) MeI, K_2CO_3, MeOH, rt, 18 h; (c) 2-bromoethyl acetate, t-BuONa, MeOH, 55 °C, 3 h; and (d) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h.

the assumption the activity on the p38α MAPK to be undesired for an improved safety profile of JNK inhibitors.

Regarding the selectivity within the JNK subfamily, the achievement of JNK isoform-selective inhibitors would be desirable to dissect the contribution of the different isoforms in various pathological conditions. However, the JNK1, 2, and 3 share more than 80% sequence identity, making the development of isoform-specific inhibitors extremely challenging.

In the last decades, pyridinylimidazoles have encountered a remarkable success in the field of p38α MAPK inhibition. This class of inhibitors counts a large number of examples starting from the precursor **SB203580** to the optimized compound **LN950** (Figure 1), until reaching derivatives with low single digit nanomolar IC_{50} values (a review on this class of compounds has recently been published).[18] As can be seen from Figure 1, the reported p38α MAPK inhibitors are also able to inhibit the JNK3 with IC_{50} values in the submicromolar range, thus offering a suitable starting point for optimization when aiming to target this enzyme. In 2016, we published compound **1a** as a balanced dual JNK3/p38α MAPK inhibitor, which served as a precursor for the synthesis of a fluorescent probe used in fluorescence polarization-based binding assays.[19,20] As it is evident from the biological activity of **1a** in comparison to the activity of previous inhibitors, modifying the substitution pattern around the pyridinylimidazole scaffold can contribute to a shift in selectivity toward the JNK3.

Some of us have recently reported the optimization of compound **1a** following a covalent inhibition approach (compound **1b**), which was based on the introduction of an electrophilic moiety able to target a noncatalytic cysteine of the JNK3 that is not conserved in the closely related p38α MAPK.[21] The aim of the herein presented work consists

instead in the achievement of a potent and selective JNK inhibitor by structural modification of the pyridinylimidazole scaffold following the canonical concept of reversible inhibition.

RESULTS AND DISCUSSION

Chemistry. Despite the overall similarity of their structures, the herein reported compounds were synthesized following considerably diverse routes, especially with regard to the construction of the five-membered heterocyclic central core. The synthesis of compounds **5** and **8** was achieved as displayed in Scheme 1. The route leading to the common intermediate **3**, starting from 2-fluoro-4-methylpyridine (**2**), is based on the Marckwald imidazole synthesis[22] and was previously reported by Laufer et al.[23] The substitution on the imidazole-C2-S position was obtained by reacting imidazole-2-thione **3** with the appropriate alkyl halide. Finally, the introduction of the 4-morpholinoaniline moiety was carried out through nucleophilic aromatic substitution in acidic conditions, this representing the final step for most of the herein presented compounds. Applying these conditions to the hydroxyethyl derivative **6** unexpectedly yielded imidazol-2-one **8**, instead of imidazole **7**, as a result of a previously described rearrangement.[24]

The preparation of 2,4,5-trisubstituted imidazole **13** and of 4,5-disubstituted imidazoles **14** and **18a−l** is outlined in Scheme 2. The route providing α-diketone **10** starting from 2-fluoro-4-methylpyridine (**2**) was recently described by Ansideri et al.,[19] whereas the synthesis of intermediates **16a−l** was achieved following a similar approach. Ethanones **15a−l** were obtained by condensation of the appropriate ethyl ester with 2-chloro-4-methylpyridine (**9**) and were subsequently oxidized

Scheme 2. Synthesis of 4,5-Disubstituted Imidazoles 13, 14, and 18a–l[a]
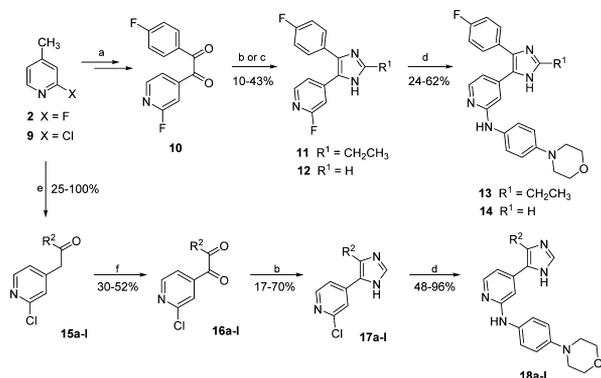
[a]Reagents and conditions: (a) route reported by Ansideri et al.;[19] (b) HCHO$_{(aq)}$, NH$_4$OAc, AcOH, 180 °C microwave irradiation, 2–5 min; (c) propionaldehyde, 7 M NH$_3$ in MeOH, 80 °C, 4 h; (d) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C 16 h; (e) ethyl arylcarboxylate or ethyl alkylcarboxylate, NaHMDS, dry THF, 0 °C 1–5 h; and (f) SeO$_2$, AcOH, 70 °C, 2–3 h; (R^2 = see Table 2).

by SeO$_2$ to the corresponding diketones (16a–l). Microwave-assisted cyclization with formaldehyde and NH$_4$OAc in Radzisewski conditions[25] then afforded the disubstituted imidazoles 12 and 17a–l, whereas propionaldehyde and

methanolic NH$_3$ were employed to obtain the 2-ethylimidazole 11. Finally, introduction of the 4-morpholinoaniline moiety at the pyridine-C2 position, giving the final compounds 13, 14, and 18a–l, was accomplished by the aforementioned nucleophilic aromatic substitution.

The synthesis of 4,5-disubstituted pyridinylimidazoles 38 and 39, featuring a linear alkyl group at the imidazole-C4 position, required a different strategy than the examples having aromatic or branched aliphatic moieties (14 and 18a–l). This was mainly due to the fact that alkyl esters of linear alkanoic acids did not undergo condensation with 2-chloro-4-methyl-pyridine (9) to give the desired ethanone intermediates.

An alternative approach to compounds 38 and 39 could also be employed for the synthesis of the 2,4(2,5)-disubstituted imidazole 43 as well as for the 2,4,5-trisubstituted imidazoles 44 and 45 (Scheme 3). This route started from the commercially available 1-(2-chloropyridin-4-yl)ethan-1-one (21) or from the acylpyridines 22 and 23, which were synthesized by Grignard reaction of the appropriate alkylmagnesium bromide with Weinreb amide 20.

Formation of the corresponding oximes 24–26 and following tosylation of the hydroxyl groups led to intermediates 27–29. Tosylated oximes 27–29 were then first converted into the α-aminoketones 30–32 through Neber rearrangement[26] and subsequently cyclized by KSCN, yielding imidazole-2-thione derivatives 33–35. From these intermediates, it was possible to achieve the disubstituted imidazoles 36

Scheme 3. Synthesis of Imidazoles 38, 39, 43–45, and 47[a]
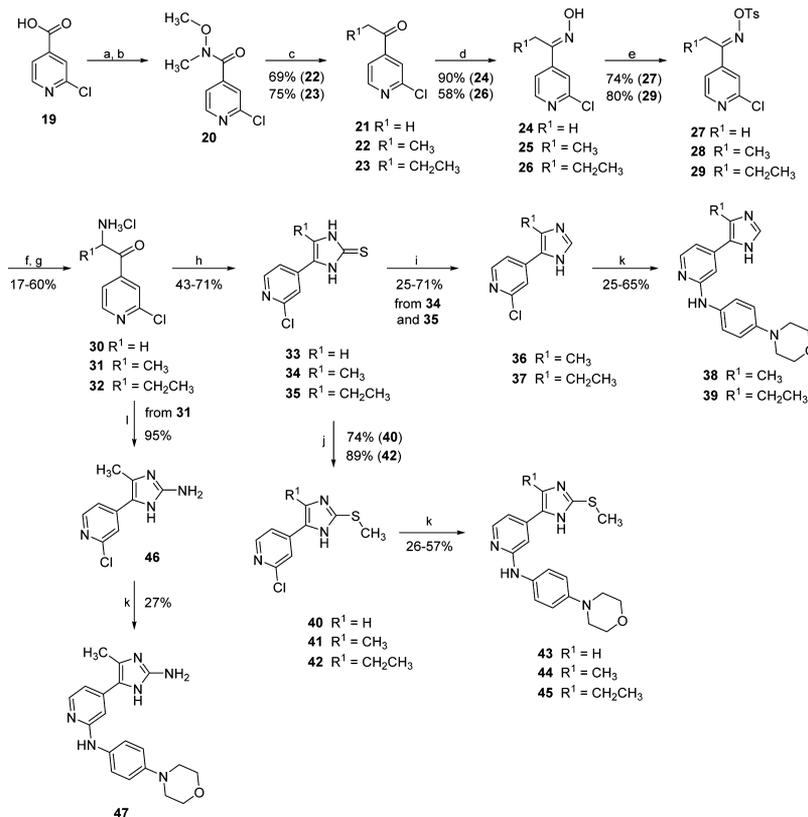
[a]Reagents and conditions: (a) SOCl$_2$, reflux temperature, 5 h; (b) N,O-dimethylhydroxylamine hydrochloride, Et$_3$N, dry DCM, 16 h; (c) EtMgBr or n-PrMgBr, dry THF, −10 °C, 1–3 h; (d) NH$_2$OH·HCl, 20% NaOH$_{(aq)}$, MeOH, H$_2$O, 0 °C, 1–2 h; (e) TsCl, pyridine, rt, 24–72 h; (f) EtOH$_{abs}$, K, 0 °C, 2–16 h; (g) concd HCl, 50 °C, 1–4 h; (h) KSCN, MeOH, reflux temperature, 4 h; (i) H$_2$O$_2$, AcOH, rt, 15 min; (j) MeI, t-BuONa, MeOH, 50 °C, 0.5–3 h; (k) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h; and (l) cyanamide, EtOH, reflux temperature, 2 h.

and **37** by oxidative desulfurization[27] as well as the 2-methylsulfanylimidazoles **40–42** via monomethylation. Alternatively, compound **46** displaying a 2-aminoimidazole core could be prepared by cyclization of the α-aminoketone **31** with cyanamide. Intermediates **36**, **37**, **40–42**, and **47** were then reacted with 4-morpholinoaniline, as previously mentioned, to afford the final compounds **38**, **39**, **43–45**, and **47**, respectively.

Several analogues of compound **44** featuring a different substituent at the pyridine-C2 position (compounds **48a–h** and **48m**, Scheme 4) could be prepared by nucleophilic

Scheme 4. Synthesis of 4(5)-Methyl-2-methylsulfanyl-5-(4)pyridin-4-ylimidazoles 48a–q[a]
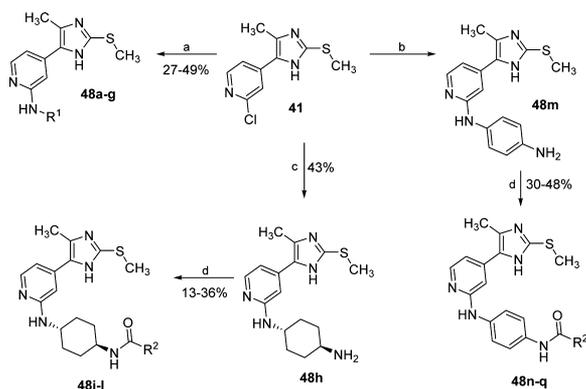
[a]Reagents and conditions: (a) cycloalkylamine (NEAT or n-BuOH), 180 °C, 24–72 h; (b) p-phenylendiamine, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h; (c) trans-diaminocyclohexane, n-BuOH, 180 °C, 72 h; and (d) acyl chloride or anhydride, dry pyridine, rt, 16 h; (R¹, R² = see Table 6).

aromatic substitution of synthone **41** with p-phenylendiamine, 1-phenylethanamine, or with diverse branched or cycloalkyl amines. In addition, compound **48h** and the previously reported **48m**[21] were coupled with different acid chlorides or anhydrides to obtain the corresponding amides **48i–l** and **48n–q** (Scheme 4).

The introduction of a methyl substituent on the imidazole-N atom, providing 1,2,4,5-tetrasubstituted imidazoles **50** and **57**, required a distinct approach depending on the desired N-methylated regioisomer. In fact, double nucleophilic substitution of imidazole-2-thione **34** using excess of methyl iodide almost exclusively afforded the regioisomer bearing the substituent on the N atom away from the pyridine ring (**49**, Scheme 5). The regioselectivity of the methylation reaction was confirmed by crystal structure analysis of intermediate **49** (see Figure S1 in the Supporting Information) and was attributed to the lower steric hindrance offered by the methyl group compared to the pyridine ring. The regioisomer **54**, having the methyl group on the N atom adjacent to the pyridine ring, was instead achieved by cyclizing the α-aminoketone **31** with methyl isothiocyanate, followed by methylation of the sulfur of the resulting N1-methylimidazole-2-thione **51**.

This approach, adapting a procedure published by Xi et al.,[27] represents an unusual route to tetrasubstituted pyridinylimidazoles and was recently reported by some of us for the preparation of tetrasubstituted imidazoles bearing two aromatic moieties at the 4 and 5 positions.[28] The same method could also be employed, using the appropriate alkyl

**Scheme 5. Synthesis of Tetrasubstituted Imidazoles 50 and 57–59[a]**
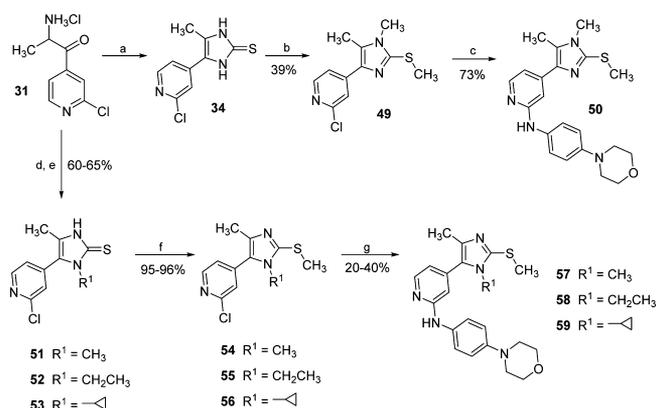
[a]Reagents and conditions: (a) KSCN, MeOH, reflux temperature, 4 h; (b) MeI, t-BuONa, MeOH, 80 °C, 3 h; (c) 4-morpholinoaniline, Pd₂(dba)₃, Xantphos, Cs₂CO₃, dry 1,4-dioxane, 100 °C, 18 h; (d) alkyl isothiocyanate, Et₃N, 60 °C, 16 h; (e) AcOH, 80 °C, 1 h; (f) MeI, t-BuONa, MeOH, 50 °C, 30 min; and (g) 4-morpholinoaniline, Pd₂(dba)₃, XPhos, Cs₂CO₃, dry 1,4-dioxane, 100 °C, 16 h.

isothiocyanate, to achieve the N-ethyl- and the N-cyclopropyl-imidazole derivatives **55** and **56**, respectively. Unlike the majority of the reported compounds, the introduction of the 4-morpholinoaniline moiety, yielding compounds **50** and **57–59**, was carried out by palladium-catalyzed Buchwald–Hartwig aryl amination.

The synthesis of the 1,5-disubstituted imidazole **66**, bearing an aromatic substituent on the imidazole-N1 atom, was performed starting from 2-bromoisonicotinaldehyde (**60**) via a two-step procedure as depicted in Scheme 6. Such a route

Scheme 6. Synthesis of 1,5-Disubstituted Imidazoles 66 and 67[a]
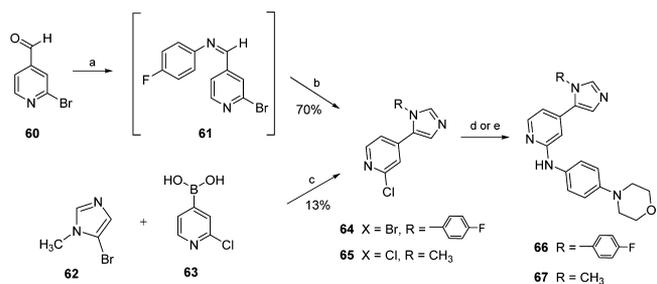
[a]Reagents and conditions: (a) 4-fluoroaniline, AcOH, EtOH, reflux temperature, 2 h; (b) TOSMIC, K₂CO₃, MeOH/dimethoxyethane 2:1, reflux temperature, 3 h; (c) Pd(PPh₃)₄, Cs₂CO₃, H₂O, DMF, 60 °C, 24 h; and (d) 4-morpholinoaniline, t-BuONa, Pd₂(dba)₃, BINAP, toluene, 80 °C, 3 h; (e) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h.

entails the formation of the imine derivative **61** and its direct cyclization through the Van Leusen reaction[29] using toluene sulfonylmethylisocyanide (TOSMIC) and K₂CO₃. The analogous route was unfortunately not accessible for the synthesis of the N1-methyl substituted derivative **67** because of the instability of the corresponding imine intermediate. As an alternative, the preformed N1-methyl imidazole group was introduced through Suzuki cross-coupling reaction[30] between 5-bromo-1-methyl-1H-imidazole (**62**) and pyridinyl-boronic

acid **63** (Scheme 6). The last step of both routes consisted of the introduction of the 4-morpholinoaniline moiety. In the case of the 4-fluorophenyl derivative **64**, this was performed by Buchwald−Hartwig amination giving compound **66**, whereas the acid-catalyzed nucleophilic aromatic substitution was employed for the synthesis of compound **67**.

The 1,2-disubstituted imidazole derivative **71** was obtained starting from 2-chloroisonicotinonitrile (**68**), which was initially reacted in a one-pot procedure described by Voss et al.[31] (Scheme 7). This reaction involves the formation of an

Scheme 7. Synthesis of Imidazol-2-yl Pyridine Derivative 71[a]
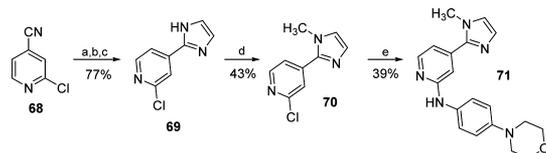
[a]Reagents and conditions: (a) 30% NaOMe in MeOH, MeOH, 40 °C, 1 h; (b) aminoacetaldehyde dimethylacetal, AcOH, MeOH, reflux temperature, 30 min; (c) 6 M HCl, reflux temperature, 18 h; (d) MeI, NaH, dry DMF, rt, 2 h; and (e) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h.

imidate, followed by substitution with acetal-protected aminoacetaldehyde and final ring closure by deprotection, affording 2-(pyridine-4-yl)imidazole **69** in good yield. At last, N-methylimidazole **70** was obtained by nucleophilic substitution with methyl iodide and subsequently reacted with 4-morpholinoaniline as previously discussed, yielding compound **71**.

For the synthesis of compounds **75** and **78**, presenting a methylaminothiazole central core, an approach related to Hantzsch thiazole synthesis[32] was employed (Schemes 8 and

Scheme 8. Synthesis of 2-Methylaminothiazole 75[a]
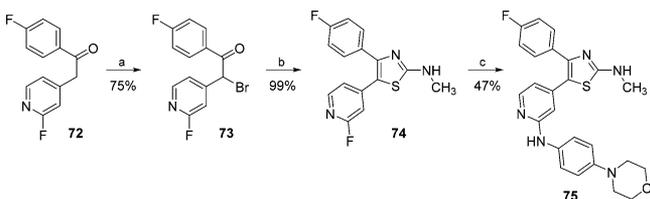
[a]Reagents and conditions: (a) Br$_2$, 30% HBr in AcOH, 75 °C, 2 h; (b) N-methylthiourea, EtOH, reflux temperature, 1 h; and (c) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h.

9). Thiazole **75** was obtained starting from 1-(4-fluorophenyl)-2-(2-fluoropyridin-4-yl)ethan-1-one (**72**),[23] whereas com-

Scheme 9. Synthesis of 2-Methylaminothiazole 78[a]
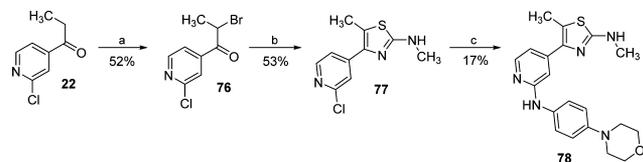
[a]Reagents and conditions: (a) Br$_2$, HBr 30% in AcOH, 75 °C, 4 h; (b) N-methylthiourea, EtOH, reflux temperature, 1 h; and (c) 4-morpholinoaniline, 1.25 M HCl in EtOH, n-BuOH, 180 °C, 16 h.

pound **78** was synthesized starting from 1-pyridinyl-propan-1-one (**22**). Both ketones **72** and **22** were monohalogenated at the α-position under acidic conditions and then cyclized via N-methylthiourea, affording intermediates **74** and **77**, respectively. Conclusively, substitution with 4-morpholinoaniline yielded the desired compounds **75** and **78**.

