

University of Tartu
FACULTY OF SCIENCE AND TECHNOLOGY
Institute of Chemistry

Katrin Kalind

**Targeting Both Substrate-Binding Sites of Mitotic Haspin
Kinase with a Single Inhibitor**

Master Thesis

Supervisors: Darja Lavõgina, *PhD*

Asko Uri, *PhD*

Tartu 2014

TABLE OF CONTENTS

ABBREVIATIONS.....	3
1. INTRODUCTION.....	4
2. LITERATURE SURVEY	5
2.1. Mitosis	5
2.1.1. General Overview: the Cell Cycle and Mitosis	5
2.1.2. Mitotic Protein Kinases.....	6
2.2. Haspin.....	6
2.2.1. Crystal Structures of Catalytical Domains of PKA and Haspin	7
2.3. Inhibitors of Enzymes.....	9
2.3.1. General Characteristics	9
2.3.2. Inhibitors of PKs	9
2.3.2.1. ATP-Site Binding Inhibitors	9
2.3.2.2. Substrate-Site Binding Inhibitors	10
2.3.2.3. Bisubstrate Inhibitors	10
3. REAGENTS, EQUIPMENT AND METHODS	12
3.1. Reagents and Equipment	12
3.1.1. Reagents	12
3.1.2. Equipment	12
3.2. Methods	13
3.2.1. Solid Phase Peptide Synthesis	13
3.2.2. Fluorescence Anisotropy-Based Binding/Displacement Assay (FA).....	14
3.2.3. Thermal Shift Assay	15
3.2.4. Protein Crystallography	16
3.2.4.1. General Features of Crystallography.....	16
3.2.4.2. Principles of Protein Crystallography	16
4. EXPERIMENTAL PART	19
4.1. Production of Proteins	19
4.1.1. Expression.....	19
4.1.2. Preparation of Lysate	19
4.1.3. Purification.....	19
4.1.4. SDS-PAGE	20
4.2. Synthesis of ARCs.....	20
4.3. Biochemical Measurements.....	21
4.3.1. FA-method	21
4.3.2. Thermal Shift Assay	21
4.4. Protein Crystallography.....	22
4.4.1. Crystallization	22
4.4.2. Preparations for the Diffraction Data Measurement and Data Processing	22
5. RESULTS AND DISCUSSION	23
5.1. General Description of the Results	23
5.2. Screening of the Initial Set of ARCs	23
5.3. Co-crystallization of Haspin and ARCs	26
5.4. Analysis of Crystal Structures	27
5.5. Synthesis and Biochemical Characterization of Novel ARCs.....	28
5.6. Prospectives	34
6. SUMMARY	35
7. KOKKUVÔTE.....	36
8. ACKNOWLEDGEMENTS	37
9. REFERENCES.....	38
10. APPENDICES.....	41

ABBREVIATIONS

AA	amino acid residue
Adc	adenosine-4'-dehydroxymethyl-4'-carboxylic acid moiety
Ahx	6-aminohexanoic acid residue
AMP	adenosine-3',5'-monophosphate
AMTH	5-(2-aminopyrimidin-4-yl)-thiophene-2-carboxylic acid moiety
ARC	adenosine analogue-oligoarginine conjugate
ATP	adenosine-5'-triphosphate
Boc	<i>tert</i> -butoxycarbonyl
DCE	1,2-dichloroethane
DIPEA	<i>N,N</i> -diisopropylethylamine
DMF	<i>N,N</i> -dimethylformamide
DMSO	dimethylsulfoxide
DSF	differential scanning fluorimetry
DTT	dithiothreitol
FA	fluorescence anisotropy
Fmoc	9-fluorenylmethoxycarbonyl
H3	histone H3
H9	<i>N</i> -aminoethyl-5-isoquinolinesulfonamide
Haspin	haploid germ cell-specific nuclear protein kinase
HBTU	2-(1 <i>H</i> -benzotriazol-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate
HOBt	1-hydroxybenzotriazole
HPLC	high-performance liquid chromatography
Ida	2,2'-iminodiacetic acid moiety
IpAdc	2',3'-O-isopropylidene-adenosine-4'-dehydroxymethyl-4'-carboxylic acid
ivDde	1-(4,4-dimethyl-2,6-dioxocyclohex-1-ylidene)isovaleryl
LB	Luria-Bertani medium
LC-MS	liquid chromatography with detection by mass spectrometry
Ni-NTA	nickel-nitrilotriacetic acid
NMM	<i>N</i> -methylmorpholine
PK	protein kinase
PKA	cAMP-dependent protein kinase
PKAc	cAMP-dependent protein kinase catalytic subunit, type α
PKI	natural heat stable protein kinase inhibitor
SAC	spindle assembly checkpoint
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SPPS	solid phase peptide synthesis
TCEP	tris(2-carboxyethyl)phosphine
TFA	trifluoroacetic acid
TIPS	triisopropylsilane
v	volume percent

1. INTRODUCTION

Protein kinases are the enzymes belonging to the class of phosphotransferases that catalyse the phosphoryl transfer from donor molecule (usually ATP [1]) to hydroxyl group in the side-chain of Ser/Thr or Tyr of substrate protein [2]. Phosphorylation functions as a molecular “switch” changing 3D-structure of the substrate protein and in this way activating or turning off the signalling pathways [3,4]. For instance, the crosstalk of different protein kinases is crucial for the correct regulation of the cell cycle, including cell division (mitosis) [5,6]. One of the protein kinases involved in mitosis is the recently discovered Haspin, which is structurally different from most other eukaryotic protein kinases, rendering it interesting as a potential target for development of selective inhibitors [6,7].

The protein X-ray crystallography has become an important tool for investigation of functioning and mechanism of action of the macromolecules, as it can provide high resolution to an atomic level. In biochemistry and pharmacology, the three-dimensional crystal structures of protein/inhibitor complexes can give information about the interactions between the protein and inhibitor, and thereby contribute to the design of more affine and selective compounds. For example, one of the strategies applied for increasing the selectivity and affinity of inhibitors takes advantage of sequences of natural substrates, which can be synthetically conjugated with small molecules targeting the ATP-site of kinase. Such efforts have already yielded several bisubstrate inhibitors, including adenosine analogue-oligoarginine conjugates (ARCs) that have been used as generic or selective bisubstrate scaffolds for targeting a variety of protein kinases [8-11]. Still, in ARCs the variation of peptidic part for achievement of higher selectivity towards the biological targets has been left relatively unexplored.

The aim of this work are as follows: 1) the production and purification of Haspin, 2) screening of the initial set of ARCs representing variable structures towards Haspin, 3) crystallization of Haspin/ARC complexes and analysis of the obtained co-crystal structures, and 4) synthesis and biochemical characterization of novel Haspin-selective ARCs based on crystallographic data.

2. LITERATURE SURVEY

2.1. Mitosis

2.1.1. General Overview: the Cell Cycle and Mitosis

The cell cycle is a period of an existence of a cell starting with the cell formation from mother cell and ending with the cell division or death [12]. The cell cycle of eukaryotic cells is divided into 4 phases: G_1 (gap 1), S (synthesis), G_2 (gap 2), and M (mitosis) (Figure 1). G_1 , S, and G_2 together are known as interphase (Figure 1) that prepares the cell for the next division: in G_1 , cell growth and preparation for DNA synthesis occurs; in S phase, DNA replication and centrosome duplication takes place; in G_2 , the cell continues to grow and protein synthesis is enhanced. In G_0 , the cells are in the resting state, but they can still re-enter G_1 [12].

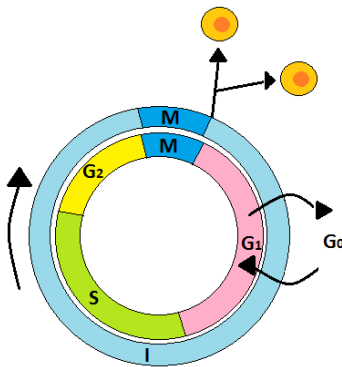


Figure 1. Phases of the cell cycle. “I” stands for interphase and “M” for mitosis.

Mitosis also consists of 4 phases: prophase, metaphase, anaphase, and telophase plus cytokinesis (Figure 2). In prophase, chromosomes condense in nuclear membrane and become visible through a transmitted light microscope; the two centrosomes move to opposite poles of the cell, and a mitotic spindle forms. The beginning of the metaphase is set by the loss of nuclear membrane, which allows some of the microtubules of mitotic spindle to attach to the binding site of sister chromatids. Subsequently, the spindle aligns the chromosomes at the middle plane of the cell. In anaphase, the sister chromatids are separated and moved towards the opposite poles. The last events of mitosis comprise the formation of the cleavage furrow in the cell membrane, the re-formation of nuclear envelopes around the separated sister chromatids, and the cell division into two (telophase and cytokinesis) [13,14].

A successful cell division cycle requires precise control by checkpoints, which act through the signalling pathways of proteins. The control mechanism is responsible for ensuring that the errors that may arise during the cycle are corrected, or the cells go to programmed cell death (apoptosis) [12,15]. The cell cycle has three major checkpoints. Before the cell enters to S phase, it must undergo G_1/S checkpoint where it is ensured that the cellular ‘machinery’ is prepared for DNA synthesis (duplication) and DNA is not damaged. G_2/M checkpoint locates before M phase and controls if the replication of DNA was successful and everything is ready for mitosis. During mitosis, in metaphase, a spindle assembly checkpoint (SAC) controls

further progression of cell division by ensuring that chromosome segregation is correct [12]. The malfunctioned regulation of the cell division leads to genetic damage and uncontrollable cell division, and thus to different diseases (e.g., cancer, Alzheimer's disease) [12,15].

2.1.2. Mitotic Protein Kinases

Protein kinases (PKs) are enzymes that catalyse phosphorylation of proteins, a reaction during which the phosphoryl group is transferred from ATP to Ser/Thr/Tyr residues of the substrate. As a result, negative charges are added to the substrate protein causing changes in its conformation; thus, phosphorylation can act as a switch turning the cellular signalling pathway “ON” or “OFF”. The mitotic events are regulated by crosstalk of different PKs, including cyclin-dependent PKs (Cdks), Polo-like PKs (Plks), Aurora PKs, and Never-in-mitosis-A-related PKs (NIMAs). The new mitotic players that have been recently discovered include Haspin and Greatwall (MAST-L) (Figure 2) [5,6]. Importantly, the elevated levels of mitotic PKs can serve as cancer biomarkers, since the cancer cells have a faster life cycle and divide uncontrollably. Thus, the biochemical tools enabling quantification of mitotic PKs and down-regulation of their activity are of great value for the cancer diagnosis and treatment [6].