Biological Evaluation. All synthesized inhibitors were evaluated by enzyme-linked immunosorbent assays[33,34] to determine their ability to inhibit JNK3 and p38α MAPK, and the results are presented in Tables 1−4 and 6.

Table 1. Core Modifications on 4-F-Phenyl-Substituted Derivatives[a]

		IC$_{50}$ ± SD [nM][b]	
Cpd	Core	JNK3	p38α MAPK
5		38 ± 2	17 ± 1
8		142 ± 28	34 ± 4
13		34 ± 0.2	9 ± 4
14		31 ± 2	21 ± 1
66		276 ± 10	323 ± 14
75		15 ± 3	2 ± 0
SB203580	---	---	58 ± 16[d]
SP600125	---	171 ± 36[c]	---

[a]Data of standard inhibitors **SB203580** (p38α MAPK) and **SP600125** (JNK3) in our in-house activity assay are included. [b]IC$_{50}$ values are the mean of three experiments. [c]n = 16. [d]n = 20.

The free terminal aniline moiety of compound **1a** is considered to be potentially responsible for aggregation and therefore might result in assay interference, as also pointed out by analysis through the ZINC 15 pattern tool.[35] For this reason, the p-phenylendiamine moiety at the pyridine-C2 position of compound **1a** was modified in a 4-morpholinoaniline group, which has already been reported as a beneficial substituent in this position.[36] Resulting compound **5** (Table 1) displayed extremely close inhibition values to its analogue **1a** (**1a**, IC$_{50(JNK3)}$ = 24 nM; IC$_{50(p38α MAPK)}$ = 17 nM) and this moiety was, therefore, maintained constant during the investigation of other positions of the scaffold.

The first attempt, which was carried out to shift the preference of compound **5** toward the JNK3, consisted of modifying the central imidazole core together with acting on the substitution at the imidazole-C2 position (Table 1). Transformation of the methylsulfanyl group at the imidazole-C2 position into an ethyl group or removal of the same group,

Table 2. Effect of Different Aryl and Alkyl Substituents at the Imidazole C4(5) Position

| Cpd | R | IC$_{50}$ ± SD [nM][a] | |
		JNK3	p38α MAPK
14	F-phenyl	31 ± 2	21 ± 1
18a	phenyl	37 ± 3	24 ± 3
18b	Cl-phenyl	60 ± 8	38 ± 2
18c	Br-phenyl	131 ± 21	61 ± 2
18d	CF$_3$-phenyl	143 ± 12	22 ± 3
18e	naphthyl	31 ± 3	16 ± 3
18f	H$_3$C-pyrazolyl	758 ± 49	3,259 ± 181
18g	cyclohexyl	1,724 ± 179	726 ± 21
18h	cyclopentyl	2,189 ± 0.136	1,716 ± 81
18i	cyclobutyl	1,757 ± 133	2,265 ± 177
18j	cyclopropyl	1,080 ± 165	4,023 ± 193
18k	H$_3$C,CH$_3$,H$_3$C	>10,000 (32%)[b]	>10,000 (41%)[b]
18l	H$_3$C,CH$_3$	2,833 ± 46	10,000 (50%)[b]
38	H$_3$C propyl	833 ± 139	>10,000 (41%)[b]
39	H$_3$C ethyl	1,198 ± 193	>10,000 (34%)[b]

[a]IC$_{50}$ values are the mean of three experiments. [b]Percent inhibition at indicated concentration.

Table 3. Modification of the Core on Methyl-Substituted Derivatives

| Cpd | Core | IC$_{50}$ ± SD [nM][a] | |
		JNK3	p38α MAPK
38	imidazole	833 ± 139	>10,000 (41%)[b]
44	S-CH$_3$ imidazole	363 ± 34	>10,000 (48%)[b]
47	NH$_2$ imidazole	1,395 ± 230	>10,000 (43%)[b]
50	CH$_3$, S-CH$_3$ imidazole	1,279 ± 179	>10,000 (37%)[b]
57	S-CH$_3$, CH$_3$ imidazole	2,514 ± 312	>10,000 (10%)[b]
58	S-CH$_3$, H$_3$C imidazole	2,091 ± 108	>10,000 (21%)[b]
59	S-CH$_3$, cyclopropyl imidazole	6,509 ± 1,326	>10,000 (40%)[b]
67	imidazole	714 ± 21	>10,000 (32%)[b]
71	imidazole	>10,000 (42%)[b]	>10,000 (15%)[b]
78	S, NH-CH$_3$ thiazole	2,500 ± 92	>10,000 (9%)[b]

[a]IC$_{50}$ values are the mean of three experiments. [b]Percent inhibition at indicated concentration.

Table 4. Effect of Small Alkyl Substituents in the HR I
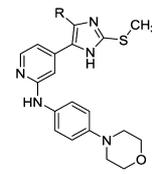
| Cpd | R | IC$_{50}$ ± SD [nM][a] | |
		JNK3	p38α MAPK
43	H	562 ± 21	>10,000 (43%)[b]
44	H$_3$C	363 ± 34	>10,000 (48%)[b]
45	H$_3$C ethyl	1,095 ± 64	>10,000 (38%)[b]

[a]IC$_{50}$ values are the mean of three experiments. [b]Percent inhibition at indicated concentration.

resulting in compounds **13** and **14**, respectively, did not seem to affect the inhibitory activity on the two enzymes. Replacement of the imidazole core with an imidazol-2-one ring instead caused a decrease in the JNK3 inhibitory activity while leaving the inhibition of p38α MAPK unchanged (**8**: IC$_{50(JNK3)}$ = 142 nM; IC$_{50(p38α\ MAPK)}$ = 34 nM). The position of the two nitrogen atoms at the central imidazole core seems to be essential for the inhibition of both enzymes, as the different arrangements of substituents around the five-membered ring of 1,5-disubstituted imidazole **66** resulted in a drop in activity on both target kinases. On the other hand, exchange of 2-sulfanylimidazole with 2-methylaminothiazole (**75**) yielded an increase in inhibitory activity of 2.5- and 8-fold for JNK3 and p38α MAPK, respectively.

To assess the effect of the substituent located in the hydrophobic region (HR) I, the 4-fluorophenyl group was replaced by different aromatic, alkyl, and cycloalkyl moieties (Table 2). In terms of both ligand efficiency (LE) as well as

lipophilic LE (LLE), the 4,5-disubstituted derivative **14** is the most efficient one out of the series of Table 1 and serves, therefore, as the optimal starting point for these modifications. Moreover, this scaffold presents a substantially equal activity compared to its S-methylated analogue **5**, along with a convenient synthetic strategy, facilitating the preparation of a broad range of derivatives.

Most of the 4,5-disubstituted pyridinylimidazoles having an aromatic moiety at the imidazole-C4 position (compounds **18a−f**) revealed to be potent inhibitors for both enzymes, displaying IC$_{50}$ values down to the low double digit nanomolar range. In general, addressing the HR I with a phenyl or monosubstituted phenyl ring resulted in dual inhibitors displaying a slight preference toward p38α MAPK over JNK3. This trend is most distinct in the case of compound **18d** having a 3-(trifluoromethyl)phenyl moiety, which presents a 6-fold higher activity for p38α MAPK than for JNK3. The only aromatic substituent stepping out of this trend was the heteroaromatic N-methylpyrazole of compound **18f**, producing an overall decrease in activity on both kinases while conserving a moderate preference toward JNK3 (**18f**: IC$_{50(JNK3)}$ = 758 nM; IC$_{50(p38\alpha\ MAPK)}$ = 3259 nM). These findings indicate that substitution on the phenyl ring is not beneficial when pursuing selectivity on JNK3 and instead seems to be counterproductive, increasing the activity on p38α MAPK. The reason behind this lack of selectivity can be intuitively explained by considering the dimensions of the hydrophobic pocket known as HR I in the two target kinases. This cleft is wider in the p38α MAPK than in JNK3 mostly because of a difference in the "gatekeeper" residue (Thr106 in p38α MAPK vs Met146 in JNK3). However, as already mentioned in some cocrystallization studies,[37] aromatic moieties can induce a shift of the flexible side chain of Met146 (JNK3), thus essentially abolishing the size differences between the two pockets. As a proof of that, even the bulky 2-naphthyl group of compound **18e** seems to be accommodated in the "reshaped" hydrophobic pocket of JNK3, therefore resulting in a high inhibitory potency. Moreover, attempts of substituting the ortho and meta positions of the phenyl ring, seeking for additional interactions, did not succeed and produced negative outcomes instead (compounds **18b−d**).

The replacement of the aromatic ring at the imidazole-C4 position by cycloalkyl moieties resulted in a dramatic decrease in activity for both enzymes, with IC$_{50}$ values in the low micromolar range. The only exception was the cyclohexyl derivative **18g** that was able to interact with p38α MAPK with an IC$_{50}$ value of 726 nM, 2-fold more potent than on JNK3. The inhibitory effect of compounds **18g−j** on p38α MAPK, decreasing alongside the reduction of the ring size, is symptomatic of a gradually diminished capability of the cyclic group to occupy the spacious cavity of the enzyme. On the other side, JNK3 activity of derivatives **18g−i**, bearing a four- to six-membered ring at the imidazole-C4 position, remained substantially constant, although significantly decreased compared to the parent compound **14**. An analogous scenario occurred in the case of compounds featuring branched aliphatic groups at the same position. The isopropyl derivative **18l**, analogously to the closely related **18j**, showed a significant drop in activity on p38α MAPK, while conserving an IC$_{50}$ value on JNK3 in the low micromolar range. On the other hand, introduction of a *tert*-butyl moiety (**18k**) resulted in a complete loss of activity on both JNK3 and p38α MAPK. Because of their flexibility and low electron density, cyclic and

branched aliphatic groups are presumably unable to promote the Met146 shift and therefore cannot fit in the narrow hydrophobic back pocket of JNK3. A reasonable consequence of this would therefore consist of the flip of the imidazole ring, directing the branched or cyclic alkyl moieties away from the hydrophobic back pocket of the JNK3, thus explaining the similarity of the inhibitory activity regardless of the substituent size.

In agreement with the trend of the series, methyl- and ethyl-substituted imidazoles **38** and **39**, respectively, displayed no inhibition of the p38α MAPK (IC$_{50}$ > 10 μM), however, preserving activity on the JNK3. In particular, the methyl derivative **38** represented the sole compound of this series reaching a submicromolar activity on JNK3 without any remarkable effect on the p38α MAPK. Moreover, this inhibitor also represents the most efficient selective inhibitor of this series in terms of LE and LLE and was therefore chosen as the starting point for further investigations.

Once the methyl substituent at the imidazole-C4 position was selected, our attention was refocused on the central core (Table 3). Because of the presence of the methyl substituent, all derivatives presented in this series lost their potency on the p38α MAPK, with each one displaying an IC$_{50}$ value higher or equal 10 μM. Altering the arrangement of the substituents around the imidazole ring proved beneficial in the case of the 1,5-disubstituted imidazole **67**, slightly increasing its potency compared to the precursor **38**, whereas it was deleterious for the 1,2-disubstituted derivative **71**. Replacement of the imidazole core with a 2-aminomethyl thiazole (**78**) also revealed to be detrimental for the inhibitory activity. A different approach consisted of the introduction of an additional substituent on the imidazole-N atom, together with a reintroduction of the S-methyl group at the C2 position, yielding the tetrasubstituted imidazole scaffold already reported in potent dual JNK3/p38α MAPK as well as JNK3 selective inhibitors.[16,21] In the case of p38α MAPK, the effect of an additional alkyl substituent on the imidazole ring has been reported to be strictly dependent on the position of the substituted N atom. Several examples have demonstrated alkylation of the imidazole-N atom away from the pyridine ring to cause a severe reduction of the activity because of the impossibility to establish a hydrogen bond with the Lys53 of the p38α MAPK.[38,39] Because the same interaction has shown to also occur in the binding to JNK3 (Lys93), an analogous effect was expected on this enzyme as well and was confirmed by the remarkably reduced JNK3 inhibition by compound **50**, carrying a methyl group on the distal imidazole N atom. On the other hand, because several tetrasubstituted JNK3/p38α MAPK inhibitors have been reported with an alkyl substituent on the imidazole N adjacent to the pyridine ring, we assumed this modification to be suitable with our 4-methyl substituted scaffold as well. However, derivatives **57** and **58**, featuring a methyl and an ethyl substituent on the N atom proximal to pyridine, respectively, unexpectedly presented an even lower potency on JNK3 than the supposedly "wrong" regioisomer **50**. The drop in activity appeared to increase with the size of the alkyl substituent, as N-cyclopropyl substituted **59** was almost 3-fold less active compared to its N-methyl analogue **57**. This outcome suggests that despite not hampering the formation of a H bond with the Lys93, alkyl substituents at the imidazole N atom proximal to pyridine reduce the tightness of the binding to the JNK3 active site.

To complete this series, starting from 4,5-disubstituted imidazole **38**, the original S-methyl group or a free amino substituent was introduced at the imidazole-C2 position, affording compounds **44** and **47**, respectively. Although the 2-amino imidazole derivative showed a drop in activity compared to the parent compound **38**, reintroduction of the S-methyl group at the imidazole-C2 position surprisingly produced a 2-fold increase in the inhibitory potency on JNK3 (**44**: $IC_{50(JNK3)}$ = 363 nM; $IC_{50(p38\alpha\ MAPK)}$ > 10 μM). This outcome prompted us to reconsider our previous assumption regarding the role of the 2-methylsulfanyl moiety. Although the S-methyl group exerts no influence on the inhibitory activity when the 4-fluorophenyl moiety is installed at the imidazole-C4 position, it has a significant impact in the case of 4-methyl imidazole derivatives.

In a closer evaluation concerning the influence of the alkyl chain in position 4 of the imidazole core combined with the 2-methylsulfanyl moiety in C2 position, the methyl group (**44**) emerged once more as the substituent presenting the optimal length to target the JNK3 HR I, when compared to the 4-unsubstituted and to the 4-ethyl derivatives **43** and **45**, respectively (Table 4). Comparison of imidazoles **5** (Table 1) and **44** (Table 4) reveals the replacement of the 4-fluorophenyl ring at the imidazole-C4(5) position by a smaller methyl group to result in a 1 order of magnitude loss in JNK3 inhibition and in a complete loss of p38α MAPK inhibitory activity.

To elucidate the binding mode of the 4-methyl-substituted-5-(pyridine-4-yl)imidazole derivatives, as well as to gain insight into the role of the 2-methylsulfanyl group, crystal structures of JNK3 in complex with compounds **38** and **44** were determined (Figure 2).
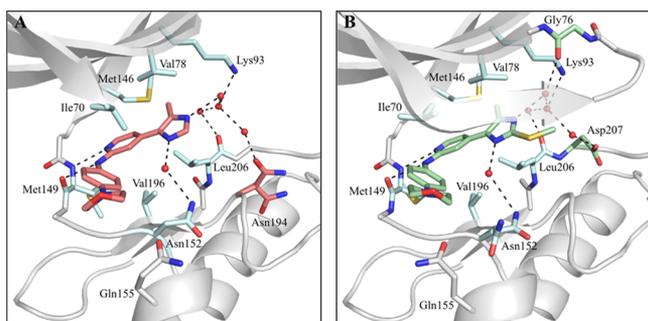

Figure 2. Crystal structures of JNK3 in complex with inhibitors **38** (A) and **44** (B) featuring a pyridinylimidazole scaffold. Only the JNK3 active site is shown. The protein backbone is displayed in gray. The compounds, the side chain of gatekeeper Met146, and a part of the Gly-rich loop are highlighted in stick display. Active site residues with common orientations and interactions are shown in light blue, whereas residues that differ between both complexes are highlighted in the same color as the respective inhibitor. Side chains for which multiple orientations are observed (Asn194 in complex with **38** and Asn152 in complex with **44**) are shown in both orientations. Water molecules are represented as red spheres and hydrogen bonds are shown as black dashed lines.

The structures revealed a similar binding mode of the inhibitors within the adenosine 5′-triphosphate (ATP) pocket of JNK3 (Figure 2). As expected, both scaffolds interacted with the hinge region of the kinase via two hydrogen bonds involving the main chain carbonyl and backbone amine groups of Met149 and mimicking the interactions of the enzyme with ATP[40] as well as with its nonhydrolyzable analogue β,γ-

methyleneadenosine-5′-triphosphate (AMP−PCP, Figure S3, Supporting Information). In both structures, the imidazole-N atom distal from the pyridine ring is part of a network of water-mediated hydrogen bonds, involving the side chain of Lys93 and the main chain of Leu206. Further water-mediated hydrogen bonds in the JNK3-**38** crystal structure (Figure 2A) include the side chain of Asn194, whereas in the JNK3-**44** structure (Figure 2B), the backbone of Gly76 and the side chain of Asp207 are involved. The structure of JNK3 in complex with inhibitor **38** also showed that the imidazole-N atom proximal to the pyridine ring participates in a water-mediated hydrogen bond with the Asn152 side chain and the same interaction seems to be present in the JNK3-**44**, thus explaining the detrimental effect produced by the substitution of this position (compounds **57**−**59**). Multiple hydrophobic interactions comprising the gatekeeper Met146 and the side chains of Ile70, Val78, Val196, and Leu206 were also observed. These interactions have been previously described by Scapin et al.[37] and confer JNK3 selectivity as they cannot be formed in the larger binding pocket of p38α MAPK. The methyl group present in both inhibitors was oriented toward the HR I, which resulted in an identical orientation of the side chain of the gatekeeper residue Met146. The 4-morpholinoaniline moiety, which occupied the solvent-exposed HR II, exhibited higher flexibility and no direct interactions with JNK3, that is, this moiety likely contributes barely or not at all to the binding.

A major structural difference between the two complex structures was observed for the Gly-rich loop. In the JNK3-**38** complex structure, no electron density for residues Gly71−Gly76 was visible because of high local flexibility, a phenomenon also encountered in other JNK3 crystal structures.[41−43] In the JNK3-**44** complex, however, the electron density for this loop was clearly defined, hinting to a structural stabilization of this region upon interaction with the 2-methylsulfanyl moiety in compound **44**.

A superposition of our inhibitor complex structures with crystal structures of JNK3 bound to AMP−PCP and the dual JNK3/p38α MAPK inhibitor by Scapin et al.[37] (PDB code: 1PMN) yielded insights into the structural basis for the observed selectivity of compounds **38** and **44** (Figure 3).

As can be seen from this structural comparison, no movement of the gatekeeper Met146 side chain is induced by compounds **38** and **44** when compared to the AMP−PCP complex, contrary to the dual kinase inhibitor studied by Scapin et al. In the latter crystal structure, an induced fit of side chain 146 occurred to accommodate the dichlorophenyl moiety of the dual kinase inhibitor. Conversely, it appears that the methyl substituent of compounds **38** and **44** was unable to occupy the wider HR I of the p38α MAPK, while possessing the optimal length to target the respective region of JNK3. Therefore, this moiety determined the selectivity achieved over p38α MAPK, demonstrated by the activities of compounds listed in Tables 2−4. In the case of AMP−PCP and compound **44**, another result of the interaction is a downward positioning of the flexible Gly-rich loop. A similar compression of the binding pocket caused by a repositioning of the Gly-rich loop was reported for a JNK3 complex crystal structure by Kamenecka et al.[44] and might be a result of hydrophobic interactions and water-mediated hydrogen bonds provided by inhibitor **44**, which stabilized this otherwise flexible section. Overall, as a result of inhibitor binding, the JNK3 ATP binding pocket in our crystal structures appears somewhat narrower in comparison to the p38α MAPK binding
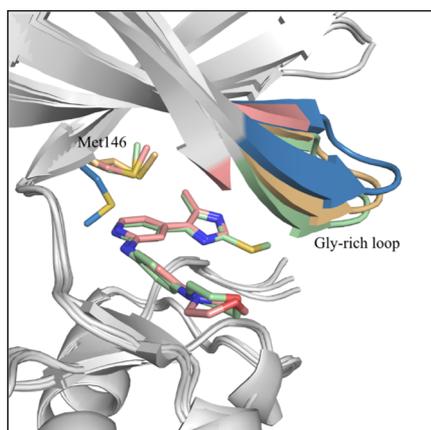
Figure 3. Comparison of the gatekeeper Met146 orientation and the Gly-rich loop positioning upon JNK3 inhibitor binding with other ligand-bound JNK3 structures. Overlay of the JNK3-**44** complex structure (light green), the JNK3-**38** complex structure (light red), the AMP−PCP-bound JNK3 structure (light orange), and the 1PMN structure reported by Scapin, et al.[37] (blue). The superposition was performed using the "align" function in PyMOL. The side chains of the gatekeeper Met146 and the Gly-rich loop are highlighted. Only compounds **38** and **44** are shown for the sake of clarity.

site (where the gatekeeper is Thr106), an effect that is less prominent for the dual kinase inhibitor (Figure 3) and probably responsible for the selectivity of compounds **38** and **44**. With respect to the 2-fold increase in the inhibitory potency on JNK3 of compound **44** over its analogue **38**, the influence of the S-methyl group on the positioning of the Gly-rich loop is the most likely structural reason for the significant gain in affinity.