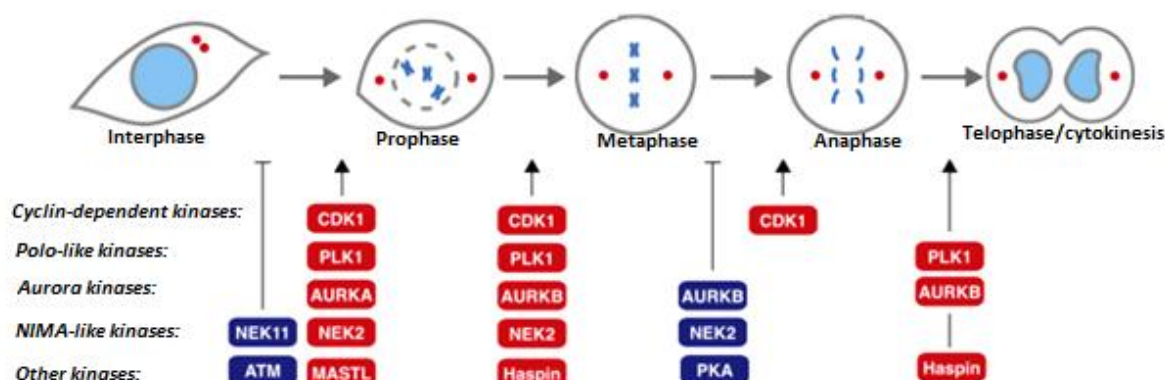


Figure 2. Regulation of mitosis by PKs [16, modified].

2.2. Haspin

Haploid germ-cell-specific nuclear protein kinase (Haspin) is a Ser/Thr PK that participates in regulation of chromosome behavior during the cell division. Haspin is encoded by the germ-cell specific gene-2 [17] and found in all proliferating somatic cells, though its expression is highest in testis [7,18]. As Haspin lacks some of the conserved structural fragments generally necessary for catalysis (discussed below) and has low sequence homology with other eukaryotic PKs, it was initially thought that Haspin is an inactive pseudokinase. However, it has been recently demonstrated that Haspin is catalytically active, although the only known substrate of Haspin known to date is histone H3 [7,19,20]. Histones are important for the “packing” of DNA, and serve as substrates for a number of mitotic PKs. The depletion of Haspin leads to misalignment of chromosomes in metaphase and, consequently, the activation

of SAC and delay in exiting the mitosis. In case of Haspin overexpression, abnormal dissociation of sister chromosomes occurs [7].

2.2.1. Crystal Structures of Catalytical Domains of PKA and Haspin

Similarly to other PKs of eukaryotes, Haspin has a bilobal structure consisting of a small N-lobe and a large C-lobe. However, Haspin has a number of structural features and specific inserts that altogether form an atypical catalytic domain (Figure 3). For example, the highly conserved DFG motif of activation segment (Asp-Phe-Gly; required for ATP/Mg²⁺ binding) of PKs is replaced in Haspin by DYT (Asp-Tyr-Thr), and the APE motif is absent [19]. To illustrate the differences between Haspin and a typical PK, the crystal structures of Haspin are compared with the crystal structures of a well-known basophilic PK represented by the catalytic subunit of cAMP-dependent protein kinase (PKAc).

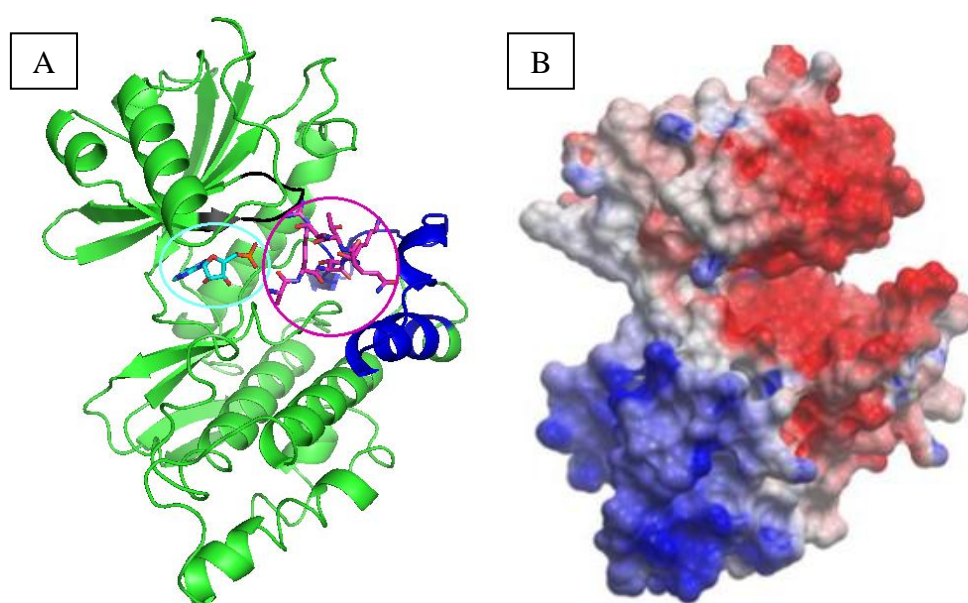


Figure 3. A. Overlay of co-crystal structures of Haspin kinase domain with AMP (PDB 3DLZ¹) and histone H3(1-7) (PDB 4OUC). Kinase is shown as cartoon and small molecules as sticks; ATP-site and substrate-site are surrounded by light blue and magenta circle, respectively. The glycine-rich loop is marked with black and activation segment with dark blue colour. **B.** Electrostatic surface potential of Haspin kinase domain [7]. Electronegative and electropositive areas are shown with red and blue colour, respectively.

As for all kinases, the ATP-site in Haspin is a relatively narrow pocket between the two lobes. No co-crystal of Haspin with ATP is available so far; still, in co-crystal structure of AMP with Haspin, amino groups of 1N and 6N are forming hydrogen bonds with *Glu606* and *Gly608* (*Glu121* and *Val123* in PKAc), similarly to ATP binding to PKAc (Figure 4). Both hydroxyl groups of ribose moiety give polar contacts with *Asp611* (*Glu127* in PKAc), and 3'-hydroxyl additionally with *Gly653* (*Glu170* in PKAc). The α -phosphate of nucleotide is coordinated by *Lys511* (*Lys72* in PKAc) [19]. In Ser/Thr PKs, γ -phosphate usually forms a charge-reinforced hydrogen bond with *Lys* (*Lys168* in PKAc), which is required for phosphoryl transfer; however, *His651* performs the same role in Haspin [21]. The phosphate ion in AMP/Haspin

¹ All PDB files are available on Protein Databank Homepage, <http://www.rcsb.org/> (last viewed May 21, 2014).

co-crystal which imitates γ -phosphate of ATP develops hydrogen bonds with **Asp687** (DYT motif) and **Asp649** (HRD motif) (**Ser53** and **Ala21** in PKAc) [19]. Unlike other PKs, the glycine-rich loop (conserved glycines in Haspin **491**, **493**, **496** and PKAc **50**, **52**, **55**) which covers the ATP-site is relatively firmly fixed in Haspin by the surrounding segments (Figure 3) [19,22]. In addition to polar contacts, the adenine ring system of ATP forms hydrophobic and π - π interactions with aromatic amino acid residues (AAs) of Haspin (e.g., **Phe607**, **Tyr688**, **Trp652**) [19]. The K_M value of ATP for Haspin is 200 μ M [23].

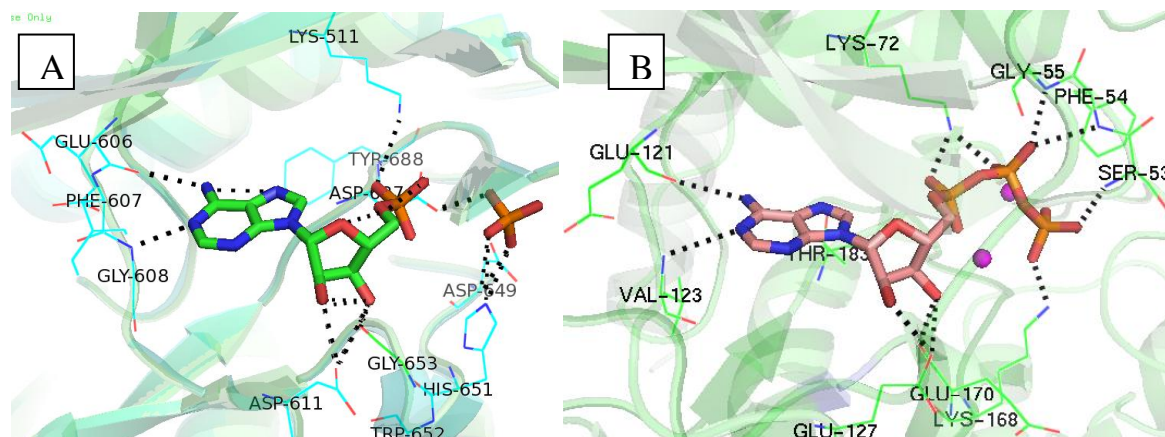


Figure 4. Co-crystal structures of Haspin with AMP (A; PDB 3IQ7, 3DLZ; [11]) and PKAc with ATP (B; PDB 1ATP; [16]). PKs are shown as cartoons; AMP, ATP and phosphate as sticks; residues of PKs forming interactions with co-crystallized small molecules are shown as lines and are labelled; hydrogen bonds are shown as black dotted lines; Mn^{2+} -ions are shown as magenta circles (no electron density can be observed for Mg^{2+} -ion(s) near ATP-site in AMP/Haspin co-crystal).

The substrate-binding site of Haspin is located on the surface of the C-lobe and formed by electronegative residues constituting a pocket suitable for binding of basic N-terminal peptide moiety of histone H3 (Figure 3) [7,19,20]. The K_M values of H3 (residues 1-21, sequence: ARTKQTARKSTGGKAPRKQLA) towards full-length and kinase domain of Haspin are 0.058 μ M and 0.35 μ M, respectively [19]. The co-crystal structure of N-terminal peptide of H3(1-7) showed that **Ala1** (hydrogen bond with **Glu613**) and **Thr3** (hydrogen bond with **Asp649** and **Gln718**) are directed into the substrate binding pocket (Figure 5). Further, **Arg2** is positioned to the relatively hydrophobic pocket of the N-terminal lobe formed between the two loops of the kinase (**Val494** from the Gly-rich loop and **Ala587** from the loop preceding α -uIH helix); still, guanidine group and the carboxyl oxygen of **Arg2** make hydrogen bonds with **Asp588** and **Gln718**, respectively. Unlike the substrates of other PKs (e.g., PKI(5-24) as an analogue of substrate of PKAc, Figure 5), the peptide makes a sharp turn at **Lys4** (ca 180°). This results in positioning of **Lys4** into pocket lined by kinase residues that develop hydrophobic interactions (**Leu690**, **Val704**, **Leu710**) as well as charge-reinforced hydrogen bonds (**Asp707**, **Asp709**) to the alkyl chain and amine group of **Lys4**, respectively. Next, **Gln5** forms hydrogen bond with **Asp714** and **Thr6** flips back over **Arg2** [24].

The majority of basophilic PKs require (auto)phosphorylation of their activation loop for full activation, as the non-phosphorylated activation loop binds to the substrate-binding site and

prevents protein substrates from binding. However, Haspin kinase is constitutively active, and no phosphorylation of activation segment is needed for its activity. Still, as the expression level of Haspin remains nearly constant all over the cycle, but its activity peaks at mitosis, it is considered that Haspin may be regulated by binding of regulatory proteins or by modifications of its N-lobe [7,19,25]. Indeed, it has been found that in mitosis, the N-terminus of Haspin is abundantly phosphorylated by Plk [7,25].

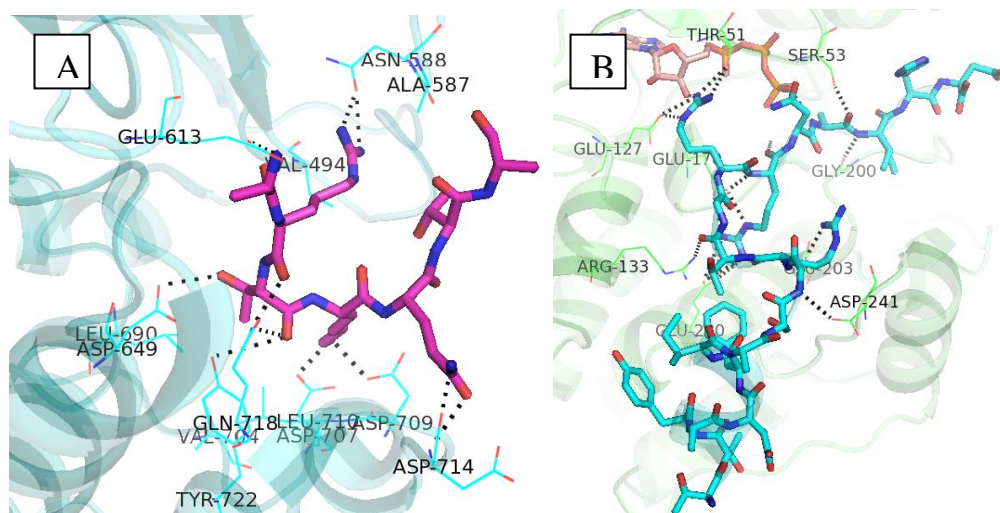


Figure 5. **A.** Co-crystal structure of Haspin/histone H3(1-7) (PDB 4OUC). **B.** Co-crystal structure of PKAc/AMP-PNP/PKI(5-24) (PDB 1ATP). PKs are shown as cartoons; H3(1-7) and PKI(5-24) as sticks; residues of PKs forming interactions with the co-crystallized peptide are labelled and shown as lines; hydrogen bonds are shown as black dotted lines.

2.3. Inhibitors of Enzymes

2.3.1. General Characteristics

An enzyme inhibitor is a compound that binds to an enzyme and thereby prevents binding of (co-)substrates, disabling enzymatic catalysis of a chemical reaction. Inhibitors of enzymes are used *in vitro* as well as *in vivo* systems for reduction the activity of enzymes; additionally, inhibitors might serve as templates for the design of enzyme-targeting probes that can be applied for characterization of different enzymes and/or quantification of their amounts in biochemical assays and in the natural milieu.

2.3.2. Inhibitors of PKs

For the transfer of phosphoryl group conducted by a PK, binding of both ATP and substrate protein to the PK is required. The PK-catalysed phosphorylation reaction can thus be directly blocked by compounds targeting the ATP- or/and the substrate-binding pocket of PK. Additionally, allosteric inhibitors exist that act *via* binding to PK pockets outside the catalytic core and interfere with PK activity by indirect disruption of 3D-structure of PK.

2.3.2.1. ATP-Site Binding Inhibitors

Most of the inhibitors of PKs bind to an ATP-binding pocket located in the active site of PK. For intracellular inhibitory potency, ATP-site binding compounds must have a sufficient affinity to compete with the high intracellular concentration of ATP (2-10 mM) [26]. Another

possible disadvantage is the lack of selectivity, as in addition to PKs there are more than 1,500 ATP-binding proteins in cell [27]. To date, it has been established that the amino acid sequence of ATP-binding pocket may vary for different PKs; thus, the selectivity of inhibitors for a specific kinase or kinase family can be substantially increased [28].

The first potential inhibitor identified for Haspin was 5-iodotubercidin (IC_{50} value of 9 nM at 250 μ M concentration of ATP [29], Figure 6), which was first identified as an inhibitor of adenosine kinase (IC_{50} =26 nM) [30]. Recently, beta-carboline derivatives and acridine analogues have also been developed as inhibitors targeting the ATP-site of Haspin and possessing moderate activity and selectivity [31,32].

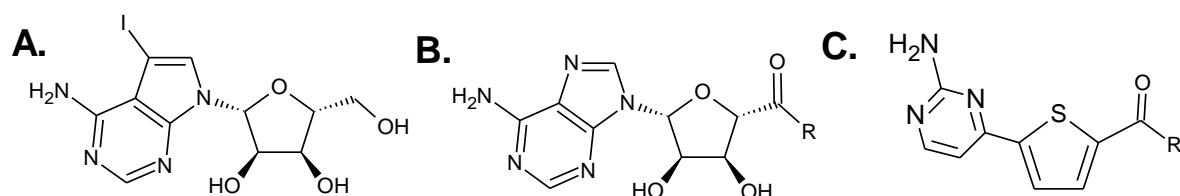


Figure 6. The structures of ATP-site targeting compounds and fragments: A) 5-iodotubercidin, B) adenosine-4'-dehydroxy-methyl-4'-carboxylic acid moiety (Adc), and C) 5-(2-aminopyrimidin-4-yl)-thiophene-2-carboxylic acid moiety (AMTH).

2.3.2.2. Substrate-Site Binding Inhibitors

A specific PK phosphorylates the substrates possessing a certain sequence (known as key-sequence) close to the phosphorylation site. Basophilic Ser/Thr kinases such as Haspin (and PKAc) preferably catalyse the phosphorylation of protein substrates that have a number of positively charged AAs next to the phosphorylatable Ser (or Thr). This fact is used to take advantage of a design of substrate-competitive inhibitors: peptides that mimic the AA sequence of substrate can serve as the substrate-competitive inhibitors [33].

Since the intracellular concentration of the protein substrate is much lower in proportion to ATP, it provides an advantage from the aspect of competition comparing to the inhibitors of ATP-site [34]. Also, a substrate with a certain key-sequence is bound only by few PKs, thus substrate-site inhibitors should generally have a better selectivity than ATP-site compounds [35]. The main downside of substrate-site inhibitors is their size. When the inhibitor binds to the enzyme, the generated interactions need to compensate for the energy loss resulting from removal of structured water layer, i.e., the inhibitor must be sufficiently large. Importantly, the large molecules of the peptidic origin cannot pass through the cell plasma membrane and may have low intracellular stability [33,36]. To increase the stability, analogues of peptides are synthesized where a variety of modifications is introduced in respect to the initial peptide structure, e.g., non-proteinogenic AA (including *D*-amino acids) [37].

2.3.2.3. Bisubstrate Inhibitors

Bisubstrate inhibitors are composed of two fragments and an interconnecting linker. The first bisubstrate inhibitors were developed in the 1970s [38,39] based on an approach in which the

fragments binding to the adjacent but still separate enzymatic sites were combined. In case of successful design, the interactions of both parts with the enzyme are preserved [40]. Bisubstrate inhibitors have the advantage of synergistic effect of both fragments as compared to the separate components [41]. However, if the length of linker is not optimal, the simultaneous binding of inhibitor fragments to their enzymatic sites cannot occur, and thus the inhibitor is no longer of bisubstrate-type. What is more, in order to maintain the mobility of the enzyme, the linker must be sufficiently long and flexible [42].

One of the most studied subset of bisubstrate inhibitors are adenosine analogue-oligoarginine conjugates (ARCs). ARCs consist of an adenosine analogue (a derivative of adenosine or inhibitor targeting ATP-site of PK), an oligo-arginine peptidic part, and hydrophobic linker(s) (Figure 7) [42]. The oligo-arginine fragment in ARC is important for the selectivity towards basophilic PKs; in addition, it provides ARCs with properties of arginine-rich delivery peptides, and hence those are able to pass through the cell plasma membrane [44-48]. The affinity of most efficient ARC-inhibitors are in low nanomolar or picomolar region [40,49].

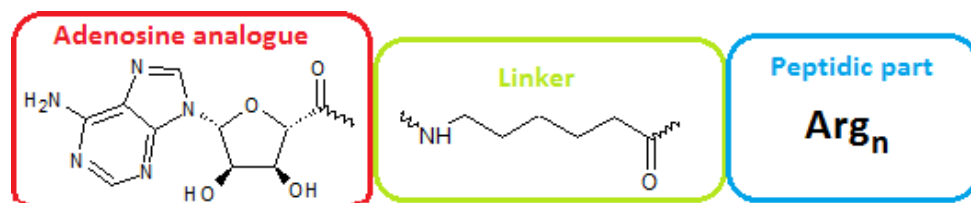


Figure 7. Representative structure of an ARC. An adenosine analogue binding to the ATP-site, a peptidic part binding to the substrate-binding site, and a hydrophobic linker are surrounded by the red, blue and green boxes, respectively.

An important goal in the development of ARCs is the variation of selectivity towards different PKs. The affinity and selectivity can be adapted by using different adenosine analogues, linker(s) and/or peptidic parts [44,49,52]. The design of novel compounds has been greatly aided by the available co-crystal structures of ARCs with their target PKs [52,53]. Some of the most potent ARCs have been linked with fluorescent dyes, and thus it has been possible to use these compounds as probes in biochemical assays [11,50,51].

To date, there are no reports of bisubstrate inhibitors or optical probes available for Haspin. The aim of the current work was to identify the bisubstrate scaffolds that are suitable for this PK. For that, the initial screening of a subset of different ARCs was performed, followed by co-crystallization of the most potent compounds and rational design and synthesis of novel Haspin-selective ARCs. The bisubstrate-analogue conjugates developed in this work incorporated variable ATP-site targeting moieties and the N-terminal peptide of histone H3 as the fragment binding to the substrate site of PK.

3. REAGENTS, EQUIPMENT AND METHODS

3.1. Reagents and Equipment

3.1.1. Reagents

All chemicals were obtained commercially unless noted otherwise. Solvents were purchased from Rathburn and Fluka. SPPS chemicals and resins were provided by Iris Biotech, Neosystem, Novabiochem, Advanced ChemTech, or AnaSpec. Other chemicals were obtained from Sigma-Aldrich. 5-Iodotubercidin was from Cayman Chemicals. PKAc type α (recombinant human protein, full sequence) was a kind gift from Prof. Richard A. Engh's group (Norwegian Structural Bioology, Univeristy of Tromsø). The production and crystallization of Haspin was performed in Prof Stefan Knapp or Dr Alex N. Bullock research groups (University of Oxford). All the materials and equipment needed for the protein production and crystallization were provided by Structural Genomics Consortium.

ARC inhibitors used were synthesized by the following persons²: Ramesh Ekambaram (ARC-3125, ARC-3119*), Erki Enkvist (ARC-0668*, ARC-0684, ARC-0685, ARC-1102*), Katrin Kalind (ARC-3009*, ARC-3010*), Darja Lavõgina (ARC-1012*, ARC-1038, ARC-0902*), Kersti Nisuma (ARC-1141), Mart Roben (ARC-1034*), Asko Uri (ARC-0341*), and Birgit Viira (ARC-1408, ARC-1411).