An additional characterization of the two compounds **38** and **44** included the determination of the protein melting temperature (T_m) in the presence and absence of inhibitors by nano differential scanning fluorimetry (nanoDSF). This methodology consists of assessing the influence of the binding event on the stability of the target protein and is carried out by monitoring temperature-dependent changes in the intrinsic protein fluorescence as a consequence of unfolding. The corresponding curves (Figure S2, Supporting Information) exhibited a significant increase in stability of JNK3 upon inhibitor binding, as can be seen from the associated T_m values (Table 5). The T_m value of JNK3 alone was determined to be 46.3 °C and increased to 53.8 and 54.8 °C in the presence of compounds **38** and **44**, respectively, which correlates well with the results concerning the inhibitory activity and stability of the Gly-rich loop.

Table 5. Determined Melting Temperatures (T_m) for JNK3 Alone and in Complex with Inhibitors 38 and 44

| sample	T_m (°C)[a]
JNK3	46.28 ± 0.58
JNK3-**38**	53.87 ± 0.04
JNK3-**44**	54.83 ± 0.04

[a]Data represent mean value ± SD of a single experiment performed in triplicate. nanoDSF measurements (Figure S2, Supporting Information) were conducted using Prometheus NT.48 (NanoTemper Technologies, Munich).

A further approach in the pursuit of a tighter binding with the JNK3 consisted of modifying the amino moiety at the pyridine-C2 position (Table 6).

An initial attempt was carried out by introducing α-methyl(phenyl)alkylamino moieties (compounds **48a−b**) as well as cycloalkylamino groups (compounds **48c−e**). The former moieties have been reported in potent p38α MAPK inhibitors, for example, **LN950** (Figure 1) and **ML3403**,[45] and were thus introduced to evaluate their effect on JNK3 inhibitory potency. In detail, these substituents were hypothesized to yield an increase in the JNK3 inhibitory activity while conserving selectivity over the p38α MAPK because of the combination with the 4-methyl substituent on the imidazole ring. However, the 3-methyl-2-butylamino group (**48a**) resulted in a loss of activity compared to the 4-morpholinoaniline precursor **44**, although maintaining some selectivity over the p38α MAPK. Substitution with the α-methylbenzylamine, giving rise to compound **48b**, was instead counterproductive as it not only caused a tremendous drop in JNK3 inhibition but also a recovery of the activity on the p38α MAPK (**48b**: IC$_{50(JNK3)}$ = 7610 nM; IC$_{50(p38α MAPK)}$ = 3460 nM). On the other hand, although not reaching the potency of the parent compound **44**, the JNK3 inhibitory activity of compounds bearing cycloalkylamino moieties at the pyridine-C2 position (**48c−e**) increased alongside the size of the aliphatic ring, a trend suggesting the importance of hydrophobic interactions in this area of the molecule. Nevertheless, replacement of the cyclohexyl ring of **48e** with the similar tetrahydropyranyl group (**48f**) yielded, unexpectedly, a remarkable loss of activity on JNK3.

A possible strategy to gain activity and selectivity on JNK3 would consist of targeting the side chain of Gln155 as this residue is replaced by a shorter Asn in the p38α MAPK.[46] As suggested by the structure of the JNK3−**44** complex (Figure 2), this amino acidic residue is located about 4 Å away from the 4-morpholinoaniline-N atom but cannot be reached because of the rigidity of this substituent. Moreover, the 4-morpholinoanilino moiety is only able to accept a hydrogen bond, whereas the Gln residue has the potential to act as both acceptor and donor of hydrogen bonds. For this reason, trans-4-aminocyclohexanol and trans-1,4-diaminocyclohexyl moieties were selected for compounds **48g** and **48h**, respectively, because of a higher flexibility and their additional capability to donate hydrogen bond interactions. In particular, the former moiety is also present in the structure of clinical candidate CC-930, wherein it is reported to interact with the aforementioned Gln155,[47] and included in potent p38α MAPK inhibitors.[45] Unfortunately, despite preserving the selectivity over the p38α MAPK, none of the two inhibitors **48g** and **48h** succeeded in overcoming the activity of the parent compound **44** on the JNK3, with the latter displaying a 3-fold drop in potency. This observation suggests an inability of the introduced moiety to form the desired interaction with the Gln155 side chain or this interaction being compensated by other factors. Additionally, it underlines the necessity of the aromatic moiety at the pyridine-C2 amino function for the binding to the JNK3. The significantly lower activity of compound **48h** could also derive from the not tolerated protonation of its terminal amino functionality. With the aim to reach the Gln155 side chain by the introduction of an additional hydrogen bond acceptor, a series of amides of compound **48h** and of its aromatic counterpart **48m** was synthesized. This approach also permits to seek additional interactions with the enzyme HR II.

Table 6. Influence of Substituents at the Pyridine-C2 Position
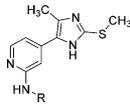
Cpd	R	IC$_{50}$ ± SD [nM][a] JNK3	IC$_{50}$ ± SD [nM][a] p38α MAPK
48a		2,590 ± 173	>10,000 (34%)[b]
48b		7,610 ± 145	3,460 ± 1,680
48c		3,040 ± 194	>10,000 (36%)[b]
48d		1,566 ± 124	>10,000 (47%)[b]
48e		671 ± 74	>10,000 (31%)[b]
48f		2,632 ± 806	>10,000 (27%)[b]
48g		457 ± 72	>10,000 (35%)[b]
48h		941 ± 75	>10,000 (46%)[b]
48i		819 ± 141	>10,000 (37%)[b]
48j		1,216 ± 723	>10,000 (12%)[b]
48k		2,231 ± 80	>10,000 (10%)[b]
48l		2,538 ± 261	>10,000 (13%)[b]
48m[c]		353 ± 58	>10,000 (38%)[b]
48n		415 ± 40	>10,000 (32%)[b]
48o		432 ± 37	>10,000 (33%)[b]
48p		1,539 ± 340	>10,000 (27%)[b]
48q		2,373 ± 504	>10,000 (15%)[b]

[a]IC$_{50}$ values are the mean of three experiments. [b]Percent inhibition at indicated concentration. [c]According to the ZINC patterns tool, compound **48m** represents a potential pan-assay interference compound. However, this compound was synthesized as the intermediate for the preparation of inhibitors **48n–q**. To estimate the impact of the amide moiety present in compounds **48n–q** on the

Table 6. continued

inhibition of the two kinases, the activities of **48m** are listed in this table.

Unfortunately, in neither of the two series, the introduction of amide moieties permitted to gain an inhibitory activity comparable with the precursor **44**. In the series featuring a cycloaliphatic amine (**48i–l**), only the small acetamide derivative **48i** exhibited an almost similar activity to the precursor, whereas bulkier alkyl and aromatic residues displayed a 2- to 3-fold decrease in potency. In an analogous fashion, when considering the series derived from the aromatic intermediate **48m**, compounds bearing a *tert*-butyl or a cyclohexyl amide (**48p** and **48q**, respectively) showed a significant drop in inhibitory activity, with IC$_{50}$ values in the micromolar range. On the other hand, both inhibitors carrying an acetamido or benzamido moiety (**48n** and **48o**, respectively) were still able to inhibit the JNK3 with a potency akin to the free amine derivative **48m**. The comparison of the two amide series also supports the theory of a higher suitability of aromatic substituents at the pyridine-C2 amino position when targeting the JNK3.

Compound **44** resulted as the best inhibitor of the synthesized series and was, therefore, further investigated to achieve a comprehensive characterization. At a first instance, to evaluate the intra-JNK selectivity, compound **44** was tested on the three JNK isoforms (Table 7). As expected, compound **44** inhibited the three isoforms with a similar potency but showed a moderate preference for JNK1 and JNK3 over JNK2.

Table 7. Inhibition Data of Compound 44 on the Three JNK Isoforms

IC$_{50}$ [nM][a]		
JNK1	JNK2	JNK3
119	468	184

[a]Compound **44** was tested by Reaction Biology corporation (Malvern, PA, USA) using a radiometric assay.

Moreover, inhibitor **44** was further screened against a panel of 45 diverse kinases to achieve a preliminary evaluation of its selectivity within the kinome. Out of the kinase panel, 10 kinases (including JNK1) were inhibited more than 50% at a testing concentration of 10 μM (Table S3, Supporting Information).

Additional studies were aimed at evaluating the inhibition of the human-ether-à-go-go related gene (hERG) potassium channels as well as liver cytochromes P450 (CYP450) to highlight potential liabilities of the synthesized scaffold. As displayed in Table 8, compound **44** showed a reduced interference with the hERG channels (IC$_{50}$ > 10 μM).

Regarding interaction with hepatic enzymes, compound **44** displayed low to moderate inhibition of four of the five tested

Table 8. Inhibitory Activity of Compound 44 on hERG Channels and on the Most Relevant CYP Isoforms

hERG inhibition [% inhibition at 10 μM]	CYP450 inhibition [% inhibition at 10 μM]				
	1A9	2C9	2C19	2D6	3A4
38.8	51.5	53.9	35.6	19.0	75.1

isoenzymes, this representing a significantly cleaner profile in comparison with previously reported inhibitors of this class.[48] However, the elevated blockage of the most abundant CYP450 isoform 3A4 still constitutes a serious limit, which needs to be solved by subsequent optimization strategies.

Finally, additional tests were performed to assess the metabolic stability of methyl-substituted pyridinylimidazole **44** upon incubation with human liver microsomes. One of the most serious limitations of previously reported 2-alkylsulfany-limidazoles is their severe metabolization consisting of oxidation of the thioether function to the corresponding sulfoxide.[48] Nevertheless, in vitro assays performed on compound **44** demonstrated a substantial metabolic stability, as approximately 80% of the unmodified compound was still present after 4 h incubation (Figure S6, Supporting Information). The major metabolite formed still appears to be represented by the sulfoxide derivative (8.49%), although modifications at the 4-morpholinoaniline substituent might also be present.

■ CONCLUSIONS

Optimization of 4-(4-fluorophenyl)-5-(pyridin-4-yl)imidazole-based p38α MAPK inhibitors by modification of the five-membered heterocyclic core, the aryl moiety at the imidazole-C4 position, and the pyridine-C2 amino function resulted in a novel series of JNK3 inhibitors exhibiting high selectivity over the closely related p38α MAPK. Biological evaluation of the different pyridinyl-substituted five-membered rings provided valuable insights into the structure activity relationship of this scaffold with respect to JNK3 and p38α MAPK inhibitory potencies. By addressing the HR I with a small methyl group, a significant selectivity toward JNK3 was achieved. This feature is not yet reported for this class of compounds, which have been generally described as p38α MAPK inhibitors. The binding mode at the ATP binding site of the enzyme for this class of compounds was confirmed by X-ray structures of JNK3 crystals incubated with imidazoles **38** and **44**. The most potent inhibitor 4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (**44**) inhibits the JNK3 in the low triple digit nanomolar range, is metabolically stable, and displays a slight selectivity over the JNK2 isoform. Further characterization of this inhibitor highlighted reduced inter-actions with the hERG channel as well with most of the tested CYP450 isoforms.

■ EXPERIMENTAL SECTION

Chemistry. *General.* All chemicals were purchased from commercial sources unless otherwise specified and used without further purification. Thin-layer chromatography (TLC) reaction controls were performed for all reactions using fluorescent silica gel 60 F$_{254}$ plates (Merck) and visualized under natural light and UV illumination at 254 and 366 nm. The purities of all tested compounds were confirmed to be >95% as determined by reverse-phase high-performance liquid chromatography (HPLC) using one of the two following methods. In the case of method 1, the instrument used was a Hewlett Packard HP 1090 Series II LC equipped with a UV diode array detector (DAD) (detection at 230 and 254 nm). The chromatographic separation was performed on a Phenomenex Luna 5u C8 column (150 mm × 4.6 mm, 5 μm) at 35 °C oven temperature. The injection volume was 5 μL and the flow

was 1.5 mL/min using the following gradient: 0.01 M KH$_2$PO$_4$, pH 2.3 (solvent A), MeOH (solvent B), 40% B to 85% B in 8 min; 85% B for 5 min; 85% to 40% B in 1 min; 40% B for 2 min; stop time 16 min. In the case of method 2, an Agilent 1100 Series HPLC system was used, equipped with a UV DAD (detection at 218, 254, and 280 nm). The chromatographic separation was performed on an XBridge C18 column (150 mm × 4.6 mm, 5 μm) and the oven temperature was set to 30 °C. The injection volume was 10 μL and the flow was 1.5 mL/min using the following gradient: 0.01 M KH$_2$PO$_4$, pH 2.3 (solvent A), MeOH (solvent B), 45% B to 85% B in 9 min; 85% B for 6 min; stop time 16 min. Flash column chromatography was performed using an Interchim puriFlash 430 automated flash chromatography system with Davisil LC60A 20−45 μm silica from Grace Davison and Geduran Si60 63−200 μm silica from Merck for the precolumn. Nuclear magnetic resonance (NMR) data were obtained on a Bruker ARX NMR spectrometer at 250 MHz, on a Bruker AVANCE III HD NMR spectrometer at 300 MHz, or on a Bruker AVANCE NMR spectrometer at 400 MHz at ambient temperature. Chemical shifts are reported in parts per million (ppm) relative to tetramethylsilane. All spectra were calibrated against the (residual proton) peak of the deuterated solvent used. Mass spectra were recorded on an Advion expression S electrospray ionization mass spectrometer (ESI-MS) with TLC interface.

Experimental Procedures. *General Procedure for the Nucleophilic Aromatic Substitution with 4-Morpholinoaniline (General Procedure A).* In a pressure vial, the 2-halide pyridine intermediate (1 equiv) and 4-morpholinoaniline (1.5 equiv) were suspended in n-butanol (3 mL) and 1.25 M HCl in EtOH (1 equiv) was added. After tightly closing the vial, the reaction mixture was heated in a heating block at 180 °C and stirred for 18 h. After removing the solvent at reduced pressure, the residue was purified by flash column chromatography.

General Procedure for the Synthesis of Compounds 15a−l (General Procedure B). In a three-neck round-bottom flask under anhydrous conditions, 2-chloro-4-methylpyridine (**9**) (1 equiv) and the appropriate ethyl ester (1 equiv) were dissolved in dry tetrahydrofuran (THF) (2 mL). After cooling the reaction mixture to 0 °C, 2 M sodium bis(trimethylsilyl)amide (NaHDMS) in dry THF (2.2 equiv) was added dropwise and the mixture was stirred at 0 °C for 1.5−5 h. After adding H$_2$O, the aqueous phase was extracted three times with dichloromethane (DCM) or EtOAc and washed with NaCl saturated solution. The combined organic layers were dried over anhydrous Na$_2$SO$_4$ and the solvent was evaporated at reduced pressure. The residue was finally purified by flash column chromatography.

General Procedure for the Synthesis of Compounds 16a−l (General Procedure C). Ethan-1-one intermediates **15a−l** (1 equiv) and SeO$_2$ (1.1 equiv) were suspended in 5−10 mL of glacial AcOH and the reaction mixture was stirred at 65 °C for 2−3 h. After cooling to room temperature (rt), the formed solid residue of Se was removed by filtration and the filtrate was diluted with EtOAc and then washed with saturated NaHCO$_3$ solution four times. Finally, the organic phase was washed with saturated NaCl solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. The residue was purified by flash column chromatography.

General Procedure for the Synthesis of Compounds 17a−l (General Procedure D). In a pressure vial, ethane-1,2-dione

derivatives **16a−l** (1 equiv) and NH$_4$OAc (10 equiv) were suspended in 3 mL of glacial AcOH and after that a 37% aqueous solution of formaldehyde (1 equiv) was added. The reaction vessel was heated in a CEM microwave reactor at 180 °C, with an initial power of 200 W, for 2−5 min. The mixture was added dropwise to NH$_4$OH concentrated solution at 0 °C. The suspension obtained was extracted three times with EtOAc and the combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure. The residue was purified by flash column chromatography.

General Procedure for the Synthesis of Compounds 48a−h (General Procedure E). In a pressure vial, 2-chloro-4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridine (**41**) was suspended in ≈2 mL of cycloalkylamine (in the case of solid amine, 20 equiv of amine was added and the mixture was suspended in ≈2 mL of *n*-butanol). The closed vial was then heated at 180 °C and stirred for 48−120 h. The reaction mixture was poured in H$_2$O and the aqueous layer was extracted three times with EtOAc. The combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure. The residue was finally purified by flash column chromatography.

General Procedure for the Synthesis of Compounds 48i−l and 48n−q (General Procedure F). Under an argon atmosphere, *trans*-N1-(4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)cyclohexane-1,4-diamine (**48h**) or N1-(4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)-benzene-1,4-diamine (**48m**) was dissolved in 1.5 mL of dry pyridine and after that the appropriate acid chloride or anhydride was added and the reaction mixture was stirred at rt for 16 h. The reaction mixture was poured in H$_2$O and the aqueous layer was extracted three times with EtOAc. The combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure. The residue was finally purified by flash column chromatography.

2-Fluoro-4-(4-(4-fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)pyridine (4).[23] The title compound was synthesized as described in the literature[23] and analytical data were in agreement with the reported ones.

4-(4-(4-Fluorophenyl)-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (5). The title compound was synthesized according to general procedure A starting from compound **4** (100 mg, 0.33 mmol) and 4-morpholinoaniline (88.1 mg, 0.49 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 9:1) afforded 61 mg of the desired compound (40% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.61 (s, 3H), 3.00 (br s, 4H), 3.73 (br s, 4H), 6.53−6.75 and 6.88−7.00 (m, 2H), 6.82 (d, J = 7.6 Hz, 2H), 7.09−7.42 (m, 4H), 7.43−7.61 (m, 2H), 7.82−8.12 (m, 1H), 8.55−8.84 (m, 1H), 12.65 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 15.0, 15.1, 49.4, 66.2, 106.1, 106.4, 111.4, 111.7, 115.2 (d, J = 21.2 Hz), 115.7, 115.9, 119.9, 120.3, 126.2, 126.9, 129.5 (d, J = 8.1 Hz), 130.7 (d, J = 8.0 Hz), 133.7, 134.2, 134.8, 137.9, 138.7, 141.9, 142.7, 145.6, 145.9, 147.3, 148.1, 156.7, 161.9 ppm (d, J = 244.4 Hz); MS−FAB *m/z*: [M] calcd for C$_{25}$H$_{24}$FN$_5$OS, 461.2; found, 461.3; HPLC (method 1): t_R = 5.326 min (100%).

2-((4-(4-Fluorophenyl)-5-(2-fluoropyridin-4-yl)-1H-imidazol-2-yl)thio)ethan-1-ol (6).[49] The title compound was synthesized as described in the literature[49] and analytical data were in agreement with the reported ones.

4-(4-Fluorophenyl)-5-(2-((4-morpholinophenyl)amino)-pyridin-4-yl)-1,3-dihydro-2H-imidazol-2-one (8). The title

compound was prepared according to general procedure A starting from **6** (300 mg, 0.90 mmol) and 4-morpholinoaniline (240.6 mg, 1.35 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 85:15) afforded 200 mg of the desired compound (64% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.81−3.11 (m, 4H), 3.58−3.85 (m, 4H), 6.42−6.65 (m, 2H), 6.80 (d, J = 6.6 Hz, 2H), 7.12−7.36 (m, 4H), 7.36−7.59 (m, 2H), 7.99 (dd, J = 4.7, 2.4 Hz, 1H), 8.65 (br s, 1H), 10.64 (br s, 1H), 10.72 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.3, 66.2, 105.2, 111.9, 115.6, 115.8, 115.9 (d, J = 19.0 Hz), 119.7, 120.5, 126.2 (d, J = 2.9 Hz), 129.9 (d, J = 8.0 Hz), 133.6, 138.2, 145.9, 147.9, 153.9, 156.6, 161.7 ppm (d, J = 245.9 Hz); MS−FAB *m/z*: [M + H]$^+$ calcd for C$_{24}$H$_{22}$FN$_5$O$_2$, 431.18; found, 431.30; HPLC (method 1): t_R = 4.552 min (96.7%).