3.1.2. Equipment

The final products were purified with Shimadzu LC Solution HPLC system (Prominence) by Gerda Raidaru using Gemini C18 reverse-phase column (5 μ m, 25 cm \times 0.46 cm), manual injection and diode array UV-vis detector (SPD M20A). Mass spectra of products were measured in positive-ion mode by using Shimadzu LCMS-2020 (ESI-MS) detector. Elution was performed with water-acetonitrile gradient (0.1% TFA, velocity of gradient 1.8%/min) and flow rate of 1 mL/min. The separated products had the purity >95% and were freeze-dried after elution. The high resolution mass spectra (HRMS) of novel ARCs were measured with Thermo Electron LTQ Orbitrap mass spectrometer. NanoDrop 2000c (Thermo Scientific) spectrophotometer was used for measuring UV-vis spectra and quantification of the compounds.

The concentration of all compounds were determined by UV-vis spectroscopy based on molar extinction coefficient (ϵ) of Adc (15000 M⁻¹ cm⁻¹ at 259 nm), AMTH (15000 M⁻¹ cm⁻¹ at 340 nm), H9 (4400 M⁻¹ cm⁻¹ at 323 nm), Arom6 (16000 M⁻¹ cm⁻¹ at 286 nm), Arom8 (16900 M⁻¹ cm⁻¹ at 250 nm), Iodo1 (10300 M⁻¹ cm⁻¹ at 300 nm), Cy3B (130000 M⁻¹ cm⁻¹ at 558 nm), or TAMRA (80000 M⁻¹ cm⁻¹ at 558 nm).

² Compounds marked with an asterisk were previously characterized as disclosed in earlier publications [8,9,48,49,54].

The fluorescence anisotropy measurements were performed by using a PHERAstar microplate reader (BMG Labtech) with optic module TAMRA [ex. 540(20) nm, em. 590(20) ja 590(20) nm]. All the solutions of samples were prepared in 384-well low-binding surface microtiter plates (Corning, code 3676). GraphPad Prism version 5.0 (GraphPad Software, Inc) was used for data processing and analysis.

3.2. Methods

3.2.1. Solid Phase Peptide Synthesis

In 1963, Robert B. Merrifield introduced the new synthesis technique – solid phase peptide synthesis (SPPS) [55]. In this method, the carboxyl group (C-terminus) of the last AA of synthesized peptide is attached to the solid carrier or resin, and SPPS takes place in the direction of C \rightarrow N. To bind the first AA, there is a certain amount of reactive centers available on resin (represented by so-called loading). With the aim of achieving the maximum yield of the reaction, the loading must be as high as possible, but the steric hindrance caused by growing chains should be avoided.

The addition of each of the following AA consists of three steps (Figure 8):

1. Removal of the N-terminus protecting group of the previous AA

A variety of protecting groups is used for the protection of N-terminus of AA and side-chain of AA containing functional groups. The orthogonality rule applies: the protecting groups of N-terminus and side-chains must be removable under different conditions. Fmoc-SPPS was used in this work.

2. Activation of the C-terminus of added AA

The formation of a peptide bond is a reaction with high activation energy. Thus, the reactants that render the carboxyl group of added AA more electrophilic by pulling electrons from it must be used to facilitate the nucleophilic attack by the amine group of the resin-bound AA.

3. Coupling of AA (acylation reaction)

Finally, the last step of synthesis is the cleavage of the peptidic chain from the resin, accompanied by the simultaneous removal of protecting groups of the side-chains.

SPPS has the advantage of speed, automation, high yield (up to 99.8% per acylation), and easy purification at intermediate stages by washing with solvent and filtering. The disadvantages include the need for a large amount of reagents (3...10 eq. per loading) and complicated analysis of a product at intermediate steps. Some studies have shown different possibilities to analyse the growing peptide on resin [57-59]. In SPPS, so-called colour tests are frequently used, which show the presence or absence of certain functional groups (e.g., Kaiser test: qualitative, shows the existence of free amino groups).

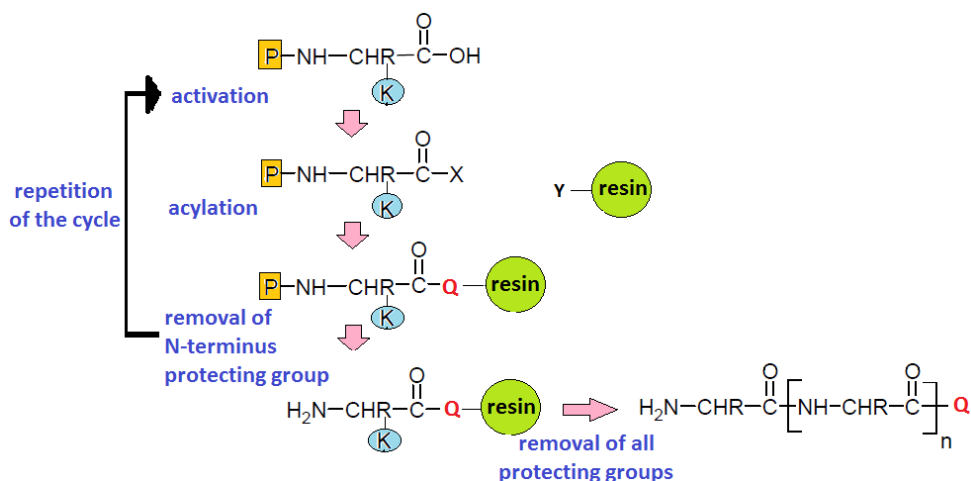


Figure 8. Scheme of SPPS. Markings: P – N-terminus protecting group of AA; K – side-chain protecting group of AA; Y – functional group attached to the resin; X – functional group reacting with activated AA; Q – functional group generated by the reaction between X and Y; N – number of AAs in peptide.

3.2.2. Fluorescence Anisotropy-Based Binding/Displacement Assay (FA)

FA method is based on the measurement of the change of fluorescence anisotropy in solution containing a fluorescence probe. The sample is excited with the linearly polarized light and the polarization of the emitted light is measured. The change in polarization can be caused by binding or displacement of the probe from its complex with a high molecular weight-molecule (e.g., kinase) (Figure 9).

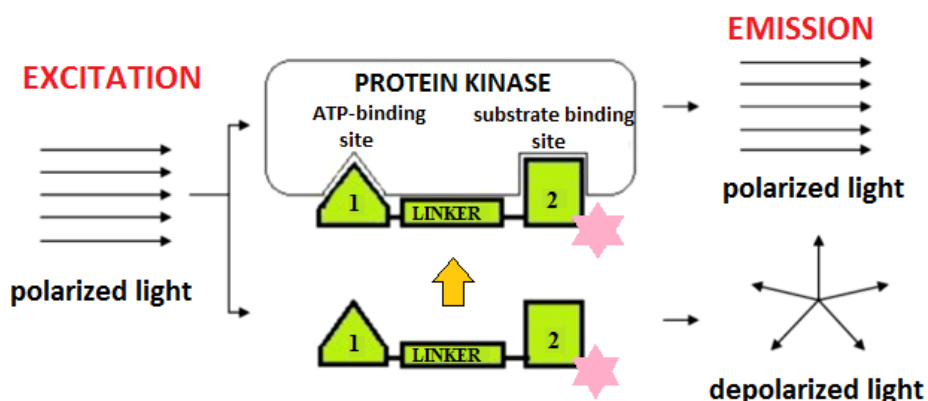


Figure 9. The scheme of FA method. The fluorescence probe is depicted in green colour; number 1 indicates a nucleosidic part, number 2 a peptidic part and a pink star a fluorescence dye.

FA method can be performed in two formats: binding and displacement assay. In the first case, a low molecular weight (1.5...2 kDa) fluorescent probe binds to the active site of the PK ($M_w > 30$ kDa). If complex is formed, the rotation of probe in solution is slowed down and the light emitted by the fluorophore stays polarized. As a result, the polarization/anisotropy of the complex-containing solution is higher than in the solution of free probe, and it is possible to calculate the dissociation constant of probe (K_D), or the active concentration of PK. In the second case, the fluorescence probe is displaced from its complex with PK by an ATP-competitive, substrate-competitive or bisubstrate non-fluorescent inhibitor, which causes the reduction in the polarization. In this case, data processing yields displacement constant of the

non-fluorescent inhibitor [50]. FA method differs from kinetic methods, where the inhibitory potential of compounds is measured by reduction of the rate of catalytic phosphorylation.

FA method has the advantage of simplicity, speed, homogeneity, and automation. The main disadvantage is the low affinity of fluorescence probes ($K_D > 10$ nM) necessitating the use of the high concentration of enzyme, which makes the analysis more expensive. In cases of high affinity probes ($K_D < 10$ nM), nanomolar concentrations of the enzymes can be used.

3.2.3. Thermal Shift Assay

Thermal shift assay (or differential scanning fluorimetry, DSF) is a fast and relatively inexpensive method for the characterization of stability of protein in its free form or in complex with low molecular weight-ligands. Thermal shift assay measures thermal unfolding (denaturation) of a protein in the presence of a fluorescence dye. In this work, SYPRO orange as a fluorescence dye (ex. 492 nm, em. 610 nm) was used due to its high excitation wavelength, which decreases the probability of excitation of other molecules in solution [59]. The measurement can be carried out using an apparatus that is simultaneously capable of controlling the temperature and measuring the fluorescence intensity, for instance, the real-time polymerase chain reaction (PCR) instrument [60].

The stability of a protein is related to its Gibbs free energy of unfolding (ΔG_u). As the temperature increases, the stability of a typical protein as well as its ΔG_u decreases [60]; at equilibrium, the concentrations of unfolded and native protein are the same and ΔG_u is equal to zero. The unfolding reveals the hydrophobic domains of the protein, where fluorescence dye can bind to, causing the increase of fluorescence quantum yield and hence the intensity of the fluorescence signal, which reaches the maximum value at the conditions when protein unfolding is complete. The temperature at which the amounts of unfolded and native proteins are equal ($\Delta G_u=0$) or the fluorescence intensity is half the maximum, is called the melting temperature (T_m). After the signal has achieved the maximum value the fluorescence intensity begins to decrease since the denatured protein aggregates or precipitates (Figure 10) [60,61].

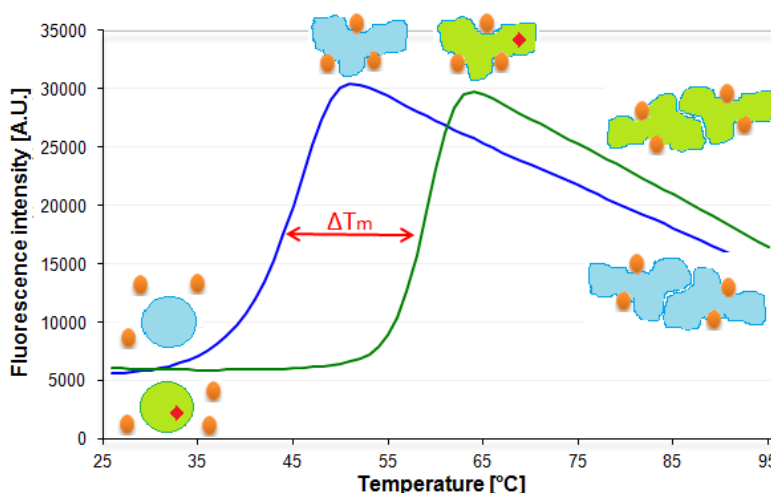


Figure 10. Dependence of the observed fluorescence intensity on the temperature of sample. The blue curve shows a native protein and the green curve a ligand-bound protein. The native and unfolded proteins are marked with a circle and an irregular shape, respectively; the protein-bound ligand is shown as a red rhombus and the molecules of SYPRO orange as orange ovals. Thermal shift value ΔT_m is depicted in red.

In most cases, the binding of a ligand to a protein causes an increase of the ΔG_u because of the contribution of the ligand binding Gibbs free energy (ΔG_l). Consequently, the stabilization of the protein takes place and the melting temperature increases. The difference between the melting temperature of ligand-bound *versus* native protein is called the thermal shift (ΔT_m) value [62,63]. Importantly, studies have shown that ΔT_m values have a relatively good correlation with the inhibition and binding constants of inhibitors (Figure 10) [62-64].

3.2.4. Protein Crystallography

3.2.4.1. General Features of Crystallography

Crystallography is a technique in which a 2D-diffraction pattern is obtained by irradiation of crystal with X-rays, which upon re-calculation provides the 3D-electron density of the crystal lattice components. X-ray radiation has a wavelength of 10^{-7} to 10^{-11} nm (or 1000 to 0.1 Å), but in term of crystallography, 0.4 to 25 Å X-ray radiation is used as it is in the range of the length of chemical bond [65].

Protein crystals differ from the inorganic crystals in several aspects. Firstly, the main difference is a high water content of the former crystals, 40 to 60% of crystal volume. The high water content helps to maintain the native physiological conformations of proteins. As liquid evaporation causes destabilization of the crystal, the crystal needs to be kept inside “mother liquid” or at sufficiently low temperatures to avoid evaporation [66]. Another difference from inorganic crystals is that atoms do not locate at nodes of the unit cell; also, protein crystals are relatively unstable and very sensitive to the environmental conditions [67]. Crystallography of macromolecules has become the interest of pharmaceutical industry, biotechnology and chemistry. The 3D-structures of biomolecules (e.g., enzymes) provide information about their functioning and mechanisms [67]. However, as the proteins differ from one other in several aspects (e.g., AA sequence, chain length, higher order structures), the parameters necessary for the crystallization vary and are often unpredictable [67].

3.2.4.2. Principles of Protein Crystallography

The acquisition process of a protein crystal structure can be divided into four steps:

1. Crystallization

Crystallization takes place when a protein precipitates slowly out from solution and forms a regular crystal lattice. The higher the purity and freshness of protein, the greater is the probability of crystal formation. An appropriate solvent must be selected, where the protein is soluble and the environment is as similar to the physiological as possible, and where salts (e.g., ammonium sulphate) or organic compounds (e.g., 2-methyl-2,4-pentanediol) can induce

precipitation. Crystallization can also be promoted by modification of pH and temperature, and addition of protein-stabilizing cofactors, metal ions, etc. An important factor is time, as the equilibrium between the soluble and crystallized protein is established slowly [66].

The crystal formation process can be divided into two parts – nucleation and crystal growth. First, it is necessary to generate a supersaturated solution where small aggregates (nucleation centres or nuclei) can be formed. After the formation of nucleation centres, the crystal growth can start. Also, supersaturation of solution must be kept at a lower level than in case of the formation of nuclei, in order to prevent the formation of further nucleation centres and thereby growth of many small crystals. The slower growth of crystals also increases the probability to achieve the maximum crystal regularity [66].

A frequently used crystallization technique is a vapour diffusion sitting drop technique, where the drop of the protein solution is located in the microwell adjacent to the reservoir. Water evaporation from the drop and condensation to reservoir solution causes increase of the concentration of precipitating agent in the droplet, which contributes to the nucleation. As water evaporation from the drop is caused by the different concentration of precipitating agent in drop and reservoir, it occurs until the concentrations are equal. The main advantage of this method is the possibility to time the arrival of equilibrium and hence to slow down the formation of supersaturated solution. This can be done by varying the distance between the droplet and reservoir. What is more, it is possible to change the concentration of the components in the droplet by diluting or concentrating the solution in the reservoir [66,67].

2. Preparation for the Diffraction Data Measurement

The high-energy X-ray photons generate free radicals (e.g., oxygen or hydroxyl radicals), which may cause further chemical reactions, and consequently the loss of crystalline regularity. Cooling of a crystal to the cryogenic temperatures (100 – 120 K) significantly reduces the radical formation and diffusion in the crystal, hence improving the resolution of the crystal diffraction pattern. In addition, the low temperatures allow the use of a longer exposure times and the reduction of the thermal motion of protein chains, both aspects also contributing to the improvement of resolution [66].

The crystal cryogenic cooling must be done very quickly (so-called flash freezing or shock cooling), because the water in a crystal must freeze in a vitreous state to avoid crystalline ice formation, which may cause cracks in the crystal, and thus decrease the quality of the diffraction pattern. Such cooling can be performed by placing of the crystal under the cold nitrogen gas stream or into liquid nitrogen. In order to avoid the formation of crystalline ice, a cryoprotective agent is also applied by addition to the crystal growth environment, by transfer

of the crystal from the mother liquid to the cryoprotectant solution (e.g., ethylene glycol), or the crystal can be grown in a solution containing cryoprotectant (e.g., methanol) [66].

Before measurement of the diffraction, crystals must be removed from the mother liquid. Usually, a nylon loop is used for mounting the crystals: a thin liquid layer spanning the loop and surface tension of the liquid holds the crystal in the nylon loop [66].

3. Data Collection

The diffraction of a macromolecular crystal is weaker as compared to crystals of small molecules [66]. The diffraction occurs due to the cooperative effect of molecules in crystal; as the protein molecules are larger than inorganic compounds, there are fewer particles in a crystal, and thus diffraction is weaker. Moreover, protein consists predominantly of atoms with a small number of electrons (C, N, O), and hence reflect X-rays weaker than heavier atoms, which overall causes lower diffraction intensity. Therefore, in order to obtain a reliable quality of the diffraction pattern, it is necessary to use a relatively high intensity of the X-ray source (e.g., synchrotron) [66,67]. Crystal exposure to a beam of X-rays generates a pattern of minima and maxima of interference, where the intensities of the spots provide information about the locations of atoms in the crystal. The pattern of reflections can be recorded with a single-photon counter, a photographic film or an area detector [66].

4. Data Processing and Analysis

As previously mentioned, the pattern of 2D diffraction picture arises from the locations of atoms in the crystal. In order to get the 3D electron-density map from 2D diffraction pattern, the Fourier transform is used. Furthermore, to obtain the electron-density map the amplitude and phase of wave must be known, but detectors used can record only the intensity of amplitude (a limitation known as phase problem). A variety of methods is used to solve the phase problem (e.g., molecular replacement). When the initial phase is found, it is possible to draw up an initial model, which is gradually converged to a simulated model until the resolution of electron-density model does not improve significantly [66,69]. The agreement is evaluated with the R-factor, which indicates the compatibility of the experimental structure and the calculated crystallographic model (normal value $R < 20\%$) [69]. The final step of data processing is refinement, during which the atoms are placed to the electron-density map [65,68]. The resolution³ of diffraction pattern is crucial for an adequate interpretation of the electron-density map, while the higher the resolution (i.e., lower the value), the more realistically the crystal structure can be solved (suitable resolution below 2.7 Å) [64].

³ Resolution represents the minimal distance between the structural elements, which can be distinguished in electron-density map [66].

4. EXPERIMENTAL PART

4.1. Production of Proteins

4.1.1. Expression

Haspin-kinase was expressed in kanamycin-resistant BL21(DE3)-R3-lambda-PPase *E. Coli* cells. An appropriate antibiotic was added to 50 mL of autoclaved LB medium (composition in Appendix 8; final conc. of kanamycin was 50 µg/mL). Thereafter, a small amount of the bacterial cells containing the plasmid was added from previously frozen glycerol stock stored at -80 °C, and incubated in shaker overnight at 180 rpm, 37 °C. Then, 7 mL of overnight bacterial suspension was added to 1 L of autoclaved LB medium with antibiotic. The flasks were incubated at 180 rpm, 37 °C until the absorbance value at 600 nm of 0.5-0.7 was achieved. Subsequently, the protein expression was induced by addition of 1 mM isopropyl-β-D-galactopyranoside (IPTG) and the flasks were incubated overnight at 180 rpm, 18 °C.

4.1.2. Preparation of Lysate

Cells were harvested by centrifugation at 6230 g (centrifuge Beckman Avanti J-20XP, rotor JA-25.50), 4 °C for 15 min. The pellet was re-suspended in binding buffer (Appendix 8) and lysed by sonification for 5 min (5 sec “ON” and 10 sec “OFF”, 35% amplitude). Polyethyleneimine (5%, pH 7) was added to precipitate DNA (1 mL per 40 mL of lysate).

4.1.3. Purification

Since the produced protein contained an oligo-histidine tag at the N-terminus, it was possible to purify it using the affinity chromatography, taking advantage of the binding of the oligo-histidine to nickel-nitrilotriacetic acid (Ni-NTA). The affinity column was prepared by adding 5 mL of Ni-NTA suspended in ethanol to a column. The beads were washed and equilibrated with distilled water and binding buffer, respectively. In order to pellet the insoluble debris, the lysate was centrifuged at 53,200 g (centrifuge Beckman Avanti J-20XP, rotor JLA-8.1000), 4 °C for 1 h. The further action took place at 4 °C. The supernatant was filtered through the 0.45 µm filter onto pre-equilibrated Ni-NTA column. After the supernatant had passed through the column, the column was washed 2 times with 50 mL of binding buffer followed by 30 mL of wash buffer (Appendix 8) with 1 mM TCEP. An elution was performed by passing buffers with the increasing imidazole concentration (50 mM, 100 mM, 150 mM and 250 mM, see Appendix 8) through the Ni-NTA column. Subsequently, the SDS-PAGE analysis of the eluted fractions was performed. Also, 1 mM DTT was added to all fractions.

The sample was concentrated to 4 mL for gel-filtration chromatography by centrifuging at 3220 g, 4 °C (centrifuge Eppendorf 5810R, rotor A-4-62) and using a 30 kDa molecular weight cut-off column (Amicon® Ultra-15 Centrifugal Filter Units, Millipore). The supernatant was purified by chromatograph (ÄKTAexpress, GE Healthcare Life Science) on a

Superdex 200 column with a flow rate of 0.5-1.0 mL/min of gel-filtration buffer (Appendix 8). The peak fractions (as detected by UV absorbance at 280 nm) were collected to a 96-well plate. The relative amount and purity of the protein was checked by SDS-PAGE analysis.

The fractions containing bigger amounts of pure protein were pooled and concentrated by using 30 kDa molecular weight cut-off column (Amicon® Ultra-15 Centrifugal Filter Units, Millipore) and centrifuging at 3220 g, 4 °C. The protein concentration was measured by absorbance at 280 nm as calculated on the basis of its molecular weight and extinction coefficient from its primary structure (Appendix 1). Then, an experimental m/z was confirmed by LC-MS (Agilent LC/MSD TOF system with reversed-phase HPLC coupled to electrospray ionisation and an orthogonal time-of-flight mass analyser). The protein was desalted prior to mass spectrometry by its elution from a C3 column with a gradient of 5-95% acetonitrile in water containing 0.1% formic acid.