1-(4-Fluorophenyl)-2-(2-fluoropyridin-4-yl)ethane-1,2-dione (10). The title compound was synthesized according to the literature and the analytical data were in agreement with the reported ones.[50]

4-(2-Ethyl-4-(4-fluorophenyl)-1H-imidazol-5-yl)-2-fluoro-pyridine (11). To a solution of **10** (250 mg, 1.01 mmol) in MeOH (5 mL), 7 M ammonia in MeOH (2.89 mL, 20.23 mmol) and propionaldehyde (88.11 mg, 1.52 mmol) were added and the reaction mixture was heated to reflux temperature and stirred for 4 h. After cooling down, the solvent was evaporated at reduced pressure and the residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 94:06), obtaining 125 mg of the desired product (43% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.28 (t, J = 7.6 Hz, 3H), 2.64−2.77 (m, 2H), 7.09 (s, 1H), 7.17−7.40 (m, 3H), 7.47−7.56 (m, 2H), 8.06 (d, J = 5.4 Hz, 1H), 12.41 ppm (br s, 1H); MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{16}$H$_{13}$F$_2$N$_3$, 286.1; found, 286.0; *m/z*: [M − H]$^-$ calcd for C$_{16}$H$_{13}$F$_2$N$_3$, 284.1; found, 284.0; HPLC (method 2): t_R = 3.680 min.

4-(2-Ethyl-4-(4-fluorophenyl)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (13). The title compound was synthesized according to general procedure A starting from 4-(2-ethyl-4-(4-fluorophenyl)-1H-imidazol-5-yl)-2-fluoropyridine (**11**) (85 mg, 0.30 mmol) and 4-morpholinoaniline (80.2 mg, 0.45 mmol). The crude residue was purified twice by flash column chromatography (SiO$_2$, DCM/EtOH 96:04 to 94:06) and (RP-C18, *iso*-propanol/H$_2$O 1:1), obtaining 32 mg of the desired compound (24% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.27 (t, J = 7.6 Hz, 3H), 2.69 (q, J = 7.6 Hz, 2H) 2.96−3.06 (m, 4H), 3.70−3.79 (m, 4H), 6.63−6.92 (m, 3H), 6.97 (br s, 1H), 7.12−7.57 (m, 6H), 7.86−8.08 (m, 1H), 8.60−8.79 (m, 1H), 12.20 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 12.7, 21.2, 49.4, 66.2, 106.2, 115.9, 119.9, 120.2, 127.6, 129.6, 130.6, 134.4, 143.5, 149.6, 156.7, 162.6 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{26}$H$_{26}$FN$_5$O, 444.2; found, 444.2; *m/z*: [M − H]$^-$ calcd for C$_{26}$H$_{26}$FN$_5$O, 442.2; found, 442.2; HPLC (method 2): t_R = 4.960 min (98.6%).

4-(4-(4-Fluorophenyl)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (14). The title compound was synthesized according to general procedure A starting from 2-fluoro-4-(4-(4-fluorophenyl)-1H-imidazol-5-yl)pyridine (**12**)[19] (70 mg, 0.27 mmol) and 4-morpholinoaniline (71.3 mg, 0.40 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 95:05 to 90:10) afforded 70 mg of the desired compound (62% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.93−3.07 (m, 4H), 3.66−3.79 (m, 4H),

6.63−7.00 (m, 4H), 7.12−7.41 (m, 4H), 7.42−7.60 (m, 2H), 7.81 (s, 1H), 7.89−8.13 (m, 1H), 8.57−8.80 (m, 1H), 12.53−12.78 ppm (m, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_{24}H_{22}FN_5O$, 415.18; found, 416.2; m/z: $[M − H]^-$ calcd for $C_{24}H_{22}FN_5O$, 414.2; found, 414.2; HPLC (method 2): t_R = 3.692 min (97.9%).

N-(4-Morpholinophenyl)-4-(4-phenyl-1H-imidazol-5-yl)-pyridin-2-amine (18a). The title compound was synthesized according to general procedure A starting from compound **17a** (100 mg, 0.39 mmol) (for the synthesis of **17a** see Supporting Information) and 4-morpholinoaniline (103.4 mg, 0.58 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 138 mg of the desired compound (89% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.92−3.08 (m, 4H), 3.66−3.80 (m, 4H), 6.68 (dd, *J* = 5.3, 0.9 Hz, 1H), 6.82 (d, *J* = 9.0 Hz, 2H), 6.93 (br s, 1H), 7.30−7.53 (m, 7H), 7.82 (s, 1H), 7.97 (d, *J* = 5.3 Hz, 1H), 8.70 (s, H), 12.66 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.4, 66.1, 106.5, 111.8, 115.9, 120.1, 127.6, 128.1, 128.6, 132.0, 134.0, 136.0, 142.4, 145.7, 147.2, 156.6 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{24}H_{23}N_5O$, 398.2; found, 398.2; m/z: $[M − H]^-$ calcd for $C_{24}H_{23}N_5O$, 396.2; found, 396.3; HPLC (method 1): t_R = 3.513 min (99.1%).

4-(4-(2-Chlorophenyl)-1H-imidazole-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18b). The title compound was synthesized according to general procedure A starting from compound **17b** (100 mg, 0.34 mmol) (for the synthesis of **17b** see Supporting Information) and 4-morpholinoaniline (90.9 mg, 0.51 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 127 mg of the desired compound (87% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.85−3.17 (m, 4H), 3.59−3.89 (m, 4H), 6.51 (d, *J* = 4.8 Hz, 1H), 6.81 (d, *J* = 8.5 Hz, 2H), 6.92 (br s, 1H), 7.28 (d, *J* = 7.5 Hz, 2H), 7.37−7.67 (m, 4H), 7.79−8.01 (m, 2H), 8.60 (s, 1H), 12.63 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.4, 66.2, 104.8, 110.4, 115.9, 120.1, 127.4, 129.8, 130.5, 130.6, 132.5, 133.3, 134.0, 136.0, 136.0, 142.8, 145.7, 146.1, 147.4, 156.7 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{24}H_{22}ClN_5O$, 432.1; found, 432.1; m/z: $[M − H]^-$ calcd for $C_{24}H_{22}ClN_5O$, 430.15; found, 429.8; HPLC (method 2): t_R = 3.671 min (99.4%).

4-(4-(2-Bromophenyl)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18c). The title compound was synthesized according to general procedure A starting from compound **17c** (100 mg, 0.30 mmol) (for the synthesis of **17c** see Supporting Information) and 4-morpholinoaniline (80.2 mg, 0.45 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 100 mg of the desired compound (71% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.91−3.15 (m, 4H), 3.62−3.95 (m, 4H), 6.51 (d, *J* = 4.2 Hz, 1H), 6.82 (m, *J* = 8.0 Hz, 3H), 7.09−7.60 (m, 5H), 7.67−8.01 (m, 3H), 8.63−8.97 (m, 1H), 12.66 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.3, 66.1, 104.7, 110.3, 115.9, 120.4, 124.0, 127.9, 130.6, 130.8, 132.5, 132.9, 133.6, 135.8, 145.9, 145.9, 146.9, 146.9, 156.4 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{24}H_{22}BrN_5O$, 476.1; found, 476.0; m/z: $[M − H]^-$ calcd for $C_{24}H_{22}BrN_5O$, 474.1; found, 473.9; HPLC (method 2): t_R = 3.669 min (99.3%).

N-(4-Morpholinophenyl)-4-(4-(3-(trifluoromethyl)phenyl)-1H-imidazol-5-yl)pyridin-2-amine (18d). The title compound was synthesized according to general procedure A starting from compound **17d** (100 mg, 0.30 mmol) (for the synthesis of **17d** see Supporting Information). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 120 mg of the desired compound (86% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.96−3.10 (m, 4H), 3.68−3.80 (m, 4H), 6.73 (d, *J* = 5.4 Hz, 1H), 6.80−6.92 (m, 3H), 7.34 (d, *J* = 8.8 Hz, H), 7.61−7.81 (m, 3H), 7.84 (br s, 1H), 7.96 (s, 1H), 7.99−8.09 (m, 1H), 8.94 (br s, 1H), 13.01 ppm (br s, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_{25}H_{22}F_3N_5O$, 466.2; found, 465.9; m/z: $[M − H]^-$ calcd for $C_{25}H_{22}F_3N_5O$, 464.18; found, 463.8; HPLC (method 2): t_R = 5.413 min (100%).

N-(4-Morpholinophenyl)-4-(4-(naphthalen-2-yl)-1H-imidazol-5-yl)pyridin-2-amine (18e). The title compound was synthesized according to general procedure A starting from compound **17e** (100 mg, 0.327 mmol) (for the synthesis of **17e** see Supporting Information) and 4-morpholinoaniline (87.5 mg, 0.49 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 90:10) afforded 88 mg of the desired compound (60% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.79−2.96 (m, 4H), 3.61−3.80 (m, 4H), 6.53 (d, *J* = 9.0 Hz, 2H), 6.80 (d, *J* = 5.1 Hz, 1H), 6.87 (br s, 1H), 7.18 (d, *J* = 8.8 Hz, 2H), 7.47−7.65 (m, 3H), 7.88 (s, 1H), 7.90−8.12 (m, 5H), 8.60 (s, 1H), 12.72 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.3, 66.1, 106.0, 111.8, 115.7, 120.1, 126.2, 126.5, 126.8, 127.6, 128.0, 128.1, 132.3, 133.1, 133.7, 136.4, 145.6, 147.7, 156.6 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{28}H_{25}N_5O$, 448.2; found, 448.3; m/z: $[M − H]^-$ calcd for $C_{28}H_{25}N_5O$, 446.2; found, 446.3; HPLC (method 2): t_R = 5.541 min (98.5%).

4-(4-(1-Methyl-1H-pyrazol-4-yl)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18f). The title compound was synthesized according to general procedure A starting from compound **17f** (105.0 mg, 0.40 mmol) (for the synthesis of **17f** see Supporting Information) and 4-morpholinoaniline (107.0 mg, 0.60 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 70:30) afforded 148 mg of the desired compound (92% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.92−3.09 (m, 4H), 3.65−3.79 (m, 4H), 3.88 (s, 3H), 6.83−6.93 (m, 3H), 7.08 (s, 1H), 7.45 (d, *J* = 8.8 Hz, 2H), 7.60 (s, 1H), 7.79 (s, 1H), 7.93 (s, 1H), 8.00 (d, *J* = 5.4 Hz, 1H), 8.87 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 38.6, 49.4, 66.2, 106.2, 111.3, 112.1, 115.9, 120.2, 123.4, 129.4, 131.1, 134.0, 135.6, 137.7, 142.9, 145.8, 146.7, 156.5 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{22}H_{23}N_7O$, 402.2; found, 402.4; m/z: $[M − H]^-$ calcd for $C_{22}H_{23}N_7O$, 400.2; found, 400.5; HPLC (method 2): t_R = 1.766 min (100%).

4-(4-Cyclohexyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18g). The title compound was synthesized according to general procedure A starting from compound **17g** (100 mg, 0.38 mmol) (for the synthesis of **17g** see Supporting Information) and 4-morpholinoaniline (107.6 mg, 0.57 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 120 mg of the desired compound (78% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.14−1.86 (m, 10H), 2.85−2.97 (m, 1H), 2.98−3.06 (m, 4H), 3.65−3.80 (m, 4H), 6.81−6.97 (m, 4H), 7.47 (d, *J* = 8.9 Hz, 2H), 7.64 (s, 1H), 8.04 (d, *J* = 5.3 Hz, 1H), 8.72 (br s, 1H), 12.25 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 25.4, 26.0, 32.4, 34.9, 49.4, 66.2, 105.7, 111.5, 116.0, 120.5, 126.4, 127.7, 134.2,

134.5, 143.2, 145.8, 147.3, 156.9 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{24}$H$_{29}$N$_5$O, 404.2; found, 404.4; *m/z*: [M − H]$^-$ calcd for C$_{24}$H$_{29}$N$_5$O, 402.2; found, 402.2; HPLC (method 2): t_R = 4.730 min (100%).

4-(4-Cyclopentyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18h). The title compound was synthesized according to general procedure A starting from compound **17h** (100 mg, 0.40 mmol) (for the synthesis of **17h** see Supporting Information) and 4-morpholinoaniline (107.0 mg, 0.60 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 9:1) afforded 107 mg of the desired compound (69% yield); ^{1}H NMR (300 MHz, DMSO-*d$_6$*): δ 1.57−1.86 (m, 6H), 1.86−2.05 (m, 2H), 2.96−3.06 (m, 4H), 3.30−3.50 (m, 1H), 3.69−3.77 (m, 4H), 6.85−6.93 (m, 3H), 6.96 (s, 1H), 7.49 (d, *J* = 8.8 Hz, 2H), 7.70 (s, 1H), 8.05 (d, *J* = 5.4 Hz, 1H), 8.78 ppm (s, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 25.1, 32.9, 36.3, 49.5, 66.2, 106.1, 111.6, 116.0, 120.3, 130.9, 131.5, 134.3, 134.8, 142.7, 145.7, 147.2, 156.8 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{23}$H$_{27}$N$_5$O, 390.2; found, 390.0; *m/z*: [M − H]$^-$ calcd for C$_{23}$H$_{27}$N$_5$O, 388.2; found, 387.9; HPLC (method 2): t_R = 4.071 min (99.6%).

4-(4-Cyclobutyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18i). The title compound was synthesized according to general procedure A starting from compound **17i** (100 mg, 0.43 mmol) (for the synthesis of **17i** see Supporting Information) and 4-morpholinoaniline (114.0 mg, 0.64 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 77 mg of the desired compound (48% yield); ^{1}H NMR (300 MHz, DMSO-*d$_6$*): δ 1.77−2.04 (m, 2H), 2.08−2.37 (m, 4H), 2.90−3.11 (m, 4H), 3.70−3.76 (m, 4H), 3.78−3.93 (m, 1H), 6.67−7.00 (m, 4H), 7.50 (d, *J* = 8.9 Hz, 2H), 7.59−7.73 (m, 1H), 8.03 (d, *J* = 5.0 Hz, 1H), 8.62−8.79 (m, 1H), 12.18−12.40 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 7.7, 28.6, 31.1, 49.4, 66.2, 106.0, 111.3, 116.0, 120.7, 128.0, 129.8, 133.7, 134.9, 142.3, 146.1, 146.5, 156.5 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{22}$H$_{25}$N$_5$O, 376.2; found, 376.1; *m/z*: [M − H]$^-$ calcd for C$_{22}$H$_{25}$N$_5$O, 374.2; found, 373.9; HPLC (method 2): t_R = 3.480 min (100%).

4-(4-Cyclopropyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18j). The title compound was synthesized according to general procedure A starting from compound **17j** (150 mg, 0.68 mmol) (for the synthesis of **17j** see Supporting Information) and 4-morpholinoaniline (181.8 mg, 1.02 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 80:20) afforded 153 mg of the desired compound (62% yield); ^{1}H NMR (250 MHz, DMSO-*d$_6$*): δ 0.70−0.81 (m, 2H), 0.91−1.02 (m, 2H), 2.06 (tt, *J* = 8.3, 5.2 Hz, 1H), 2.92−3.10 (m, 4H), 3.60−3.84 (m, 4H), 6.88 (d, *J* = 9.0 Hz, 2H), 7.08 (d, *J* = 5.1 Hz, 1H), 7.22 (s, 1H), 7.51 (d, *J* = 8.8 Hz, 2H), 7.55 (s, 1H), 8.05 (d, *J* = 5.6 Hz, 1H), 8.74 (s, 1H), 12.12 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 7.4, 7.5, 49.5, 66.2, 105.4, 110.9, 116.0, 120.0, 134.1, 134.5, 145.6, 147.2, 156.8 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{21}$H$_{23}$N$_5$O, 362.2; found, 362.6; *m/z*: [M − H]$^-$ calcd for C$_{21}$H$_{23}$N$_5$O, 360.2; found, 360.5; HPLC (method 2): t_R = 2.699 min (100%).

4-(4-(tert-Butyl)-1H-imidazol-y-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18k). The title compound was synthesized according to general procedure A starting from compound **17k** (100 mg, 0.43 mmol) (for the

synthesis of **17k** see Supporting Information) and 4-morpholinoaniline (115.0 mg, 0.64 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 155 mg of the desired compound (96% yield); ^{1}H NMR (300 MHz, DMSO-*d$_6$*): δ 1.20−1.35 (m, 9H), 3.04 (br s, 4H), 3.74 (br s, 4H), 6.70−6.94 (m, 4H), 7.47−7.52 (m, 2H), 8.14−8.20 (m, 1H), 9.04 (br s, 1H), 9.10 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 29.9, 31.6, 49.3, 66.1, 111.4, 114.9, 116.0, 120.6, 125.6, 133.4, 137.5, 138.7, 147.0, 156.1, 158.3, 158.7 ppm; ESI-MS *m/z*: [M + H]$^+$ calcd for C$_{22}$H$_{27}$N$_5$O, 378.2; found, 378.3; ESI-MS *m/z*: [M − H]$^-$ calcd for C$_{22}$H$_{27}$N$_5$O, 376.2; found, 376.1; HPLC (method 2): t_R = 2.860 min (100%).

4-(4-Isopropyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (18l). The title compound was synthesized according to general procedure A starting from compound **17l** (100 mg, 0.45 mmol) (for the synthesis of **17l** see Supporting Information) and 4-morpholinoaniline (119.4 mg, 0.67 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 124 mg of the desired compound (76% yield); ^{1}H NMR (300 MHz, methanol-*d$_4$*): δ 1.31 (d, *J* = 7.0 Hz, 6H), 3.07−3.16 (m, 4H), 3.25−3.45 (m, 1H), 3.80−3.91 (m, 4H), 6.86−6.94 (m, 2H), 6.96−7.03 (m, 2H), 7.30−7.39 (m, 2H), 7.65 (s, 1H), 8.03 ppm (d, *J* = 6.2 Hz, 1H); ^{13}C NMR (101 MHz DMSO-*d$_6$*): δ 22.4, 24.6, 49.5, 66.1, 106.2, 111.6, 115.9, 119.8, 131.9, 134.4, 134.5, 134.6, 143.8, 145.5, 147.2, 156.8 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_{21}$H$_{25}$N$_5$O, 364.2; found, 364.5; *m/z*: [M − H]$^-$ calcd for C$_{21}$H$_{25}$N$_5$O, 362.2; found, 362.3; HPLC (method 2): t_R = 2.492 min (98.6%).

2-Chloro-4-(4-methyl-1H-imidazol-5-yl)pyridine (36). Compound **34**[21] (1.0 g, 4.43 mmol) was suspended in glacial AcOH (10 mL) and subsequently 30% H$_2$O$_2$ (602.7 mg, 17.72 mmol) was added dropwise and the reaction mixture was stirred at rt for 15 min. After adding H$_2$O, the pH was adjusted to 8 using K$_2$CO$_3$ saturated solution and the aqueous phase was extracted five times with EtOAc. The combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure, affording 230 mg of the product which was used in the following step without further purification (25% yield); ^{1}H NMR (300 MHz, DMSO-*d$_6$*): δ 2.47 (s, 3H), 7.62 (dd, *J* = 5.3, 1.3 Hz, 1H), 7.65 (br s, 1H), 7.69 (s, 1H), 8.33 ppm (d, *J* = 5.2 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 11.7, 118.8, 119.1, 127.9, 130.5, 134.9, 145.9, 149.8, 150.8 ppm; MS-ESI *m/z*: [M + H]$^+$ calcd for C$_9$H$_8$ClN$_3$, 194.0; found, 194.0; *m/z*: [M − H]$^-$ calcd for C$_9$H$_8$ClN$_3$, 192.0; found, 191.8; HPLC (method 2): t_R = 1.375 min.