4.1.4. SDS-PAGE

5 µL of NuPAGE® loading buffer (5x) with a reducing agent DTT (Appendix 8) was added to 15 µL of a sample. The samples were heated at 70 °C for 5 min. Then, 10 µL of each sample and a molecular weight standard [Precision Plus Protein™ Unstained Standard (BioRad)] were added to the gel [Novex® NuPAGE® 4-12% Bis-Tris Midi Gel (1.0 mm x 26 well)]. The electrophoresis was performed by using NuPAGE® MES (1x) buffer at 170 V for 45 min. The gel was stained with InstantBlue™ (Expedeon) for 1 h and de-stained overnight.

4.2. Synthesis of ARCs

Synthesis of peptidic fragments was performed using traditional SPPS [71] on Fmoc Rink Amide MBHA resin in DMF (1 mL per 100 mg of resin with loading of 0.45 mmol/g). All reactions were carried out at room temperature. Firstly, resin was swelled in DMF for 45 min. Fmoc protective group was removed from N-terminus by treatment with 20% piperidine solution in DMF (5+15 min). When necessary, ivDde protective group was removed from side-chain amine of Lys by treatment with 2% hydrazine solution in DMF (5+15 min). A *D*- or *L*-amino acid or a linker-forming AA (3 eq. per loading) was activated with mixture of HOBt (2.9 eq), HBTU (2.9 eq) and NMM (9 eq); the acylation reaction was carried out for 45-180 min. When necessary, introduction of N-terminal Boc group was performed by using Boc₂O (20 eq) and DIEA (4 eq) in DCE; reaction was carried out for 60 min. Each acylation and removal of the protective group was followed by washing resin five times with DMF. The completeness of coupling was checked by the Kaiser test. IpAdc or AMTH were added to resin-bound peptidic fragment using the activation mixture of HOBt (1.9 eq), HBTU (1.9 eq) and DIPEA (5 eq); reaction time 3 h or 12 h, respectively.

Prior to the standard cleaving procedure, the washing step (five times with DMF, propan-2-ol and DCE) was performed and the resin was dried under vacuum (1 h or more). The cleavage was carried out using the mixture of TFA/H₂O/TIPS (90/5/5, v/v/v) for 3 h. Then, the collected product was washed three times with water and two times with methyl-*tert*-butyl ether by rotating under vacuum. The purification with HPLC followed. The monoisotopic masses of the obtained products were determined by Orbitrap mass spectrometer, and the concentrations were measured with a UV-vis spectrophotometer.

4.3. Biochemical Measurements

4.3.1. FA-method

The binding and displacement curves were measured according to the previously published protocol [49]. The composition of FA buffer used is given in Appendix 8. The concentration of active PK and the K_D values of probes were determined by the binding format of FA-method performing the titration of the PK solution at low (1 or 2 nM) or high concentration (20 nM) of a fluorescence probe, respectively. For this purpose, 2-fold dilution series of enzyme in 384-well microplate using FA buffer were prepared. Subsequently, the fluorescence probe solution was added to both series (final volume 20 μ L). The microplate was incubated at 30 °C for 15 min and then fluorescence anisotropy values were measured with a PHERAstat microplate reader. Binding curves obtained were fitted to equations of enzyme titration [50] using software GraphPad Prism version 5.0 (GraphPad Software, Inc). For displacement curves, 3-fold dilution series of inhibitors were prepared. The enzyme solution together with fluorescence probe ARC-1081 in FA buffer [Haspin and ARC-1081 concentrations were 6 nM and 2 nM (K_D =1.0 nM); PKAc and ARC-1081 concentrations were 6 nM and 2 nM (K_D =0.4 nM)] was transferred to each well. Thereafter, microplate was incubated at 30 °C for 15 min and fluorescence anisotropy measurements with a PHERAstar microplate reader followed. The obtained displacement curves were fitted to equations of logarithmic dose-response model using software GraphPad Prism version 5.0 (GraphPad Software, Inc) in order to determine the IC₅₀ values of inhibitors [50].

4.3.2. Thermal Shift Assay

Thermal shift assay was performed in Prof Stefan Knapp's research group (Structural Genomics Consortium, Oxford). The measurements were performed using a real-time PCR instrument (Mx3005p RT-PCR, Stratagene). 500 μ M solutions of ARCs (in DMSO) were prepared in a 96-well microplate. The reference and control wells containing 100% (v/v) DMSO and distilled water, respectively, were added to the same microplate. Then, 0.6 mL of 2.2 μ M kinase solution in thermal shift assay buffer (Appendix 8) was prepared and 0.6 μ L of SYPRO orange in DMSO was added to kinase solution (Invitrogen, 1:1000 dilution of the

stock solution). Next, 19.5 μL of protein and 0.5 μL of ARC inhibitor solutions (final concentrations 12.5 μM and 2 μM , respectively) were transferred to the required number of PCR low-profile microplate wells (ABgene). The PCR plate was sealed with the optical seal (BioRad) and rotated at 210 g for 1 min (room temperature, centrifuge Eppendorf 5810R, rotor S-4-104). The temperature scan was run from 25 $^{\circ}\text{C}$ to 95 $^{\circ}\text{C}$, at 1 $^{\circ}\text{C}/\text{min}$. GraphPad Prism version 6.0 (GraphPad Software, Inc) and Microsoft Excel version 2007 softwares were used for data processing and analysis.

4.4. Protein Crystallography

4.4.1. Crystallization

The protein was concentrated in gel-filtration buffer to the required concentration ($\sim 370 \mu\text{M}$ or $\sim 15 \text{ mg/mL}$) and approximately two-fold excess of ARC was added. The protein-ARC solution was centrifuged at 20,800 g (centrifuge Eppendorf 5417R, rotor FA-45-30-11), 4 $^{\circ}\text{C}$ for 10 min. Then, 20 μL of the precipitating solution was added to each well of 96-well microtiter crystallization plate (Corning). Firstly, a coarse screen (Appendix 4) was used for crystallization. If crystals were formed but quality was unsatisfactory (e.g., too small size), the precipitating solution properties were accordingly modified for the so-called fine screen (Appendix 5). All the crystallizations were carried out using vapour-diffusion sitting drop method at 4 $^{\circ}\text{C}$ by mixing 100 nL of sample (protein-ARC solution) with 50 nL of precipitating solution. The latter procedure was performed with Mosquito nanolitre pipetting robot (TTP Labtech). The formation of crystals took several days.

4.4.2. Preparations for the Diffraction Data Measurement and Data Processing

Prior to the data collection, the crystals were cryoprotected using the precipitating solution containing ethylene glycol (final conc. 25%). For that, 1 μL of cryoprotectant solution was added to the crystal mother liquid. Thereafter, the crystal was mounted with a nylon loop from mother liquid and flash-frozen in liquid nitrogen.

The diffraction datasets were collected by Dr Eleanor Williams, Dr Apirat Chaikuad or Dr Jola Kopec in Diamond Light Source Ltd synchrotron science facility (Harwell Science and Innovation Campus, Didcot, UK); the detector and wavelength used were a single photon counting detector Pilatus M6 (Dectris) and 0.97949 \AA , respectively.

The diffraction data processing was performed by Dr Apirat Chaikuad. The co-crystal structures of Haspin/ARC were obtained by the molecular replacement technique. The analysis of crystal structures obtained was performed with a molecular visualization program PyMOL (DeLano Scientific LLC).

5. RESULTS AND DISCUSSION

5.1. General Description of the Results

The aim of this study was the generation bisubstrate inhibitors possessing high affinity and selectivity towards Haspin. The work is divided into four parts:

1. Screening towards Haspin of the initial set of ARCs representing variable structural scaffolds (31 compounds), and selection of the candidate compounds for co-crystallization with Haspin;
2. Crystallization of Haspin/ARC complexes (3 different crystals obtained);
3. Analysis of the obtained crystal structures (using 2 co-crystals obtained from the preliminary studies and a published co-crystal);
4. Rational design, synthesis and biochemical characterization of novel ARCs containing N-terminal peptide sequence of histone H3 (14 new compounds).

5.2. Screening of the Initial Set of ARCs

The first step of the studies involved selection of fluorescence probe for Haspin in the FA binding assay. The fluorescence probes representing different structures, but incorporating fluorescent dyes with similar optical properties (TAMRA or Cy3B; codes and schematic structures are shown in Table 1) were selected; all of these compounds have been previously synthesized for studies of other PKs. In this step, the high affinity of the probe was primarily pursued, which would enable screening of a variety of non-labeled ARCs and application of low concentration of kinase for the experiments. Note that all probes contained an oligo-arginine fragment, as it was presumed that basophilic kinase Haspin should possess high affinity towards probes comprising positively charged arginine residues.

Table 1. Schematic structures of fluorescent ARC-probes used for the initial screening, together with values of dissociation constants (K_D) and brightness change (Q)

<i>Code</i>	Structure	K_D (nM)^a (- BSA)	Q	K_D (nM)^a (+ BSA)	Q
ARC-0669	AMTH-Ahx-dArg-Ahx-dArg ₆ -dLys[TAMRA]-NH ₂	8,0 [1,5]	3,3	13 [1,6]	1,4
ARC-0583	Adc-Ahx-dArg ₆ -dLys[TAMRA]-NH ₂	1,4 [0,1]	1,4	2,5 [0,2]	1,3
ARC-1081	Adc-Ahx-dArg-Ahx-dArg ₆ -dLys[Cy3B]-NH ₂	1,0 [0,1]	3,2	n.d.	n.d.
ARC-1042	Adc-Ahx-dArg-Ahx-dArg ₆ -dLys[TAMRA]-NH ₂	1,0 [0,2]	3,6	n.d.	n.d.
ARC-1059	H9-(CH ₂) ₅ -C(=O)-dArg ₆ -dLys[TAMRA]-NH ₂	1,2 [0,7]	3,2	1,8 [0,1]	1,4
ARC-1144	Sele1-Ahx-dArg-Ahx-dArg ₆ -dLys[TAMRA]-NH ₂	26 [3,8]	4,1	11 [1,5]	1,4

^a K_D values measured by FA binding assay at 1 or 2 concentration of fluorescent probe. Standard errors are in parentheses. n.d. - not determined.

In addition to the dissociation constant K_D value, the brightness change (Q) was determined for each fluorescent probe (Table 1). The Q value represents the ratio of emission intensity of the probe bound to the kinase to the emission intensity of the free probe in solution. The

results showed that Q values in FA buffer (Appendix 8) without BSA were higher than in buffer that contained BSA, which may be caused by the non-specific binding of the fluorescence probe to other components of assay or surfaces (e.g., walls of the plastic tubes, wells of the measurement plates, pipette tips). However, the presence of BSA reduces the size of so-called dynamic range or measurement window (i.e., the difference of maximum and minimum values), and may also serve as a component binding ARCs non-specifically; hence, it was decided to perform the further experiments in the absence of BSA.

The most potent probes were **ARC-1081** and **ARC-1042**, which comprised Adc as the adenosine analogue moiety, Ahx as the first and second linker, a chiral spacer *D*-Arg between them, the (*D*-Arg)₆ peptide as the peptidic part with amidated C-terminus, and Cy3B or TAMRA as the fluorescent dye, respectively. The dissociation constants of both compounds toward Haspin remained the same (K_D values of 1 nM), which indicates that there was no influence of the fluorescence dye to the affinity of the compound.

The subsequent screening of non-labelled ARCs was performed by FA displacement assay using **ARC-1081** as the fluorescent probe. The screening set consisted of ARCs incorporating different ATP-site-targeting fragments; other structural variations included different number of Arg residues (0, 2, 6 or 8) and linkers (1 or 2), incorporation of a chiral spacer between the two linker moieties, and attachment of a fatty acid moiety. The results are shown in Table 2.

The data revealed that the affinity of compounds towards Haspin strongly depends on the number of arginine residues; in general, the addition of 2 Arg approximately adds the affinity of two orders of magnitude (e.g., series: **ARC-1034** → **ARC-0582** → **ARC-0902**). Still, 8 Arg in peptidic part (**ARC-1090**) did not improve the affinity compared two compound with 6 Arg (**ARC-0902**). In addition, *D*-Arg seems to be preferred over the *L*-isomer in the peptidic fragment of the compounds (**ARC-0902** vs **ARC-0341**). The effect of addition of myristoyl moiety (Myr) was not uniform: the affinity decreased upon incorporation of Myr in compounds containing Arom8 moiety (**ARC-0684** → **ARC-0685**), but no effect of Myr was observed in compounds containing AMTH moiety as the ATP-site targeting fragment (**ARC-1141** → **ARC-1143**). Surprisingly, the ATP-site targeting fragment itself (Arom6, Arom8, AMTH, Adc, or H9) did not seem to have any significant impact on affinity (the low affinity of H9-containing compounds could rather be attributed to the different linker structure, as the H9-containing **ARC-0903** and fluorescent probe **ARC-1059** previously showed relatively high affinity to the kinase). Additionally, Adc fragment with a linker (**ARC-1010**), or peptide alone (dArg₉-NH₂) were tested. Expectedly, these fragments separately were unable to displace the ligand from its complex with the kinase; therefore, it was demonstrated that in

order to obtain the considerable affinity towards Haspin, the linking of these fragments was required. These results confirmed indirectly the bisubstrate nature of the inhibitors.

Table 2. Schematic structures of non-labelled ARCs used for the initial screening, their affinities obtained by FA displacement assay with the kinase domain of Haspin (N = 2), and the values of thermal shift assay (N=3). The compounds are grouped according to the incorporated ATP-site targeting fragment.

Code	Structure	log IC ₅₀ ^a	K _d (μM)	ΔT _m [°C] ^b
ARC-1408	Arom6-C(=O)-(CH ₂) ₇ -C(=O)-dArg-Ahx-dArg-NH ₂	-4,62 [0,07]	3,7	3,9 [0.0]
ARC-1411	Arom6-C(=O)-(CH ₂) ₇ -C(=O)-dArg ₆ -dLys-NH ₂	-7,12 [0,07]	0,011	7,8 [0.1]
ARC-0684	Arom8-Ahx-dArg-Ahx-dArg ₆ -dLys-NH ₂	-6,95 [0,07]	0,016	5,7 [0.1]
ARC-0685	Arom8-Ahx-dArg-Ahx-dArg ₆ -dLys[Myr]-NH ₂	-5,60 [0,06]	0,38	2,8 [0.1]
ARC-1176	AMTH-Ahc-dArg-NH ₂	-3,85 [0,12]	21,1	n.d.
ARC-1102	AMTH-Ahx-dLys-Ahx-dArg ₂ -NH ₂	-3,99 [0,23]	15,7	2,8 [0.0]
ARC-0668	AMTH-Ahx-dArg-Ahx-dArg ₆ -dLys-NH ₂	-7,00 [0,09]	0,015	8,0 [0.1]
ARC-1141	AMTH-Ahx-dAla-dArg ₆ -dLys-Gly	-6,97 [0,07]	0,016	6,0 [0.1]
ARC-1143	AMTH-Ahx-dAla-dArg ₆ -dLys[Myr]-Gly	-6,36 [0,03]	0,066	n.d.
ARC-1197	AMTH-Ahx-dArg-Ahx-dArg ₆ -dLys[-C(=O)-(CH ₂) ₇ -C(=O)-dArg ₆ -NH ₂]-NH ₂	-7,45 [0,04]	0,005	n.d.
ARC-1034	Adc-Ahx-dArg ₂ -NH ₂	-4,78 [0,15]	2,4	4,5 [0.1]
ARC-0582	Adc-Ahx-dArg ₄ -NH ₂	-6,14 [0,03]	0,11	n.d.
ARC-0902	Adc-Ahx-dArg ₆ -NH ₂	-7,67 [0,04]	0,003	8,6 [0.2]
ARC-1090	Adc-Ahx-dArg ₈ -NH ₂	-7,20 [0,05]	0,009	n.d.
ARC-0341	Adc-Ahx-lArg ₆ -NH ₂	-6,58 [0,04]	0,039	5,7 [0.1]
ARC-0342	Adc-Ahx-lArg ₆ -lLys-NH ₂	-6,66 [0,03]	0,033	n.d.
ARC-1012	Adc-Ahx-dLys-Ahx-dArg ₂ -NH ₂	-4,48 [0,11]	5,0	4,6 [0.1]
ARC-1038	Adc-Ahx-lLys-Ahx-dArg ₂ -NH ₂	-4,52 [0,11]	4,6	4,2 [0.1]
ARC-1041	Adc-Ahx-dArg-Ahx-dArg ₆ -dLys-NH ₂	-8,22 [0,05]	0,0002	n.d.
ARC-3009	H9-C(=O)-CH ₂ -NH-CH ₂ -C(=O)-dArg ₂ -NH ₂	-3,08 [0,59]	81	1,1 [0.1]
ARC-3010	H9-C(=O)-CH ₂ -NH-CH ₂ -C(=O)-dArg ₆ -NH ₂	-5,29 [0,39]	0,78	2,0 [0.0]
ARC-0903	H9-(CH ₂) ₅ -C(=O)-dArg ₆ -NH ₂	-6,61 [0,04]	0,037	n.d.
ARC-3125	Iodo1-CH ₂ -C(=O)-Ahx-dArg ₆ -dLys-NH ₂	-8,09 [0,03]	0,008	10,1 [0.1]
ARC-1010	Adc-Ahx-OH	no binding at 33 μM		
	dArg ₉ -NH ₂	no binding at 66 μM		

^a Logarithms of IC₅₀ values measured with FA-assay by displacement of 2 nM fluorescent probe ARC-1081 from its complex with 6 nM Haspin or PKAc. ^b Thermal shift value. Standard errors are in parentheses. n.d. - not determined.

For all of the aforementioned FA-assays, Haspin was provided by SGC; next, after the initial screening and choice of most effective compounds, the studies were continued in University of Oxford and the kinase was produced in order to obtain very pure and fresh protein for the subsequent crystallographic studies (see below). The production of the His-tagged kinase domain of Haspin was performed in BL(DE3) *E.Coli* bacterial cells co-expressing λ-phosphatase, and the purification involved Ni-affinity and gel-filtration chromatography.

In order to confirm that compounds survived the transportation and that the new portion of PK was active, thermal shift assay was performed. The results (Table 2) were compared to those obtained from FA displacement assay (Figure 11). While thermal shift assay indicated that the produced kinase was active and confirmed the major trends in affinities of compounds toward Haspin, the correlation between the characteristic values obtained from thermal shift assay *versus* FA assay (ΔT_m *versus* $\log IC_{50}$) remained in the moderate range (R^2 value of 0.56, Figure 11). Interestingly, it was noted that compounds containing H9 (**ARC-3009**, **ARC-3010**) had the major deviation from the results. This probably indicates a different binding mode of these compounds to Haspin as compared to other compounds. When the data points corresponding to these compounds were excluded from the linear correlation analysis, the correlation improved significantly (R^2 value of 0.77, Figure 11). Thereafter, three most potent compounds (except **ARC-0902**) were chosen for crystallization trials with Haspin kinase (see below). **ARC-0902** was excluded from this set, as this compound had by this time been already co-crystallized with Haspin by our collaboration partners in Oxford; another ARC that had been co-crystallized by SGC was **ARC-1141**.

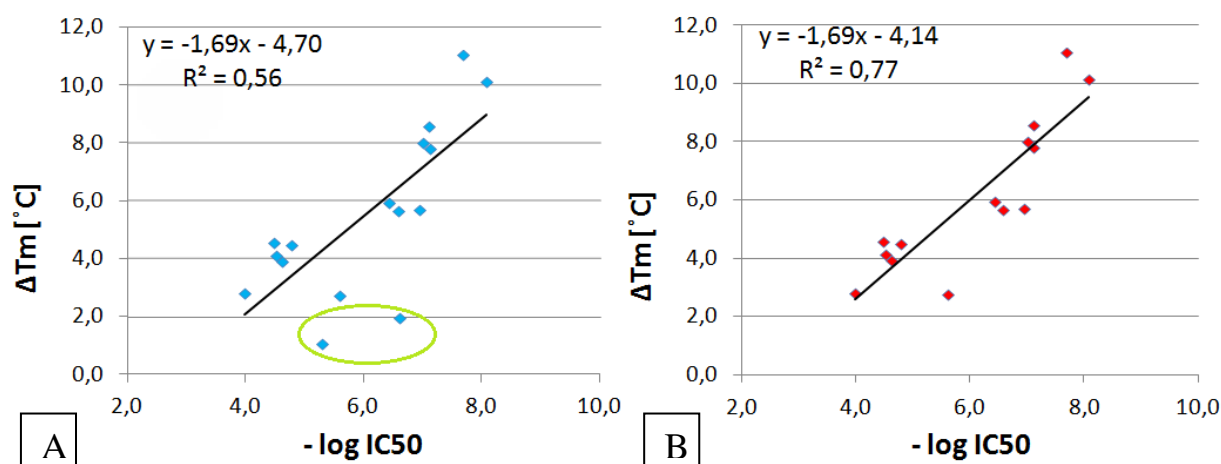


Figure 11. The linear correlation between values of thermal shift (on the y-axis) and $-\log IC_{50}$ (on the x-axis). **A.** Correlation between all compounds analysed. **B.** Correlation excluding the data points corresponding to H9-containing compounds. The data points are shown as blue or red diamonds; the black line indicates linear fit; the data points excluded from the second graph are surrounded by green circle.

5.3. Co-crystallization of Haspin and ARCs

The ARC-inhibitors used for co-crystallization with kinase domain of Haspin were chosen based on results of FA displacement assay and thermal shift assay (the compounds presented in Table 3). The crystals were obtained with all three compounds (the pictures of the crystals are presented in Appendix 6), and the X-ray diffraction data were collected with high resolution (1.5 Å). Unfortunately, as the data processing is extremely time-consuming process, it has not been finished by the deadline of this study. Still, for the further steps of the given study we could take advantage of the three co-crystal structures that had been obtained

previously by our collaborators: two of these crystals represented Haspin/ARC complexes, and one crystal represented Haspin/H3(1-7) complex.