2-Chloro-4-(4-ethyl-1H-imidazol-5-yl)pyridine (37). Compound **35** (400 mg, 1.67 mmol) (for the synthesis of compound **35** see Supporting Information) was suspended in glacial AcOH (10 mL) and subsequently 30% H$_2$O$_2$ (227.2 mg, 6.68 mmol) was added dropwise and the reaction mixture was stirred at rt for 40 min. The reaction mixture was concentrated at reduced pressure and after that 20 mL of K$_2$CO$_3$ saturated solution was added. The aqueous layer was extracted five times with EtOAc and the combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure, affording 230 mg of the product which was used in the following step without further purification (71% yield); ^{1}H NMR (300 MHz, DMSO-*d$_6$*): δ 1.22 (t, *J* = 7.4 Hz, 3H), 2.85 (q, *J* = 7.4 Hz, 2H), 7.58 (d, *J* = 5.0 Hz, 1H), 7.62 (s, 1H), 7.70 (s, 1H), 8.33 ppm (d, *J* = 5.1 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-*d$_6$*): δ 13.4, 18.8, 119.1, 119.4, 129.7,

133.8, 135.2, 145.8, 149.9, 150.8 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{10}H_{10}ClN_3$, 208.0; found, 208.1; m/z: $[M − H]^−$ calcd for $C_{10}H_{10}ClN_3$, 206.0; found, 205.9; HPLC (method 2): t_R = 1.653 min.

4-(4-Methyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)-pyridin-2-amine (38). The title compound was synthesized according to general procedure A starting from compound **36** (100 mg, 0.47 mmol). The crude product was purified twice by flash column chromatography (SiO₂, DCM/EtOH 90:10 to 80:20), (SiO₂, EtOAc), obtaining 38 mg of the desired compound (25% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 2.42 (s, 3H), 2.94−3.09 (m, 4H), 3.63−3.81 (m, 4H), 6.88 (d, J = 9.0 Hz, 2H), 6.94 (d, J = 5.1 Hz, 1H), 7.07 (s, 1H), 7.51 (d, J = 9.1 Hz, 2H), 7.60 (s, 1H), 8.04 (d, J = 5.4 Hz, 1H), 8.69 (s, 1H), 12.15 ppm (br s, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 11.4, 49.5, 66.2, 105.3, 110.8, 116.0, 119.7, 124.7, 133.0, 133.9, 134.6, 143.6, 145.5, 147.2, 156.8 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{19}H_{21}N_5O$, 336.2; found, 336.2; m/z: $[M − H]^−$ calcd for $C_{19}H_{21}N_5O$, 334.2; found, 334.1; HPLC (method 2): t_R = 1.871 min (100%).

4-(4-Ethyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)-pyridin-2-amine (39). The title compound was synthesized according to general procedure A starting from compound **37** (100 mg, 0.48 mmol). The crude product was purified twice by flash column chromatography (SiO₂, DCM/EtOH 90:10 to 80:20), (SiO₂, EtOAc), obtaining 110 mg of the desired compound (65% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 1.22 (t, J = 7.5 Hz, 3H), 2.70−2.90 (m, 2H), 2.92−3.11 (m, 4H), 3.62−3.85 (m, 4H), 6.80−6.98 (m, 3H), 7.07 (br s, 1H), 7.51 (d, J = 9.0 Hz, 2H), 7.61 (s, 1H), 8.04 (d, J = 5.0 Hz, 1H), 8.70 (br s, 1H), 12.03−12.48 ppm (m, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 13.9, 18.5, 49.5, 66.2, 105.6, 111.0, 116.0, 119.8, 130.6, 132.5, 134.2, 134.6, 143.7, 145.5, 147.3, 156.9 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{20}H_{23}N_5O$, 350.4; found, 350.4; m/z: $[M − H]^−$ calcd for $C_{20}H_{23}N_5O$, 348.2; found, 348.2; HPLC (method 2): t_R = 1.774 min (99.4%).

2-Chloro-4-(2-(methylthio)-1H-imidazol-5-yl)pyridine (40). Under an argon atmosphere, compound **33** (500 mg, 2.36 mmol) (for the synthesis of compound **33** see Supporting Information) and *t*-BuONa (454 mg, 4.72 mmol) were dissolved in dry MeOH (20 mL) and after cooling the reaction mixture to 0 °C, methyl iodide (147.5 μL, 2.36 mmol) was added and the reaction mixture was stirred at 0 °C for 30 min. The reaction mixture was then heated to 55 °C and stirred for 3 h. After cooling to rt, the solvent was evaporated at reduced pressure and H₂O was added. The aqueous phase was then extracted two times with EtOAc and the combined organic layers were dried over anhydrous Na₂SO₄ and concentrated at reduced pressure. The residue was finally purified by flash column chromatography (SiO₂, DCM/EtOH 100:0 to 90:10) giving 396 mg of the desired compound (74% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 2.59 (s, 3H), 7.64−7.72 (m, 1H), 7.73−7.79 (m, 1H), 8.03 (s, 1H), 8.31 (dd, J = 5.3, 1.8 Hz, 1H), 12.70 ppm (br s, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_9H_8ClN_3S$, 226.0; found, 225.9; m/z: $[M − H]^−$ calcd for $C_9H_8ClN_3S$, 224.0; found, 223.9; HPLC (method 1): t_R = 4.096 min.

2-Chloro-4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridine (41). The title compound was prepared as previously described[21] and analytical data were in agreement with the reported ones.

2-Chloro-4-(4-ethyl-2-(methylthio)-1H-imidazol-5-yl)-pyridine (42). In a pressure vial, compound **35** (400 mg, 1.67 mmol) (for the synthesis of compound **35** see Supporting Information) and *t*-BuONa (160.5 mg, 1.67 mmol) were dissolved in dry MeOH (15 mL) and after cooling the reaction mixture to 0 °C, methyl iodide (203 μL, 3.26 mmol) was added. The vial was tightly closed and the mixture was stirred at 50 °C for 1 h. The solvent was evaporated at reduced pressure and the residue was purified by flash column chromatography (SiO₂, DCM/EtOH 99:01 to 95:05), affording 378 mg of the product (89% yield); ¹H NMR (300 MHz, CDCl₃): δ 1.32 (t, J = 7.6 Hz, 3H), 2.63 (s, 3H), 2.89 (q, J = 7.6 Hz, 2H), 7.50 (dd, J = 5.3, 1.5 Hz, 1H), 7.64 (br s, 1H), 8.35 ppm (dd, J = 5.3, 0.4 Hz, 1H); ¹³C NMR (101 MHz, CDCl₃): δ 13.5, 16.6, 19.4, 119.3, 120.7, 132.4, 135.1, 141.7, 145.0, 149.4, 151.9 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{11}H_{12}ClN_3S$, 254.0; found, 254.0; m/z: $[M − H]^−$ calcd for $C_{11}H_{12}ClN_3S$, 252.0; found, 252.0; HPLC (method 2): t_R = 3.575 min.

4-(2-(Methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (43). The title compound was synthesized according to general procedure A starting from **40** (100 mg, 0.44 mmol) and 4-morpholinoaniline (117.6 mg, 0.66 mmol). Purification by flash column chromatography (SiO₂, DCM/EtOH 100:0 to 90:10) afforded 92 mg of the desired compound (57% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 2.54−2.62 (m, 3H), 2.92−3.09 (m, 4H), 3.63−3.78 (m, 4H), 6.87 (d, J = 7.8 Hz, 2H), 6.96 (d, J = 4.5 Hz, 1H), 7.17 (br s, 1H), 7.53 (d, J = 7.8 Hz, 2H), 7.76 (br s, 1H), 8.00 (d, J = 4.5 Hz, 1H), 8.75 (br s, 1H), 12.34−12.62 ppm (m, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 15.3, 49.5, 66.2, 104.1, 109.5, 116.0, 116.8, 119.7, 120.3, 134.6, 139.2, 142.0, 145.4, 147.3, 156.9 ppm; MS−FAB m/z: $[M]$ calcd for $C_{19}H_{21}N_5OS$, 367.1; found, 367.2; HPLC (method 1): t_R = 2.501 min (100%).

4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (44). The title compound was synthesized according to general procedure A starting from **41** (100 mg, 0.42 mmol) and 4-morpholinoaniline (112.3 mg, 0.63 mmol). Purification by flash column chromatography (SiO₂, DCM/EtOH 100:0 to 80:20) afforded 42 mg of the desired compound (26% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 2.39 (br s, 3H), 2.52−2.60 (m, 3H), 2.93−3.09 (m, 4H), 3.61−3.83 (m, 4H), 6.78−6.94 (m, 3H), 7.06 (br s, 1H), 7.52 (d, J = 7.8 Hz, 2H), 8.03 (d, J = 4.3 Hz, 1H), 8.74 (br s, 1H), 12.27 ppm (br s, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 15.4, 25.4, 49.5, 66.2, 105.3, 110.5, 116.0, 119.6, 134.6, 145.4, 147.3, 156.7 ppm; MS−FAB m/z: $[M + H]^+$ calcd for $C_{20}H_{23}N_5OS$, 382.2; found, 382.3; HPLC (method 1): t_R = 3.024 min (96.4%).

4-(4-Ethyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (45). The title compound was synthesized according to general procedure A starting from compound **42** (100 mg, 0.39 mmol) and 4-morpholinoaniline (103.4 mg, 0.58 mmol). Purification by flash column chromatography (SiO₂, DCM/EtOH 99:01 to 90:10) afforded 51 mg of the desired compound (33% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 1.21 (t, J = 7.5 Hz, 3H), 2.55 (s, 3H), 2.78 (q, J = 7.4 Hz, 2H), 2.96−3.08 (m, 4H), 3.66−3.81 (m, 4H), 6.69−6.96 (m, 3H), 7.03 (br s, 1H), 7.52 (d, J = 8.8 Hz, 2H), 8.03 (d, J = 5.3 Hz, 1H), 8.67−8.81 (m, 1H), 12.07−12.37 ppm (m, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 13.9, 15.3, 18.7, 49.5, 66.2, 105.5, 110.8, 116.0,

119.8, 133.2, 134.5, 139.7, 143.0, 145.5, 147.3, 151.6, 156.8 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{21}H_{25}N_5OS$, 396.2; found, 396.3; m/z: $[M − H]^−$ calcd for $C_{21}H_{25}N_5OS$, 394.2; found, 394.1; HPLC (method 2): t_R = 3.499 min (97.0%).

4-(2-Chloropyridin-4-yl)-5-methyl-1H-imidazol-2-amine (46). Cyanamide (652 mg, 15.52 mmol) was dissolved in EtOH (30 mL) and after heating at reflux temperature, compound **31** was added portionwise over 1 h and the mixture was stirred at the same temperature further for 3 h. After cooling down, the solvent was evaporated at reduced pressure and the residue was purified by flash column chromatography (SiO₂, DCM/EtOH/Et₃N 95:05:0 to 80:18:2), obtaining 900 mg of the desired product (95% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 2.37 (s, 3H), 7.53 (d, J = 4.9 Hz, 1H), 7.62 (br s, 3H), 8.42 (d, J = 5.1 Hz, 1H), 12.86 ppm (br s, 1H); ¹³C NMR (75 MHz, DMSO-d_6): δ 10.7, 117.6, 118.9, 119.4, 124.4, 139.2, 147.0, 150.3, 151.1 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_9H_9ClN_4$, 209.0; found, 208.9; m/z: $[M − H]^−$ calcd for $C_9H_9ClN_4$, 207.0; found, 206.9; HPLC (method 2): t_R = 1.524 min.

4-(2-Amino-5-methyl-1H-imidazol-4-yl)-N-(4-morpholinophenyl)pyridin-2-amine (47). The title compound was synthesized according to general procedure A starting from compound **46** (100 mg, 0.48 mmol) and 4-morpholinoaniline (128.3 mg, 0.72 mmol). Purification by flash column chromatography (SiO₂, DCM/MeOH 99:01 to 90:10) and (SiO₂, DCM/MeOH 95:05 to 80:20) afforded 46 mg of the desired compound (27% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 2.30 (s, 3H), 2.95−3.07 (m, 4H), 3.67−3.79 (m, 4H), 6.76 (dd, J = 5.4, 1.4 Hz, 1H), 6.83 (s, 1H), 6.89 (d, J = 9.0 Hz, 2H), 7.25 (br s, 2H), 7.50 (d, J = 9.0 Hz, 2H), 8.11 (d, J = 5.4 Hz, 1H), 8.91 (s, 1H), 12.37 ppm (br s, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 11.0, 49.9, 66.7, 105.9, 110.6, 116.4, 120.1, 120.6, 122.2, 134.4, 137.5, 146.4, 147.4, 148.4, 157.2 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{19}H_{22}N_6O$, 351.2; found, 351.1; m/z: $[M Z H]^−$ calcd for $C_{21}H_{25}N_5OS$, 349.2; found, 349.1; HPLC (method 2): t_R = 1.876 min (96%).

4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(3-methylbutan-2-yl)pyridin-2-amine (48a). The title compound was synthesized according to general procedure E starting from **41** (85 mg, 0.355 mmol) and 3-methylbutan-2-amine. The crude residue was purified by flash column chromatography (SiO₂, DCM/EtOH 100:0 to 80:20), affording 35 mg of pure product (34% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 0.79−0.94 (m, 6H), 1.03 (d, J = 6.6 Hz, 3H), 1.70−1.86 (m, 1H), 2.36 (br s, 3H), 2.54 (br s, 3H), 3.72−3.89 (m, 1H), 6.35 (d, J = 7.1 Hz, 1H), 6.53−6.92 (m, 2H), 7.86 (d, J = 5.3 Hz, 1H), 12.23 ppm (br s, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 15.4, 16.7, 17.9, 19.2, 32.1, 50.5, 104.1, 108.4, 127.0, 134.5, 139.0, 142.9, 146.5, 158.6 ppm; HPLC (method 1): t_R = 3.355 min (95%).

4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(1-phenylethyl)pyridin-2-amine (48b). The title compound was synthesized according to general procedure E starting from **41** (100 mg, 417 mmol) and 1-phenylethan-1-amine. The crude residue was purified by flash column chromatography (SiO₂, DCM/EtOH 100:0 to 80:20), affording 42 mg of pure product (31% yield); ¹H NMR (400 MHz, CDCl₃): δ 1.38−1.55 (m, 3H), 1.96−2.13 (m, 3H), 2.36−2.54 (m, 3H), 4.50−4.69 (m, 1H), 5.50 (br s, 1H), 6.33 (br s, 1H), 6.77 (br s, 1H), 7.08−7.34 (m, 6H), 7.84−7.98 ppm (m, 1H); ¹³C NMR (101 MHz, CDCl₃): δ 12.2, 16.8, 24.4, 52.3, 103.5, 110.7, 125.8, 127.0,

128.7, 140.7, 142.5, 144.6, 147.1, 158.0 ppm; HPLC (method 1): t_R = 2.748 min (100%).

N-Cyclobutyl-4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-amine (48c). The title compound was synthesized according to general procedure E starting from **41** (150 mg, 0.62 mmol) and cyclobutylamine (48 h). The crude residue was purified by flash column chromatography (SiO₂, DCM/EtOH 97:03 to 90:10), affording 83 mg of pure product (49% yield); ¹H NMR (300 MHz, DMSO-d_6): δ 1.55−1.73 (m, 2H), 1.77−1.95 (m, 2H), 2.19−2.41 (m, 5H), 2.53 (s, 3H), 4.16−4.40 (m, 1H), 6.41−6.77 (m, 3H), 7.84−7.97 (m, 1H), 12.08−12.34 ppm (m, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 11.3, 14.7, 15.5, 30.7, 46.1, 103.3, 109.0, 126.6, 134.5, 138.7, 142.7, 147.5, 158.3 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{14}H_{18}N_4S$, 275.1; found, 275.0; m/z: $[M − H]^−$ calcd for $C_{14}H_{18}N_4S$, 273.1; found, 273.0; HPLC (method 2): t_R = 2.499 min (99%).

N-Cyclopentyl-4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-amine (48d). The title compound was synthesized according to general procedure E starting from **41** (100 mg, 0.42 mmol) and cyclopentylamine (120 h). The crude residue was purified by flash column chromatography (SiO₂, DCM/EtOH 95:05 to 90:10), affording 51 mg of pure product (42% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 1.36−1.74 (m, 6H), 1.84−1.98 (m, 2H), 2.37 (s, 3H), 2.54 (s, 3H), 3.99−4.15 (m, 1H), 6.60−7.02 (m, 3H), 7.88 (d, J = 5.6 Hz, 1H), 12.32 ppm (br s, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_{15}H_{20}N_4S$, 289.1; found, 289.0; m/z: $[M − H]^−$ calcd for $C_{15}H_{20}N_4S$, 287.1; found, 287.0; HPLC (method 2): t_R = 3.265 min (98%).

N-Cyclohexyl-4-(4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-amine (48e). The title compound was synthesized according to general procedure E starting from **41** (100 mg, 0.42 mmol) and cyclohexylamine (72 h). The crude residue was purified twice by flash column chromatography (SiO₂, DCM/EtOH 97:03 to 90:10) and (SiO₂, DCM/EtOH 95:05 to 90:10), affording 35 mg of pure product (27% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 1.11−1.32 (m, 5H), 1.35 (br s, 1H), 1.52−1.63 (m, 1H), 1.64−1.76 (m, 2H), 1.80−1.96 (m, 2H), 2.35 (br s, 3H), 2.53 (s, 3H), 3.62−3.74 (m, 1H), 6.29 (d, J = 7.6 Hz, 1H), 6.53−6.83 (m, 2H), 7.88 (d, J = 5.3 Hz, 1H), 12.19 ppm (br s, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_{16}H_{22}N_4S$, 303.2; found, 303.1; m/z: $[M − H]^−$ calcd for $C_{16}H_{22}N_4S$, 301.2; found, 301.2; HPLC (method 2): t_R = 4.347 min (100%).

4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(tetrahydro-2H-pyran-4-yl)pyridin-2-amine (48f). The title compound was synthesized according to general procedure E starting from **41** (80 mg, 0.33 mmol) and 4-aminotetrahydropyran (120 h). The crude residue was purified by flash column chromatography (SiO₂, DCM/EtOH 95:05 to 90:10), affording 30 mg of pure product (30% yield); ¹H NMR (400 MHz, DMSO-d_6): δ 1.33−1.48 (m, 2H), 1.87 (d, J = 10.6 Hz, 2H), 2.36 (br s, 3H), 2.53 (s, 3H), 3.40−3.45 (m, 2H), 3.76−3.99 (m, 3H), 6.29−6.86 (m, 3H), 7.90 (d, J = 4.5 Hz, 1H), 12.20 ppm (br s, 1H); ¹³C NMR (101 MHz, DMSO-d_6): δ 11.3, 15.4, 32.9, 46.2, 66.0, 104.2, 108.9, 126.6, 134.4, 138.7, 142.6, 147.2, 158.4 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{15}H_{20}N_4OS$, 305.1; found, 305.0; m/z: $[M − H]^−$ calcd for $C_{15}H_{20}N_4OS$, 303.1; found, 303.1; HPLC (method 2): t_R = 1.570 min (96%).

trans-4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)amino)cyclohexan-1-ol (48g). The title com-

pound was synthesized according to general procedure E starting from **41** (100 mg, 0.42 mmol) and *trans*-4-amino-cyclohexanol (484 mg, 4.20 mmol) and adding 2 mL of *n*-butanol (120 h). The crude residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 92:08 to 80:20), affording 37 mg of pure product (27% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.10−1.35 (m, 4H), 1.76−2.00 (m, 4H), 2.22−2.42 (m, 3H), 2.52 (s, 3H), 3.41 (br s, 1H), 3.61 (br s, 1H), 4.47−4.70 (m, 1H), 6.14−6.85 (m, 3H), 7.78−8.02 (m, 1H), 12.06−12.36 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.3, 15.5, 30.6, 34.1, 48.5, 68.5, 104.1, 108.7, 126.5, 134.5, 138.6, 142.5, 147.3, 158.7 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{16}$H$_{22}$N$_4$OS, 319.1; found, 319.1; m/z: [M − H]$^-$ calcd for C$_{16}$H$_{22}$N$_4$OS, 317.1; found, 317.2; HPLC (method 2): t_R = 1.640 min (98%).

trans-N1-(4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)cyclohexane-1,4-diamine (**48h**). The title compound was synthesized according to general procedure E starting from **41** (300 mg, 1.25 mmol) and *trans*-1,4-diaminocyclohexane (2.8 g, 25 mmol) and adding 2 mL of *n*-butanol (72 h). The crude residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 95:05 to 90:10), affording 172 mg of pure product (43% yield); ^{1}H NMR (300 MHz, methanol-d_4): δ 1.23−1.42 (m, 4H), 1.85−2.13 (m, 4H), 2.41 (s, 3H), 2.55 (s, 3H), 2.58−2.75 (m, 1H), 3.56−3.67 (m, 1H), 6.67 (br s, 1H), 6.73 (dd, J = 5.6, 1.5 Hz, 1H), 7.89 ppm (dd, J = 5.6, 0.6 Hz, 1H); ^{13}C NMR (75 MHz, DMSO-d_6): δ 12.6, 15.9, 32.0, 35.5, 49.3, 50.4, 104.4, 109.0, 129.2, 133.0, 139.6, 142.0, 148.0, 159.2 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{16}$H$_{23}$N$_5$S, 318.2; found, 318.0; m/z: [M − H]$^-$ calcd for C$_{16}$H$_{23}$N$_5$S, 316.2; found, 316.1; HPLC (method 2): t_R = 1.234 min (100%).

N-(*trans*-4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)cyclohexyl)acetamide (**48i**). The title compound was synthesized according to general procedure F starting from **48h** (180 mg, 0.57 mmol) and acetic anhydride (116 mg, 1.14 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 90:10 to 80:20) afforded 74 mg of the desired product (36% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.14−1.35 (m, 4H), 1.74−1.87 (m, 5H), 1.90−2.06 (m, 2H), 2.26−2.41 (m, 3H), 2.53 (s, 3H), 3.43−3.57 (m, 1H), 3.63 (br s, 1H), 6.21−6.80 (m, 3H), 7.75 (d, J = 7.8 Hz, 1H), 7.84−7.98 (m, 1H), 12.11−12.32 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.3, 15.4, 22.7, 31.2, 31.4, 47.3, 48.4, 104.2, 108.7, 126.6, 134.4, 138.7, 142.5, 147.1, 158.5, 168.1 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{18}$H$_{25}$N$_5$OS, 360.2; found, 360.1; m/z: [M − H]$^-$ calcd for C$_{18}$H$_{25}$N$_5$OS, 358.2; found, 358.1; HPLC (method 2): t_R = 1.647 min (99%).

N-(*trans*-4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)cyclohexyl)benzamide (**48j**). The title compound was synthesized according to general procedure F starting from **48h** (120 mg, 0.38 mmol) and benzoyl chloride (80 mg, 0.57 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 90:10 to 80:20) afforded 22 mg of the desired product (13% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.20−1.57 (m, 4H), 1.78−2.11 (m, 4H), 2.28−2.43 (m, 3H), 2.54 (s, 3H), 3.60−3.90 (m, 2H), 6.26−6.88 (m, 3H), 7.36−7.58 (m, 3H), 7.75−8.01 (m, 3H), 8.27 (d, J = 7.9 Hz, 1H), 12.04−12.44 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.3, 15.4, 31.1, 31.6, 48.1, 48.6, 104.2, 108.7, 126.6, 127.2, 128.1, 130.9, 134.4, 134.8, 138.7, 142.7, 147.0, 158.5, 165.5 ppm; MS-ESI m/z: [M + H]$^+$ calcd for

C$_{23}$H$_{27}$N$_5$OS, 422.2; found, 422.0; m/z: [M − H]$^-$ calcd for C$_{23}$H$_{27}$N$_5$OS, 420.2; found, 420.0; HPLC (method 2): t_R = 4.076 min (97%).

N-(*trans*-4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)cyclohexyl)cyclohexanecarboxamide (**48k**). The title compound was synthesized according to general procedure F starting from **48h** (120 mg, 0.38 mmol) and cyclohexane carbonyl chloride (83 mg, 0.57 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 95:05 to 80:20) afforded 29 mg of the desired product (18% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.05−1.43 (m, 9H), 1.54−1.87 (m, 7H), 1.90−2.10 (m, 3H), 2.23−2.42 (m, 3H), 2.53 (s, 3H), 3.45−3.72 (m, 2H), 6.15−6.84 (m, 3H), 7.56 (d, J = 7.6 Hz, 1H), 7.76−8.03 (m, 1H), 12.20 ppm (br s, 1H); MS-ESI m/z: [M + H]$^+$ calcd for C$_{23}$H$_{33}$N$_5$OS, 428.2; found, 428.0; m/z: [M − H]$^-$ calcd for C$_{23}$H$_{33}$N$_5$OS, 426.2; found, 426.1; HPLC (method 2): t_R = 5.391 min (98%).

N-(*trans*-4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)amino)cyclohexyl)pivalamide (**48l**). The title compound was synthesized according to general procedure F starting from **48h** (120 mg, 0.38 mmol) and pivaloyl chloride (69 mg, 0.57 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 90:10 to 80:20) afforded 29 mg of the desired product (19% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.08 (s, 9H), 1.15−1.40 (m, 4H), 1.64−1.77 (m, 2H), 1.92−2.05 (m, 2H), 2.25−2.41 (m, 3H), 2.53 (s, 3H), 3.46−3.70 (m, 2H), 6.23−6.77 (m, 3H), 7.13 (d, J = 8.1 Hz, 1H), 7.89 (d, J = 4.5 Hz, 1H), 12.04−12.35 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.3, 15.5, 27.4, 31.0, 31.6, 37.8, 47.5, 48.6, 104.1, 108.6, 126.6, 134.4, 138.7, 142.6, 147.1, 158.5, 176.5 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{21}$H$_{31}$N$_5$OS, 402.2; found, 402.0; m/z: [M − H]$^-$ calcd for C$_{21}$H$_{31}$N$_5$OS, 400.2; found, 400.0; HPLC: t_R = 4.114 min (99%).

N1-(4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)pyridin-2-yl)benzene-1,4-diamine (**48m**).[21] The title compound was prepared as previously described[21] and analytical data were in agreement with the reported ones.

N-(4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)amino)phenyl)acetamide (**48n**). The title compound was synthesized according to general procedure F starting from **48m** (100 mg, 0.32 mmol) and acetic anhydride (49 mg, 0.48 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 37 mg of the desired product (33% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.01 (s, 3H), 2.28−2.44 (m, 3H), 2.54 (s, 3H), 6.75−7.04 (m, 1H), 7.10 (s, 1H), 7.43 (d, J = 8.8 Hz, 2H), 7.58 (d, J = 8.8 Hz, 2H), 8.01−8.13 (m, 1H), 8.80−8.98 (m, 1H), 9.78 (s, 1H), 12.19−12.47 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.5, 15.6, 23.8, 105.9, 111.1, 118.4, 119.9, 127.3, 132.3, 134.2, 137.6, 139.2, 143.1, 147.3, 156.5, 167.9 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{18}$H$_{19}$N$_5$OS, 354.1; found, 354.1; m/z: [M − H]$^-$ calcd for C$_{18}$H$_{19}$N$_5$OS, 352.1; found, 352.2; HPLC (method 2): t_R = 2.035 min (98%).

N-(4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)amino)phenyl)benzamide (**48o**). The title compound was synthesized according to general procedure F starting from **48m** (100 mg, 0.32 mmol) and benzoyl chloride (67 mg, 0.48 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 98:02 to 90:10) afforded 27 mg of the desired product (33% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.41 (s, 3H), 2.56 (s, 3H), 6.97 (br s, 1H), 7.14 (br s,

1H), 7.46−7.73 (m, 7H), 7.90−8.00 (m, 2H), 8.10 (d, J = 5.3 Hz, 1H), 9.01 (s, 1H), 10.12 (s, 1H), 12.29 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 14.2, 15.9, 106.4, 111.5, 118.6, 121.6, 128.0, 128.8, 129.2, 131.4, 131.7, 132.5, 135.6, 138.6, 139.9, 143.7, 147.7, 156.9, 165.5 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_{23}H_{21}N_5OS$, 416.1; found, 415.7; m/z: [M − H]$^-$ calcd for $C_{23}H_{21}N_5OS$, 414.1; found, 413.7; HPLC (method 2): t_R = 4.357 min (97%).

N-(4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)amino)phenyl)cyclohexanecarboxamide (48p). The title compound was synthesized according to general procedure F starting from **48m** (100 mg, 0.32 mmol) and cyclohexane carbonyl chloride (70 mg, 0.48 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 40 mg of the desired product (30% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 1.11−1.50 (m, 5H), 1.56−1.85 (m, 5H), 2.22−2.36 (m, 1H), 2.40 (s, 3H), 2.55 (s, 3H), 6.94 (d, J = 5.0 Hz, 1H), 7.09 (br s, 1H), 7.47 (d, J = 9.0 Hz, 2H), 7.58 (d, J = 9.0 Hz, 2H), 8.06 (d, J = 5.4 Hz, 1H), 8.91 (s, 1H), 9.62 (s, 1H), 12.28 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.5, 15.4, 25.2, 25.4, 29.2, 44.7, 105.7, 110.9, 118.3, 119.7, 127.1, 132.5, 134.2, 137.3, 139.4, 142.8, 147.2, 156.4, 173.6 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_{23}H_{27}N_5OS$, 422.2; found, 422.0; m/z: [M − H]$^-$ calcd for $C_{23}H_{27}N_5OS$, 420.2; found, 420.1; HPLC (method 2): t_R = 4.646 min (99%).

N-(4-((4-(4-Methyl-2-(methylthio)-1H-imidazol-5-yl)-pyridin-2-yl)amino)phenyl)pivalamide (48q). The title compound was synthesized according to general procedure F starting from **48m** (120 mg, 0.39 mmol) and pivaloyl chloride (70 mg, 0.58 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to 90:10) afforded 74 mg of the desired product (48% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 1.22 (s, 9H), 2.41 (s, 3H), 2.55 (s, 3H), 6.96 (d, J = 5.6 Hz, 1H), 7.11 (s, 1H), 7.50 (d, J = 8.8 Hz, 2H), 7.58 (d, J = 8.8 Hz, 2H), 8.07 (d, J = 5.3 Hz, 1H), 9.00 (s, 1H), 9.07 (s, 1H), 12.35 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 11.9, 15.4, 27.3, 38.9, 105.7, 110.9, 118.2, 121.1, 128.8, 132.5, 137.3, 139.5, 142.6, 146.7, 156.2, 175.9 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_{21}H_{25}N_5OS$, 396.2; found, 395.7; m/z: [M − H]$^-$ calcd for $C_{21}H_{25}N_5OS$, 394.2; found, 393.7; HPLC (method 2): t_R = 3.999 min (100%).

2-Chloro-4-(1,5-dimethyl-2-(methylthio)-1H-imidazol-4-yl)pyridine (49). Under an argon atmosphere, compound **34** (250 mg, 1.11 mmol) and *t*-BuONa (213 mg, 2.22 mmol) were dissolved in dry MeOH (10 mL), and after cooling the reaction mixture to 0 °C, methyl iodide (205 μL, 3.32 mmol) was added and the reaction mixture was let to heat to rt. The reaction mixture was then heated to 80 °C and stirred for 3 h. After cooling to rt, the solvent was evaporated at reduced pressure and H$_2$O was added. The aqueous phase was then extracted two times with EtOAc and the combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure. The residue was finally purified by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 90:10) giving 110 mg of the desired compound (39% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.44 (s, 3H), 2.54−2.62 (m, 3H), 3.34 (s, 3H), 7.47−7.77 (m, 2H), 8.34 ppm (d, J = 4.5 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 10.5, 15.4, 30.6, 119.1, 119.4, 130.5, 132.4, 142.3, 145.6, 149.8, 150.8 ppm; HPLC (method 1): t_R = 4.812 min.

4-(1,5-Dimethyl-2-(methylthio)-1H-imidazol-4-yl)-N-(4-morpholinophenyl)pyridin-2-amine (50). Under an argon

atmosphere, tris(dibenzylideneaceton)dipalladium(0) (Pd$_2$(dba)$_3$) (17.5 mg, 0.02 mmol) and 9,9-dimethyl-4,5-bis(diphenylphosphino)xanten (Xantphos) (22.1 mg, 0.04 mmol) were dissolved in dry 1,4-dioxane (5 mL) and stirred for 10 min. After that compound **49** (50 mg, 0.21 mmol), Cs$_2$CO$_3$ (138.1 mg, 0.42 mmol), and 4-morpholinoaniline (56.7 mg, 0.32 mmol) were added and the reaction mixture was heated to 100 °C and stirred for 15 h. After cooling to rt, the reaction mixture was diluted with DCM and the solid residue was removed by filtration. The filtrate was then concentrated at reduced pressure and the residue was purified by flash column chromatography (DCM/EtOH 100:0 to 90:10) giving 61 mg of the desired product (73% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.36−2.44 (m, 3H), 2.54 (m, 3H), 2.93−3.05 (m, 4H), 3.45−3.55 (m, 3H), 3.67−3.76 (m, 4H), 6.79−6.96 (m, 3H), 7.07 (s, 1H), 7.54 (d, J = 7.3 Hz, 2H), 8.00−8.09 (m, 1H), 8.76 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 10.5, 15.8, 30.5, 49.5, 66.2, 106.0, 111.0, 115.9, 119.5, 128.4, 134.4, 134.6, 140.9, 142.9, 145.4, 147.2, 156.7 ppm; MS−FAB m/z: [M] calcd for $C_{21}H_{25}N_5OS$, 395.2; found, 395.3; HPLC (method 1): t_R = 3.156 min (98.7%).

5-(2-Chloropyridin-4-yl)-1,4-dimethyl-1,3-dihydro-2H-imidazole-2-thione (51). In a pressure vial, compound **31** (200 mg, 0.77 mmol) (for the synthesis of compound **31** see Supporting Information) and methyl isothiocyanate (284 mg, 3.88 mmol) were suspended in triethylamine (2 mL), and after closing the vial tightly, the reaction mixture was stirred at 60 °C for 16 h. The excess of triethylamine was evaporated at reduced pressure and the residue was suspended in glacial AcOH and stirred at 80 °C for 1.5 h. The reaction mixture was concentrated at reduced pressure and after that NaHCO$_3$ saturated solution (20 mL) was added and the aqueous phase was extracted four times with EtOAc. The combined organic layers were washed with H$_2$O and NaCl saturated solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 95:05), affording 110 mg of the desired product (60% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 2.11 (s, 3H), 3.45 (s, 3H), 7.47 (dd, J = 5.2, 1.4 Hz, 1H), 7.55−7.63 (m, 1H), 8.47 (d, J = 5.1 Hz, 1H), 12.51 ppm (br s, 1H); ^{13}C NMR (75 MHz, DMSO-d_6): δ 9.6, 32.4, 122.7, 122.9, 123.3, 124.4, 139.5, 150.2, 150.9, 161.8 ppm; MS-ESI m/z: [M − H]$^-$ calcd for $C_{10}H_{10}ClN_3S$, 238.0; found, 238.0; HPLC (method 2): t_R = 2.353 min.

5-(2-Chloropyridin-4-yl)-1-ethyl-4-methyl-1,3-dihydro-2H-imidazole-2-thione (52). The title compound was prepared following the same procedure of compound **51** starting from **31** (200 mg, 0.77 mmol) (for the synthesis of compound **31** see Supporting Information) and ethyl isothiocyanate (335.5 mg, 3.85 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 99:01 to 95:05) afforded 128 mg of the desired compound (65% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 1.15 (t, J = 6.9 Hz, 3H), 2.14 (s, 3H), 4.04 (q, J = 6.7 Hz, 2H), 7.11 (d, J = 4.5 Hz, 1H), 7.21 (s, 1H), 8.46 (d, J = 4.8 Hz, 1H), 12.32 ppm (br s, 1H); ^{13}C NMR (101 MHz, CDCl$_3$): δ 9.6, 14.1, 40.3, 122.5, 123.2, 124.2, 124.8, 139.5, 150.4, 152.4, 160.1 ppm; MS-ESI m/z: [M − H]$^-$ calcd for $C_{11}H_{12}ClN_3S$, 252.0; found, 252.0; HPLC (method 2): t_R = 3.168 min.

5-(2-Chloropyridin-4-yl)-1-cyclopropyl-4-methyl-1,3-dihydro-2H-imidazole-2-thione (53). The title compound was prepared following the same procedure of compound **51** starting from **31** (500 mg, 1.94 mmol) (for the synthesis of

compound **31** see Supporting Information) and cyclopropyl isothiocyanate (962.4 mg, 9.70 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 95:05) afforded 335 mg of the desired compound (60% yield); ^{1}H NMR (300 MHz, DMSO-d_6): δ 0.43−0.53 (m, 2H), 0.79−0.96 (m, 2H), 2.09 (s, 3H), 3.17−3.29 (m, 1H), 7.49 (dd, J = 5.2, 1.4 Hz, 1H), 7.55−7.64 (m, 1H), 8.45 (d, J = 5.2 Hz, 1H), 12.42 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 6.5, 9.6, 27.0, 122.5, 122.9, 123.3, 124.4, 139.6, 149.6, 150.4, 163.9 ppm; MS-ESI m/z: [M − H]$^-$ calcd for C$_{12}$H$_{12}$ClN$_3$S, 264.0; found, 264.0; HPLC (method 2): t_R = 3.057 min.

2-Chloro-4-(1,4-dimethyl-2-(methylthio)-1H-imidazol-5-yl)pyridine (54). In a pressure vial, compound **51** (285 mg, 1.19 mmol) and t-BuONa (114.3 mg, 1.19 mmol) were dissolved in dry MeOH (15 mL), and after cooling the reaction mixture to 0 °C, methyl iodide (217 μL, 3.48 mmol) was added. The vial was tightly closed and the mixture was stirred at 50 °C for 30 min. After evaporating the solvent at reduced pressure, H$_2$O was added and the aqueous phase was extracted four times with EtOAc. The combined organic layers were washed with H$_2$O and NaCl saturated solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure, giving 290 mg of the product which was used in the following step without further purification (96% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 2.28 (s, 3H), 2.66 (s, 3H), 3.52 (s, 3H), 7.12 (dd, J = 5.2, 1.5 Hz, 1H), 7.22−7.24 (m, 1H), 8.44 ppm (d, J = 5.1 Hz, 1H); ^{13}C NMR (101 MHz, CDCl$_3$): δ 13.7, 15.8, 32.3, 122.0, 123.5, 126.8, 138.5, 141.1, 145.5, 149.9, 152.0 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{11}$H$_{12}$ClN$_3$S, 254.0; found, 254.0; HPLC (method 2): t_R = 1.720 min.

2-Chloro-4-(1-ethyl-4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridine (55). The title compound was synthesized following the same procedure of compound **54** starting from **52** (125 mg, 0.49 mmol), t-BuONa (47 mg, 0.49 mmol), and methyl iodide (90 μL, 1.44 mmol) giving 120 mg of the product, which was used in the following step without further purification (95% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 1.13 (t, J = 7.2 Hz, 3H), 2.15 (s, 3H), 2.57 (s, 3H), 3.85 (q, J = 7.2 Hz, 2H), 7.06 (dd, J = 5.1, 1.5 Hz, 1H), 7.13−7.18 (m, 1H), 8.35 ppm (dd, J = 5.1, 0.5 Hz, 1H); ^{13}C NMR (75 MHz, CDCl$_3$): δ 13.3, 15.5, 15.6, 39.7, 122.0, 123.5, 125.6, 138.3, 141.3, 144.4, 149.7, 151.8 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{12}$H$_{14}$ClN$_3$S, 268.0; found, 268.0; HPLC (method 2): t_R = 2.250 min.

2-Chloro-4-(1-cyclopropyl-4-methyl-2-(methylthio)-1H-imidazol-5-yl)pyridine (56). The title compound was synthesized following the same procedure of compound **54** starting from **53** (210 mg, 0.74 mmol), t-BuONa (71.4 mg, 0.74 mmol), and methyl iodide (135 μL, 2.16 mmol) giving 200 mg of the product, which was used in the following step without further purification (95% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 0.59−0.71 (m, 2H), 0.93−1.03 (m, 2H), 2.27 (s, 3H), 2.67 (s, 3H), 3.03−3.14 (m, 1H), 7.19 (dd, J = 5.2, 1.5 Hz, 1H), 7.27−7.31 (m, 1H), 8.39 ppm (d, J = 5.1 Hz, 1H); ^{13}C NMR (75 MHz, CDCl$_3$): δ 9.6, 13.9, 14.6, 26.1, 121.7, 123.1, 126.9, 138.1, 141.3, 148.3, 149.3, 151.5 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{13}$H$_{14}$ClN$_3$S, 280.1; found, 280.0; HPLC (method 2): t_R = 2.763 min.

4-(1,4-Dimethyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (57). In an argon-flushed pressure tube, compound **54** (100 mg, 0.39 mmol), 4-morpholinoaniline (105.3 mg, 0.59 mmol), Pd$_2$(dba)$_3$ (36.1 mg, 0.04 mmol), 2-dicyclohexylphosphino-2′,4′,6′-triisopro-

pylbiphenyl (XPhos) (37.18 mg, 0.08 mmol), and Cs$_2$CO$_3$ (770.2 mg, 2.36 mmol) were suspended in dry 1,4-dioxane, and after closing the vial tightly, the mixture was stirred at 100 °C for 36 h. The solvent was evaporated at reduced pressure and after that NH$_4$Cl saturated solution was added to the residue and the aqueous phase was extracted five times with EtOAc. The combined organic layers were washed twice with H$_2$O and NaCl saturated solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. Finally, the residue was purified twice by flash column chromatography (SiO$_2$, DCM/EtOH 95:05 to 90:10) and (SiO$_2$ DCM/EtOH 97:03) giving 30 mg of the desired product (20% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 2.23 (s, 3H), 2.62 (s, 3H), 3.07−3.23 (m, 4H), 3.46 (s, 3H), 3.83−3.94 (m, 4H), 6.52−6.62 (m, 2H), 6.72 (br s, 1H), 6.92 (d, J = 8.3 Hz, 2H), 7.24 (d, J = 8.4 Hz, 2H), 8.19 ppm (d, J = 5.1 Hz, 1H); ^{13}C NMR (101 MHz, CDCl$_3$): δ 13.7, 15.9, 32.2, 49.6, 66.8, 107.2, 114.1, 116.7, 124.2, 128.4, 131.6, 137.5, 140.4, 144.3, 146.7, 148.5, 157.1 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{21}$H$_{25}$N$_5$OS, 396.2; found, 396.5; m/z: [M − H]$^-$ calcd for C$_{21}$H$_{25}$N$_5$OS, 394.2; found, 394.3; HPLC (method 2): t_R = 2.116 min (99.3%).

4-(1-Ethyl-4-methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (58). Under an argon atmosphere, 4-morpholinoaniline (98.8 mg, 0.55 mmol), Pd$_2$(dba)$_3$ (16.94 mg, 0.02 mmol), XPhos (17.64 mg, 0.04 mmol), and Cs$_2$CO$_3$ (365 mg, 1.12 mmol) were placed and after that compound **55** (100 mg, 0.37 mmol) previously dissolved in 5 mL of dry 1,4-dioxane was added and the reaction mixture was stirred at 100 °C for 18 h. The solvent was evaporated at reduced pressure and after that NH$_4$Cl saturated solution was added to the residue and the aqueous phase was extracted three times with EtOAc. The combined organic layers were washed with H$_2$O and NaCl saturated solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 95:05) giving 30 mg of the desired product (20% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 1.16 (t, J = 7.1 Hz, 3H), 2.20 (s, 3H), 2.63 (s, 3H), 3.05−3.22 (m, 4H), 3.73−3.99 (m, 6H), 6.49−6.64 (m, 2H), 6.81−7.03 (m, 3H), 7.24 (d, J = 8.7 Hz, 2H), 8.19 ppm (d, J = 5.0 Hz, 1H); ^{13}C NMR (101 MHz, CDCl$_3$): δ 13.5, 15.8, 16.1, 39.8, 49.8, 66.9, 107.2, 114.6, 116.8, 124.0, 127.9, 132.3, 137.2, 140.3, 143.0, 148.3, 148.4, 157.7 ppm; MS-ESI m/z: [M + H]$^+$ calcd for C$_{22}$H$_{27}$N$_5$OS, 410.2; found, 410.1; m/z: [M − H]$^-$ calcd for C$_{22}$H$_{27}$N$_5$OS, 408.2; found, 408.1; HPLC (method 2): t_R = 2.427 min (98.0%).

4-(1-Cyclopropyl-4-methyl-2-(methylthio)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine (59). The title compound was synthesized following the same procedure used for the preparation of compound **58** starting from **56** (150 mg, 0.53 mmol), 4-morpholinoaniline (141.7 mg, 0.79 mmol), Pd$_2$(dba)$_3$ (24.7 mg, 0.03 mmol), XPhos (25.2 mg, 0.05 mmol), and Cs$_2$CO$_3$ (524 mg, 1.61 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 100:0 to 95:05) afforded 86 mg of the desired product (40% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 0.63−0.72 (m, 2H), 0.87−0.97 (m, 2H), 2.22 (s, 3H), 2.65 (s, 3H), 2.91−3.02 (m, 1H), 3.09−3.19 (m, 4H), 3.80−3.93 (m, 4H), 6.63 (br s, 1H), 6.66 (d, J = 5.1 Hz, 1H), 6.81 (br s, 1H), 6.91 (d, J = 8.9 Hz, 2H), 7.24 (d, J = 8.9 Hz, 2H), 8.16 ppm (d, J = 5.1 Hz, 1H); ^{13}C NMR (101 MHz, CDCl$_3$): δ 9.4, 13.8, 14.7, 26.1, 49.7, 66.8, 107.0, 114.3, 116.7, 124.0, 128.6, 132.0, 137.0, 140.7, 146.4,

147.1, 148.3, 156.9 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{23}H_{27}N_5OS$, 422.2; found, 422.0; m/z: $[M − H]^−$ calcd for $C_{23}H_{27}N_5OS$, 420.2; found, 420.1; HPLC (method 2): t_R = 2.989 min (98.0%).

2-Bromo-4-(1-(4-fluorophenyl)-1H-imidazol-5-yl)pyridine **(64)**. 2-Bromoisonicotinaldehyde (**60**, 300 mg, 1.61 mmol), 4-fluoroaniline (179 mg, 1.61 mmol), and AcOH (160 μL), were dissolved in EtOH and the reaction mixture was stirred at reflux temperature for 2 h. After cooling to rt, the solvent was evaporated at reduced pressure and the residue was resuspended in a mixture 2:1 of MeOH and 1,2-dimethoxy-ethane (8 mL) and transferred into a three-neck round-bottom flask under an argon atmosphere. TOSMIC (471.5 mg, 2.41 mmol) and K_2CO_3 (445 mg, 3.22 mmol) were added and the mixture was stirred at reflux temperature for 3 h. The mixture was cooled at rt and the solvent was evaporated at reduced pressure. The residue was suspended in DCM and the organic layer was washed three times with H_2O and one time with NaCl saturated solution. The organic phase was dried over anhydrous Na_2SO_4 and evaporated at reduced pressure. Finally, the residue was purified by flash column chromatography (SiO_2, DCM to DCM/EtOH 95:05) yielding 360 mg of the desired compound (70% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.03 (d, J = 4.3 Hz, 1H), 7.30−7.50 (m, 5H), 7.69 (br s, 1H), 8.08 (br s, 1H), 8.24 ppm (d, J = 4.3 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 116.6 (d, J = 22.7 Hz), 120.5, 124.6, 128.3 (d, J = 8.8 Hz), 128.4, 131.9, 132.1 (d, J = 2.9 Hz), 139.5, 141.7, 142.0, 150.3, 161.8 ppm (d, J = 246.6 Hz); MS-ESI m/z: $[M + H]^+$ calcd for $C_{14}H_9BrFN_3$, 318.0; found, 317.8; HPLC (method 1): t_R = 5.51 min.

2-Chloro-4-(1-methyl-1H-imidazol-5-yl)pyridine **(65)**. Tetrakis(triphenylphosphine)palladium (367 mg, 0,317 mmol) was dissolved in dimethylformamide (DMF) (50 mL) and after that 5-bromo-1-methyl-1*H*-imidazole (**62**) (2.04 g, 12.7 mmol), (2-chloropyridin-4-yl)boronic acid (**63**) (1.0 g, 6.35 mmol), Cs_2CO_3 (4.13 g, 12.7 mmol), and H_2O (228 mg, 12.7 mmol) were added and the reaction mixture was stirred at 60 °C for 24 h. The mixture was poured in H_2O and the aqueous phase was extracted five times with EtOAc. The combined organic layers were washed with NaCl saturated solution, dried over anhydrous Na_2SO_4, and concentrated at reduced pressure. The residue obtained was purified by flash column chromatography (SiO_2, DCM/EtOH 95:05 to 90:10), affording 160 mg of the desired product (13% yield); ^{1}H NMR (250 MHz, CDCl_3): δ 3.79 (s, 3H), 7.27 (dd, J = 5.2, 1.6 Hz, 2H), 7.34 (br s, 1H), 7.38 (dd, J = 1.6, 0.6 Hz, 1H), 7.60 (br s, 1H), 8.43 ppm (dd, J = 5.2, 0.6 Hz, 1H); MS-ESI m/z: $[M + H]^+$ calcd for $C_9H_8BrN_3$, 194.0; found, 194.0; HPLC (method 2): t_R = 1.162 min.

4-(1-(4-Fluorophenyl)-1H-imidazol-5-yl)-N-(4-morpholinophenyl)pyridin-2-amine **(66)**. Under an argon atmosphere, compound **64** (200 mg, 0.62 mmol), 4-morpholinoaniline (134.5 mg, 0.75 mmol), *t*-BuONa (83.8 mg, 0.87 mmol), $Pd_2(dba)_3$ (20 mg, 0.02 mmol), and 2,2′-bis(difenilfosfino)-1,1′-binaftile (BINAP) (30 mg, 0.02 mmol) were dissolved in dry toluene (15 mL) and the reaction mixture was then stirred for 3 h at 80 °C. After removing the solvent at reduced pressure, the residue was suspended in H_2O and the aqueous phase was extracted with EtOAc. The combined organic layers were then dried over anhydrous Na_2SO_4 and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (DCM/MeOH 100:0 to 95:05), affording 62 mg of the product (24%

yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 3.01 (s, 4H), 3.73 (s, 4H), 6.41−6.44 (m, 2H), 6.80−6.82 (m, 2H), 7.22−7.24 (m, 2H), 7.37−7.42 (m, 5H), 7.99 (s, 2H), 8.70 ppm (s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 49.3, 66.2, 106.3, 111.8, 115.8, 116.6 (d, J = 22.0 Hz), 127.4 (d, J = 9.0 Hz), 129.9, 132.5, 133.5, 137.3, 140.8, 146.0, 147.9, 156.6, 161.6 ppm (d, J = 244.0 Hz); FAB−MS m/z: $[M]$ calcd for $C_{24}H_{22}FN_5O$, 415.2; found, 415.3; HPLC (method 1): t_R = 5.08 min (100%).

4-(1-Methyl-1H-imidazol-5-yl)-N-(4-morpholinophenyl)-pyridin-2-amine **(67)**. The title compound was synthesized according to general procedure A starting from **65** (140 mg, 0.72 mmol) and 4-morpholinoaniline (192.5 mg, 1.08 mmol). Purification by flash column chromatography (SiO_2, DCM/EtOH 100:0 to 90:10) afforded 50 mg of the desired compound (35% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.97−3.07 (m, 4H), 3.69−3.79 (m, 7H), 6.78−6.85 (m, 2H), 6.89 (m, J = 9.0 Hz, 2H), 7.22 (d, J = 1.2 Hz, 1H), 7.52 (m, J = 9.0 Hz, 2H), 7.76 (br s, 1H), 8.11 (d, J = 5.1 Hz, 1H), 8.85 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 32.9, 49.4, 66.2, 106.9, 111.7, 115.9, 119.9, 128.9, 130.7, 134.0, 137.8, 141.1, 145.8, 147.8, 156.8 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_{19}H_{21}N_5O$, 336.2; found, 336.3; m/z: $[M − H]^−$ calcd for $C_{19}H_{21}N_5O$, 334.2; found, 334.3; HPLC (method 1): t_R = 2.048 min (100%).

2-Chloro-4-(1H-imidazol-2-yl)pyridine **(69)**. To a solution of 2-chloroisonicotinonitrile (**68**) (2.0 g, 14.44 mmol) in MeOH (8 mL), a 30% solution of NaOMe in MeOH (260 μL, 1.44 mmol) was added and the reaction mixture was stirred at 40 °C for 1 h. After that both 2,2-dimethoxyethan-1-amine (1.56 mL, 14.44 mmol) and AcOH (1.56 mL, 27.27 mmol) were added dropwise and the mixture was stirred at reflux temperature for 30 min. After cooling to rt, the mixture was diluted with MeOH (8 mL) and then 6 N HCl solution (7.2 mL, 43.2 mmol) was added and the mixture was stirred at reflux temperature for 18 h. The solvent was evaporated at reduced pressure and after that a 10% solution of K_2CO_3 was added to the residue until reaching pH = 10. The precipitate obtained was filtered off and washed with H_2O, affording 2.01 g of the product which was used for the following step without further purification (77% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 7.30 (br s, 2H), 7.78−7.91 (m, 1H), 7.95 (br s, 1H), 8.36−8.49 (m, 1H), 13.03 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 118.1, 118.6, 124.8, 140.7, 141.9, 150.5, 151.1 ppm; MS-ESI m/z: $[M + H]^+$ calcd for $C_8H_6ClN_3$, 180.0; found, 179.8; m/z: $[M − H]^−$ calcd for $C_8H_6ClN_3$, 178.0; found, 177.8; HPLC (method 1): t_R = 2.346 min.

2-Chloro-4-(1-methyl-1H-imidazol-2-yl)pyridine **(70)**. Under an argon atmosphere, compound **69** (1.72 g, 9.61 mmol) was dissolved in dry DMF (20 mL) and after cooling the reaction mixture to 0 °C, NaH (231 mg, 9.61 mmol) was added and the mixture was stirred at 0 °C for 15 min. After that methyl iodide (1.61 mL, 25.6 mmol) was added dropwise and the reaction mixture was let to heat to rt and stirred for 90 min. The mixture was poured in H_2O and the aqueous phase was extracted three times with DCM. The combined organic layers were dried over anhydrous Na_2SO_4 and the solvent was evaporated at reduced pressure. Finally, the residue was treated with a mixture of *n*-hexane/EtOAc 40:1 and the solid obtained was filtered off and dried in vacuo, affording 793 mg of the product which was used for the following step without further purification (43% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 3.82−3.93 (m, 3H), 7.09 (br s, 1H), 7.40 (br s, 1H), 7.71−

7.78 (m, 1H), 7.80 (br s, 1H), 8.40−8.53 ppm (m, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 34.8, 121.0, 121.6, 125.8, 128.7, 140.9, 142.3, 150.1, 150.8 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_9H_8ClN_3$, 194.0; found, 193.8; HPLC (method 1): t_R = 1.161 min.

4-(1-Methyl-1H-imidazol-2-yl)-N-(4-morpholinophenyl)-pyridin-2-amine (71). The title compound was synthesized according to general procedure A starting from **70** (150 mg, 0.77 mmol) and 4-morpholinoaniline (205 mg, 1.15 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 95:05) afforded 100 mg of the desired compound (39% yield); ^{1}H NMR (400 MHz, DMSO-d_6): δ 2.93−3.09 (m, 4H), 3.66−3.77 (m, 4H), 3.81 (s, 3H), 6.90 (d, J = 8.3 Hz, 2H), 6.95−7.04 (m, 2H), 7.09 (s, 1H), 7.31 (s, 1H), 7.52 (d, J = 8.1 Hz, 2H), 8.15 (d, J = 4.8 Hz, 1H), 8.89 ppm (br s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 34.6, 49.4, 66.1, 107.8, 111.9, 115.9, 120.0, 124.5, 127.9, 133.9, 138.5, 144.3, 145.8, 147.6, 156.7 ppm; FAB−MS m/z: [M] calcd for $C_{19}H_{21}N_5O$, 335.2; found, 335.3; HPLC (method 1): t_R = 2.359 min (100%).

2-Bromo-1-(4-fluorophenyl)-2-(2-fluoropyridin-4-yl)-ethan-1-one (73).[51] Compound **72**[19] (1.0 g, 4.29 mmol) was dissolved in 30% HBr in AcOH (6 mL). After cooling the reaction mixture to 0 °C, Br$_2$ (220 μL, 4.29 mmol) was added dropwise and the reaction mixture was heated for 6 h at 40 °C. After evaporating the solvent at reduced pressure, H$_2$O was added and the pH was adjusted to 8 using NH$_4$OH solution. The water layer was then extracted three times by DCM and the combined organic layers were dried over anhydrous Na$_2$SO$_4$ and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (SiO$_2$, n-hexane/EtOAc 7:3), affording 1.0 g of the desired compound (75% yield). Analytical data were in agreement with the reported ones.[51]

4-(4-Fluorophenyl)-5-(2-fluoropyridin-4-yl)-N-methylthia-zol-2-amine (74). Compound **73** (513 mg, 1.64 mmol) and *N*-methylthiourea (148 mg, 1.64 mmol) were dissolved in EtOH and the reaction mixture was stirred at reflux temperature for 1 h. After cooling to rt, the solvent was evaporated at reduced pressure and then H$_2$O was added. The solution was alkalized to pH = 8 and then extracted three times with DCM. The combined organic layers were dried over anhydrous Na$_2$SO$_4$, concentrated at reduced pressure, and the residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 95:05), obtaining 495 mg of the desired compound (99% yield): ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.89 (d, J = 4.6 Hz, 3H), 6.77 (br s, 1H), 6.93−7.02 (m, 1H), 7.14−7.29 (m, 2H), 7.40−7.52 (m, 2H), 8.04 (d, J = 5.4 Hz, 1H), 8.10 ppm (q, J = 4.8 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 30.8, 106.8 (d, J = 39.5 Hz), 114.2 (d, J = 3.7 Hz), 115.4 (d, J = 21.2 Hz), 120.4 (d, J = 3.7 Hz), 130.9 (d, J = 8.1 Hz), 131.4 (d, J = 3.7 Hz), 145.9 (d, J = 8.8 Hz), 147.6 (d, J = 16.1 Hz), 149.2, 162.0 (d, J = 245.9 Hz), 163.4 (d, J = 234.2 Hz), 168.3 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_{15}H_{11}F_2N_3S$, 304.3; found, 304.1; m/z: [M − H]$^-$ calcd for $C_{15}H_{11}F_2N_3S$, 302.3; found, 302.1; HPLC (method 2): t_R = 7.123 min.

4-(4-Fluorophenyl)-N-methyl-5-(2-((4-morpholinophenyl)amino)pyridin-4-yl)thiazol-2-amine (75). The title compound was synthesized according to general procedure A starting from **74** (200 mg, 0.66 mmol) and 4-morpholinoaniline (176.4 mg, 0.99 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 97:03 to

90:10) afforded 143 mg of the desired compound (47% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.87 (d, J = 4.9 Hz, 3H), 2.94−3.07 (m, 4H), 3.64−3.82 (m, 4H), 6.37 (dd, J = 5.4, 1.7 Hz, 1H), 6.54 (d, J = 1.0 Hz, 1H), 6.82 (m, J = 9.0 Hz, 2H), 7.12−7.25 (m, 2H), 7.27−7.38 (m, 2H), 7.42−7.55 (m, 2H), 7.85 (q, J = 4.6 Hz, 1H), 7.93 (d, J = 5.4 Hz, 1H), 8.69 ppm (s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 30.7, 49.3, 66.1, 107.6, 112.7, 115.2 (d, J = 21.2 Hz), 115.8, 116.3, 120.0, 130.8 (d, J = 8.1 Hz), 131.8 (d, J = 2.9 Hz), 133.7, 141.0, 145.8, 146.7, 147.8, 156.7, 161.7 (d, J = 244.4 Hz), 167.3 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_{25}H_{24}FN_5OS$, 462.2; found, 462.1; m/z: [M − H]$^-$ calcd for $C_{25}H_{24}FN_5OS$, 460.2; found, 460.1; HPLC (method 2): t_R = 8.518 min (98.2%).

2-Bromo-1-(2-chloropyridin-4-yl)propan-1-one (76). 1-(2-Chloropyridin-4-yl)propan-1-one (**22**) (3.0 g, 17.68 mmol) was dissolved in a 30% solution of HBr in AcOH (20 mL), and after cooling the mixture to 0 °C, bromine (900 μL, 17.68 mmol) was added and the reaction mixture was stirred at 45 °C for 2 h and then heated to 75 °C and stirred for additional 2 h. After evaporating the solvent at reduced pressure, H$_2$O was added and the pH was adjusted to 9 using NH$_4$OH solution. The aqueous phase was extracted three times with DCM and the combined organic layers were washed with H$_2$O, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (n-hexane/EtOAc 90:10 to 80:20), affording 2.3 g of the desired product (52% yield); ^{1}H NMR (300 MHz, CDCl$_3$): δ 1.85 (d, J = 6.6 Hz, 3H), 5.07 (q, J = 6.6 Hz, 1H), 7.64 (dd, J = 5.1, 1.5 Hz, 1H), 7.76 (dd, J = 1.4, 0.7 Hz, 1H), 8.53 ppm (dd, J = 5.1, 0.7 Hz, 1H); ^{13}C NMR (75 MHz, CDCl$_3$): δ 19.5, 41.2, 120.4, 123.0, 143.2, 150.9, 152.9, 191.2 ppm; MS-ESI m/z: [M + H]$^+$ calcd for $C_8H_7BrClNO$, 247.9; found, 248.0; m/z: [M + MeOH]$^+$ calcd for $C_8H_7BrClNO$, 279.9; found, 280.0; HPLC (method 2): t_R = 5.761 min.

4-(2-Chloropyridin-4-yl)-N,5-dimethylthiazol-2-amine (77). Compound **76** (1.0 g, 4.0 mmol) and *N*-methylthiourea (362.7 mg, 4.0 mmol) were dissolved in EtOH (20 mL) and the reaction mixture was stirred at reflux temperature for 1 h. The solvent was evaporated at reduced pressure and after that the residue was suspended in H$_2$O and the pH was adjusted to 8 using NH$_4$OH solution. The resulting suspension was extracted three times with DCM and the combined organic layers were washed with H$_2$O and NaCl saturated solution, dried over anhydrous Na$_2$SO$_4$, and concentrated at reduced pressure. Finally, the residue was purified by flash column chromatography (SiO$_2$, DCM/EtOH 95:05) giving 510 mg of the desired product (53% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.43 (s, 3H), 2.83 (d, J = 4.9 Hz, 3H), 7.47 (q, J = 4.9 Hz, 1H), 7.61 (dd, J = 5.1, 1.5 Hz, 1H), 7.64−7.66 (m, 1H), 8.40 ppm (dd, J = 5.1, 0.7 Hz, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 12.2, 30.5, 120.4, 121.2, 121.8, 141.0, 145.7, 149.8, 150.6, 165.5 ppm; MS-ESI m/z: [M + H]$^+$ calcd, for $C_{10}H_{10}ClN_3S$, 240.0; found, 239.9; HPLC (method 2): t_R = 9.091 min.

N,5-Dimethyl-4-(2-((4-morpholinophenyl)amino)pyridin-4-yl)thiazol-2-amine (78). The title compound was synthesized according to general procedure A starting from **77** (150 mg, 0.62 mmol) and 4-morpholinoaniline (165.8 mg, 0.93 mmol). Purification by flash column chromatography (SiO$_2$, DCM/EtOH 95:05) afforded 42 mg of the desired compound (17% yield); ^{1}H NMR (250 MHz, DMSO-d_6): δ 2.40 (s, 3H), 2.82 (d, J = 4.9 Hz, 3H), 2.95−3.07 (m, 4H), 3.67−3.80 (m, 4H), 6.81−6.93 (m, 3H), 7.01 (br s, 1H), 7.34 (q, J = 4.6 Hz,

1H), 7.46−7.58 (m, 2H), 8.07 (d, *J* = 5.4 Hz, 1H), 8.78 ppm (s, 1H); ^{13}C NMR (101 MHz, DMSO-d_6): δ 12.3, 30.6, 49.5, 66.2, 108.2, 112.6, 115.9, 117.3, 119.7, 134.4, 143.3, 143.3, 145.5, 147.2, 156.6, 165.4 ppm; MS-ESI *m/z*: $[M + H]^+$ calcd for $C_{20}H_{23}N_5OS$, 382.2; found, 382.2; *m/z*: $[M − H]^-$ calcd for $C_{20}H_{23}N_5OS$, 380.2; found, 380.2; HPLC (method 2): t_R = 5.343 (99.4%).

■ ASSOCIATED CONTENT

ⓢ Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acsomega.8b00668.

> Detailed procedures for the preparation of intermediates **17a−l** and **33−35**; crystal structure of compound **49**; melting curves and method description of the nanoDSF experiments; experimental procedures for the crystallization and structure determination of JNK3 in the complex with compounds **38**, **44**, and AMP−PCP; selectivity screening of **44**; and in vitro metabolic stability study for compound **44** (PDF)
>
> SMILES strings of tested compounds (CSV)

Accession Codes

The atomic coordinates and structures factors of complex structures containing compounds **38** and **44** and AMP−PCP were deposited in the Protein Data Bank (PDB) with the respective accession codes 6EMH, 6EKD, and 6EQ9, respectively. The authors will release the atomic coordinates and experimental data upon article publication.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: pierre.koch@uni-tuebingen.de. Phone: +49 7071 2974579 (P.K.).

ORCID ⓞ

Stefan A. Laufer: 0000-0001-6952-1486
Pierre Koch: 0000-0003-4620-4650

Present Addresses

§Chemistry Department, Faculty of Science, Menofia University, Gamal Abdel-Nasser Street, 32511 Shebin El-Koum, Menofia, Egypt.

‖Department of Applied Sciences, Chemistry Branch, University of Technology-Baghdad, Sinaa' Street, 10066, Baghdad, Iraq.

Notes

The authors declare no competing financial interest.

■ ABBREVIATIONS

MAPK, mitogen activated protein kinase; JNK, c-Jun N terminal kinase; TOSMIC, toluene sulfonylmethylisocyanide; HR, hydrophobic region; AMP−PCP, β,γ-methyleneadenosine-5′-triphosphate; nanoDSF, nano differential scanning fluorimetry; hERG, human-ether-à-go-go related gene; CYP450, cytochrome P450; TLC, thin layer chromatography; HPLC, high performance liquid chromatography; DAD, diode array detector; NMR, nuclear magnetic resonance; ESI-MS, electrospray ionization mass spectrometer

■ REFERENCES

(1) Margutti, S.; Laufer, S. A. Are MAP kinases drug targets? Yes, but difficult ones. *ChemMedChem* **2007**, *2*, 1116−1140.

(2) Barr, R. K.; Bogoyevitch, M. A. The c-Jun N-terminal protein kinase family of mitogen-activated protein kinases (JNK MAPKs). *Int. J. Biochem. Cell Biol.* **2001**, *33*, 1047−1063.

(3) Davis, R. J. Signal transduction by the JNK group of MAP kinases. *Cell* **2000**, *103*, 239−252.

(4) Bogoyevitch, M. A. The isoform-specific functions of the c-Jun N-terminal kinases (JNKs): differences revealed by gene targeting. *BioEssays* **2006**, *28*, 923−934.

(5) Graczyk, P. P. JNK inhibitors as anti-inflammatory and neuroprotective agents. *Future Med. Chem.* **2013**, *5*, 539−551.

(6) Resnick, L.; Fennell, M. Targeting JNK3 for the treatment of neurodegenerative disorders. *Drug Discovery Today* **2004**, *9*, 932−939.

(7) Manning, A. M.; Davis, R. J. Targeting JNK for therapeutic benefit: from junk to gold? *Nat. Rev. Drug Discovery* **2003**, *2*, 554−565.

(8) Ebelt, N. D.; Cantrell, M. A.; Van Den Berg, C. L. c-Jun N-Terminal Kinases Mediate a Wide Range of Targets in the Metastatic Cascade. *Genes Cancer* **2013**, *4*, 378−387.

(9) Wagner, E. F.; Nebreda, Á. R. Signal integration by JNK and p38 MAPK pathways in cancer development. *Nat. Rev. Cancer* **2009**, *9*, 537−549.

(10) Gehringer, M.; Muth, F.; Koch, P.; Laufer, S. A. c-JunN-terminal kinase inhibitors: a patent review (2010 - 2014). *Expert Opin. Ther. Pat.* **2015**, *25*, 849−872.

(11) Koch, P.; Gehringer, M.; Laufer, S. A. Inhibitors of c-Jun N-terminal kinases: an update. *J. Med. Chem.* **2015**, *58*, 72−95.

(12) Siddiqui, M. A.; Reddy, P. A. Small Molecule JNK (c-Jun N-Terminal Kinase) Inhibitors. *J. Med. Chem.* **2010**, *53*, 3005−3012.

(13) Cuenda, A.; Rousseau, S. p38 MAP-kinases pathway regulation, function and role in human diseases. *Biochim. Biophys. Acta, Mol. Cell Res.* **2007**, *1773*, 1358−1375.

(14) Corrêa, S. A. L.; Eales, K. L. The Role of p38 MAPK and Its Substrates in Neuronal Plasticity and Neurodegenerative Disease. *J. Signal Transduction* **2012**, *2012*, 1.

(15) Kumar, S.; Boehm, J.; Lee, J. C. p38 MAP kinases: key signalling molecules as therapeutic targets for inflammatory diseases. *Nat. Rev. Drug Discovery* **2003**, *2*, 717−726.

(16) Muth, F.; Günther, M.; Bauer, S. M.; Döring, E.; Fischer, S.; Maier, J.; Drückes, P.; Köppler, J.; Trappe, J.; Rothbauer, U.; Koch, P.; Laufer, S. A. Tetra-Substituted Pyridinylimidazoles As Dual Inhibitors of p38α Mitogen-Activated Protein Kinase and c-Jun N-Terminal Kinase 3 for Potential Treatment of Neurodegenerative Diseases. *J. Med. Chem.* **2015**, *58*, 443−456.

(17) Genovese, M. C. Inhibition of p38: Has the fat lady sung? *Arthritis Rheum.* **2009**, *60*, 317−320.

(18) Koch, P.; Ansideri, F. 2-Alkylsufanyl-4(5)-aryl-5(4)-heteroarylimidazoles: An Overview on Synthetic Strategies and Biological Activity. *Arch. Pharm.* **2017**, *350*, 1700258.

(19) Ansideri, F.; Lange, A.; El-Gokha, A.; Boeckler, F. M.; Koch, P. Fluorescence polarization-based assays for detecting compounds binding to inactive c-Jun N-terminal kinase 3 and p38α mitogen-activated protein kinase. *Anal. Biochem.* **2016**, *503*, 28−40.

(20) Ansideri, F.; Dammann, M.; Boeckler, F. M.; Koch, P. Fluorescence polarization-based competition binding assay for c-Jun N-terminal kinases 1 and 2. *Anal. Biochem.* **2017**, *532*, 26−28.

(21) Muth, F.; El-Gokha, A.; Ansideri, F.; Eitel, M.; Döring, E.; Sievers-Engler, A.; Lange, A.; Boeckler, F. M.; Lämmerhofer, M.; Koch, P.; Laufer, S. A. Tri- and Tetrasubstituted Pyridinylimidazoles as Covalent Inhibitors of c-Jun N-Terminal Kinase 3. *J. Med. Chem.* **2017**, *60*, 594−607.

(22) Marckwald, W. Ein Beitrag zur Kenntniss der Imidazole und der Constitution des Glyoxalins. *Ber. Dtsch. Chem. Ges.* **1892**, *25*, 2354−2373.

(23) Laufer, S. A.; Wagner, G. K.; Kotschenreuther, D. A.; Albrecht, W. Novel substituted pyridinyl imidazoles as potent anticytokine agents with low activity against hepatic cytochrome P450 enzymes. *J. Med. Chem.* **2003**, *46*, 3230−3244.

(24) Koch, P.; Laufer, S. Unexpected Reaction of 2-Alkylsulfanyli-midazoles to Imidazol-2-ones: Pyridinylimidazol-2-ones as Novel Potent p38α Mitogen-Activated Protein Kinase Inhibitors. *J. Med. Chem.* **2010**, *53*, 4798−4802.

(25) Radziszewski, B. Ueber Glyoxalin und seine Homologe. *Ber. Dtsch. Chem. Ges.* **1882**, *15*, 2706−2708.

(26) Neber, P. W.; Friedolsheim, A. V. Über eine neue Art der Umlagerung von Oximen. *Justus Liebigs Ann. Chem.* **1926**, *449*, 109−134.

(27) Xi, N.; Xu, S.; Cheng, Y.; Tasker, A. S.; Hungate, R. W.; Reider, P. J. Regio-controlled synthesis of N-substituted imidazoles. *Tetrahedron Lett.* **2005**, *46*, 7315−7319.

(28) Ansideri, F.; Andreev, S.; Kuhn, A.; Albrecht, W.; Laufer, S.; Koch, P. A Diverse and Versatile Regiospecific Synthesis of Tetrasubstituted Alkylsulfanylimidazoles as p38α Mitogen-Activated Protein Kinase Inhibitors. *Molecules* **2018**, *23*, 221.

(29) Van Leusen, A. M.; Wildeman, J.; Oldenziel, O. H. Chemistry of sulfonylmethyl isocyanides. 12. Base-induced cycloaddition of sulfonylmethyl isocyanides to carbon,nitrogen double bonds. Synthesis of 1,5-disubstituted and 1,4,5-trisubstituted imidazoles from aldimines and imidoyl chlorides. *J. Org. Chem.* **1977**, *42*, 1153−1159.

(30) Miyaura, N.; Suzuki, A. Palladium-Catalyzed Cross-Coupling Reactions of Organoboron Compounds. *Chem. Rev.* **1995**, *95*, 2457−2483.

(31) Voss, M. E.; Beer, C. M.; Mitchell, S. A.; Blomgren, P. A.; Zhichkin, P. E. A simple and convenient one-pot method for the preparation of heteroaryl-2-imidazoles from nitriles. *Tetrahedron* **2008**, *64*, 645−651.

(32) Hantzsch, A.; Weber, J. H. Ueber Verbindungen des Thiazols (Pyridins der Thiophenreihe). *Ber. Dtsch. Chem. Ges.* **1887**, *20*, 3118−3132.

(33) Goettert, M.; Graeser, R.; Laufer, S. A. Optimization of a nonradioactive immunosorbent assay for p38α mitogen-activated protein kinase activity. *Anal. Biochem.* **2010**, *406*, 233−234.

(34) Goettert, M.; Luik, S.; Graeser, R.; Laufer, S. A. A direct ELISA assay for quantitative determination of the inhibitory potency of small molecules inhibitors for JNK3. *J. Pharm. Biomed. Anal.* **2011**, *55*, 236−240.

(35) Sterling, T.; Irwin, J. J. ZINC 15 - Ligand Discovery for Everyone. *J. Chem. Inf. Model.* **2015**, *55*, 2324−2337.

(36) He, Y.; Duckett, D.; Chen, W.; Ling, Y. Y.; Cameron, M. D.; Lin, L.; Ruiz, C. H.; LoGrasso, P. V.; Kamenecka, T. M.; Koenig, M. Synthesis and SAR of novel isoxazoles as potent c-jun N-terminal kinase (JNK) inhibitors. *Bioorg. Med. Chem. Lett.* **2014**, *24*, 161−164.

(37) Scapin, G.; Patel, S. B.; Lisnock, J. M.; Becker, J. W.; LoGrasso, P. V. The Structure of JNK3 in Complex with Small Molecule Inhibitors. *Chem. Biol.* **2003**, *10*, 705−712.

(38) Thaher, B. A.; Koch, P.; Schattel, V.; Laufer, S. Role of the Hydrogen Bonding Heteroatom−Lys53 Interaction between the p38α Mitogen-Activated Protein (MAP) Kinase and Pyridinyl-Substituted 5-Membered Heterocyclic Ring Inhibitors. *J. Med. Chem.* **2009**, *52*, 2613−2617.

(39) Wagner, G. K.; Kotschenreuther, D.; Zimmermann, W.; Laufer, S. A. Identification of Regioisomers in a Series of N-Substituted Pyridin-4-yl Imidazole Derivatives by Regiospecific Synthesis, GC/MS, and1H NMR. *J. Org. Chem.* **2003**, *68*, 4527−4530.

(40) Toledo, L. M.; Lydon, N. B.; Elbaum, D. The structure-based design of ATP-site directed protein kinase inhibitors. *Curr. Med. Chem.* **1999**, *6*, 775−805.

(41) Kamenecka, T.; Habel, J.; Duckett, D.; Chen, W.; Ling, Y. Y.; Frackowiak, B.; Jiang, R.; Shin, Y.; Song, X.; LoGrasso, P. Structure-activity relationships and X-ray structures describing the selectivity of aminopyrazole inhibitors for c-Jun N-terminal kinase 3 (JNK3) over p38. *J. Biol. Chem.* **2009**, *284*, 12853−12861.

(42) Probst, G. D.; Bowers, S.; Sealy, J. M.; Truong, A. P.; Hom, R. K.; Galemmo, R. A.; Konradi, A. W.; Sham, H. L.; Quincy, D. A.; Pan, H.; Yao, N.; Lin, M.; Tóth, G.; Artis, D. R.; Zmolek, W.; Wong, K.; Qin, A.; Lorentzen, C.; Nakamura, D. F.; Quinn, K. P.; Sauer, J.-M.; Powell, K.; Ruslim, L.; Wright, S.; Chereau, D.; Ren, Z.; Anderson, J. P.; Bard, F.; Yednock, T. A.; Griswold-Prenner, I. Highly selective c-Jun N-terminal kinase (JNK) 2 and 3 inhibitors with in vitro CNS-like pharmacokinetic properties prevent neurodegeneration. *Bioorg. Med. Chem. Lett.* **2011**, *21*, 315−319.

(43) Swahn, B.-M.; Xue, Y.; Arzel, E.; Kallin, E.; Magnus, A.; Plobeck, N.; Viklund, J. Design and synthesis of 2′-anilino-4,4′-bipyridines as selective inhibitors of c-Jun N-terminal kinase-3. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 1397−1401.

(44) Kamenecka, T.; Jiang, R.; Song, X.; Duckett, D.; Chen, W.; Ling, Y. Y.; Habel, J.; Laughlin, J. D.; Chambers, J.; Figuera-Losada, M.; Cameron, M. D.; Lin, L.; Ruiz, C. H.; LoGrasso, P. V. Synthesis, Biological Evaluation, X-ray Structure, and Pharmacokinetics of Aminopyrimidine c-jun-N-terminal Kinase (JNK) Inhibitors. *J. Med. Chem.* **2010**, *53*, 419−431.

(45) Laufer, S. A.; Hauser, D. R. J.; Domeyer, D. M.; Kinkel, K.; Liedtke, A. J. Design, synthesis, and biological evaluation of novel Tri- and tetrasubstituted imidazoles as highly potent and specific ATP-mimetic inhibitors of p38 MAP kinase: focus on optimized interactions with the enzyme's surface-exposed front region. *J. Med. Chem.* **2008**, *51*, 4122−4149.

(46) Fricker, M.; LoGrasso, P.; Ellis, S.; Wilkie, N.; Hunt, P.; Pollack, S. J. Substituting c-Jun N-terminal kinase-3 (JNK3) ATP-binding site amino acid residues with their p38 counterparts affects binding of JNK- and p38-selective inhibitors. *Arch. Biochem. Biophys.* **2005**, *438*, 195−205.

(47) Krenitsky, V. P.; Nadolny, L.; Delgado, M.; Ayala, L.; Clareen, S. S.; Hilgraf, R.; Albers, R.; Hegde, S.; D'Sidocky, N.; Sapienza, J.; Wright, J.; McCarrick, M.; Bahmanyar, S.; Chamberlain, P.; Delker, S. L.; Muir, J.; Giegel, D.; Xu, L.; Celeridad, M.; Lachowitzer, J.; Bennett, B.; Moghaddam, M.; Khatsenko, O.; Katz, J.; Fan, R.; Bai, A.; Tang, Y.; Shirley, M. A.; Benish, B.; Bodine, T.; Blease, K.; Raymon, H.; Cathers, B. E.; Satoh, Y. Discovery of CC-930, an orally active anti-fibrotic JNK inhibitor. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 1433−1438.

(48) Heider, F.; Haun, U.; Döring, E.; Kudolo, M.; Sessler, C.; Albrecht, W.; Laufer, S.; Koch, P. From 2-Alkylsulfanylimidazoles to 2-Alkylimidazoles: An Approach towards Metabolically More Stable p38α MAP Kinase Inhibitors. *Molecules* **2017**, *22*, 1729.

(49) Koch, P.; Bäuerlein, C.; Jank, H.; Laufer, S. Targeting the ribose and phosphate binding site of p38 mitogen-activated protein (MAP) kinase: Synthesis and biological testing of 2-alkylsulfanyl-, 4(5)-aryl-, 5(4)-heteroaryl-substituted imidazoles. *J. Med. Chem.* **2008**, *51*, 5630−5640.

(50) Koch, P.; Jahns, H.; Schattel, V.; Goettert, M.; Laufer, S. Pyridinylquinoxalines and Pyridinylpyridopyrazines as Lead Compounds for Novel p38α Mitogen-Activated Protein Kinase Inhibitors. *J. Med. Chem.* **2010**, *53*, 1128−1137.

(51) Revesz, L.; Blum, E.; Di Padova, F. E.; Buhl, R.; Feifel, R.; Gram, H.; Hiestand, P.; Manning, U.; Rucklin, G. Novel p38 inhibitors with potent oral efficacy in several models of rheumatoid arthritis. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 3595−3599.