Table 3. The ARC-inhibitors used for co-crystallization with kinase domain of Haspin and the general crystallization results

Code	Structure	Resolution of X-ray diffraction data
ARC-1411 ^a	Arom6-C(=O)-(CH ₂) ₇ -C(=O)-dArg ₆ -dLys-NH ₂	1.5 Å
ARC-0668 ^a	AMTH-Ahx-dArg-Ahx-dArg ₆ -dLys-NH ₂	1.5 Å
ARC-1141 ^b	AMTH-Ahx-dAla-dArg ₆ -dLys-Gly	1.87 Å
ARC-0902 ^b	Adc-Ahx-dArg ₆ -NH ₂	1.7 Å
ARC-3125 ^a	Iodo1-CH ₂ -C(=O)-Ahx-dArg ₆ -dLys-NH ₂	1.5 Å

^a New diffraction data. ^b Diffraction data obtained previously by SGC.

5.4. Analysis of Crystal Structures

As mentioned above, due to the issues with the data processing, the rational design of novel ARCs was based on analysis of the previously obtained co-crystal structures of Haspin with **ARC-0902** or **ARC-1141**. In both cases, the resolution of crystal structure was high (1.7 and 1.87 Å, respectively) and both structures showed typical space group P2₁2₁2₁ for Haspin kinase domain (more structural factors are given in Appendix 7).

The electron densities corresponding to the linker and the inhibitor fragments targeted to the ATP-site of Haspin were well-defined. However, the electron densities corresponding to the peptidic parts are only defined for three (**ARC-0902**; Figure 12) or two (**ARC-1141**; Figure 13) arginine residues, indicating that arginine residues of the peptide are relatively mobile and can adopt multiple conformations. The compound **ARC-0902** binds to Haspin in two conformations, the differences involving particularly the peptidic part of the inhibitor. The adenosine analogue (Adc) of **ARC-0902** binding to the ATP-site of Haspin gives mainly the same interactions as ATP (Figure 12); however, the interaction with **Gly608** is missing.

The linker moiety is positioned under the Gly-rich loop of the kinase, which is similar to the previously reported co-crystals of ARCs with PKAc [52]. The 1st Arg of the peptidic part (i.e., counting from the linker) forms hydrogen bonds with **Asp649** and **Thr689**, which is also similar to the typical binding pattern of ARCs observed in co-crystal structures with PKAc [51]. Interestingly, the next arginine residue flips over the Gly-rich loop of the kinase and forms hydrogen bonds with Haspin residues **Ala587** and **Asn588** positioned in the N-terminal lobe of the kinase. Similarly, **Arg2** of H3(1-7) form hydrogen bond with **Asn588** (Figure 5). Finally, 3rd Arg has two different positions in the co-crystal also forming different hydrogen bond patterns – **Glu613** and **Gln718** or **Gly653** and **Asn654**. In case of the former conformation, a π - π stacking occurs between the guanidine group of arginine moiety and the imidazole ring of **His651**.

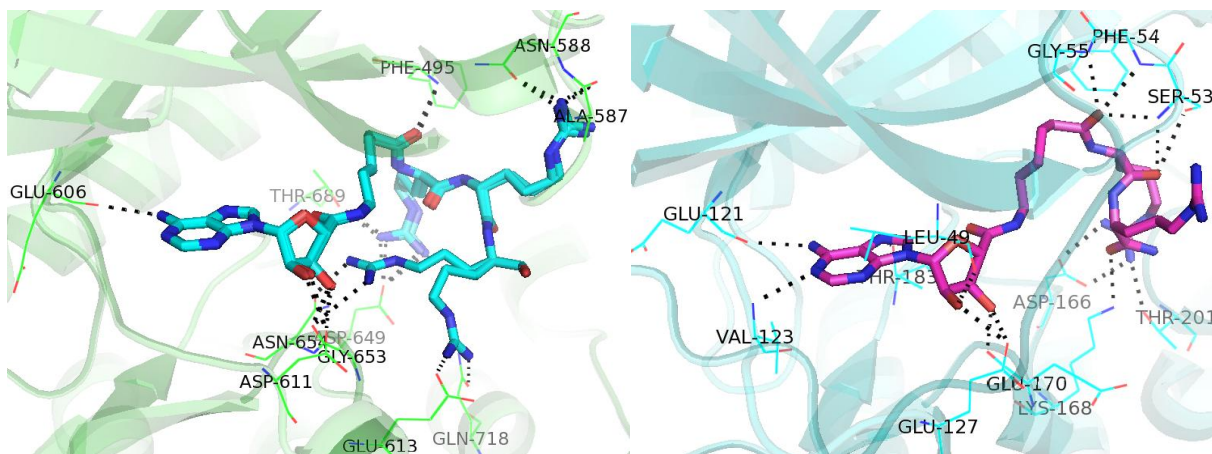


Figure 12. Co-crystal structure of ARC-0902 (blue sticks) with Haspin kinase domain (green cartoon) and co-crystal structure of ARC-1034 (magenta sticks) with PKAc domain (blue cartoon) (PDB code 4BWJ, Protein Databank Homepage). Residues of PK forming interactions with the co-crystallized ARC are shown as lines and

Unlike the binding of **ARC-0902** to kinase domain of Haspin, **ARC-1141** gives remarkably smaller number of hydrogen bonds. The exocyclic nitrogen of AMTH moiety behaves as a donor of hydrogen bond for **Glu606** (the same for binding of ATP is shown in Figure 4). Interestingly, the hydrophobic linker following the AMTH moiety makes a sharp turn ($> 90^\circ$), which is a conformation that has not been seen before in any co-crystal structures of ARCs with PKs. Also, the backbone NH of 1st Arg counted from the linker forms a hydrogen bond with **Gln718**, but there is no interaction visible for the side-chain of 1st Arg. However, guanidine group of this Arg might be involved in π - π stacking with the amide bond between **Gly715** and **Asp716** of Haspin. The 2nd Arg residue of **ARC-1141** protrudes up to the Gly-rich loop and forms two charge-reinforced hydrogen bonds with the backbone carbonyl of **Glu492**.

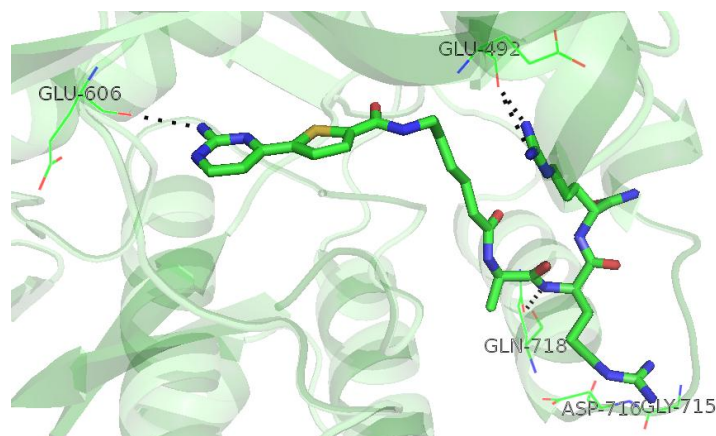


Figure 13. Co-crystal structures of ARC-1141 (green sticks) with Haspin kinase domain (green cartoon). Residues of PK forming interactions with co-crystallized ARC are shown as lines and are labelled; hydrogen bonds are shown as black dotted lines.

5.5. Synthesis and Biochemical Characterization of Novel ARCs

The design of novel ARCs was based on co-crystal structures of **ARC-1141** and **ARC-0902** (unpublished data), which were overlaid with the co-crystal of Haspin kinase domain with histone H3(1-7) (PDB 4OUC). The peptidic part of novel compounds was identical with or

similar to the sequence of H3(1-7) (Figure 5). The structures of novel ARCs are presented in Table 4 and Appendix 9.

The overlay of co-crystals of **ARC-0902** *versus* H3(1-7) with Haspin (Figure 14) showed that the 1st *D*-Arg (counted from the linker) of ARC was positioned close to the **Thr3** of H3(1-7). Therefore, the compound **ARC-3323** was designed incorporating Adc as the ATP-site-targeting fragment, followed by linker (Ahx) and one lysine residue; this part of the novel conjugate was expected to imitate the binding mode of **ARC-0902**. Another part of the novel conjugate was expected to mimic H3(1-7); however, on synthetic rationale (e.g., to avoid non-desired side-reactions with side-chains of AAs) **Thr6** of H3 was replaced by *L*-Ala, and the N- and C-termini of the substrate peptide were acetylated and amidated, respectively. In order to link the two aforementioned parts, the **Thr3** of H3 was replaced by *L*-Glu. Synthesis of the compound was accomplished as follows: the peptidic part was synthesized on solid phase and, thereafter, cleaved and purified by HPLC. Next, the fragment IpAdc-Ahx-dLys(ivDde) was synthesized on solid phase, and the purified peptide was conjugated to the side chain of *D*-Lys on resin after removal of the orthogonal protecting group ivDde which is not cleavable by 20% piperidine/DMF used for Fmoc-deprotection. Finally, the product was cleaved from the resin by TFA, accompanied by the simultaneous removal of the isopropylidene protection from the hydroxyl groups of ribose moiety of Adc.

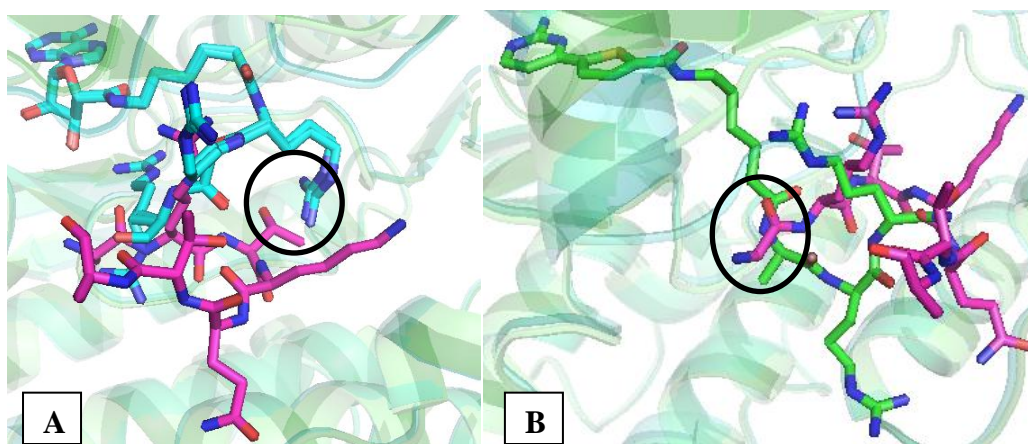


Figure 14. Co-crystal structures of ARC and histone H3/Haspin kinase domain. A, Overlay of co-crystals of ARC-0902 (blue) versus histone H3 (magenta; residues 1-7, PDB 4OUC) with Haspin. B, Overlay of co-crystals of ARC-1141 (green) *versus* histone H3 (magenta; residues 1-7, PDB 4OUC) with Haspin. Kinase domain is indicated as cartoon; ARCs and H3 as sticks; the black circle shows the place, where the linkage between the adenosine analogue and peptide was introduced in novel ARCs.

The overlay of co-crystals of **ARC-1141** *versus* H3(1-7) with Haspin (Figure 14) indicated that the **Ala1** was located almost at the same place as *D*-Ala of **ARC-1141** functioning as a 1st AA of peptidic part (counted from the linker). Thus, a novel compound (**ARC-3324**) was designed incorporating AMTH as the ATP-site targeting fragment linked *via* a Gly residue with the H3(1-7) peptide. The linkage point in the peptide was the side-chain of *D*-Lys replacing the **Ala1**; *D*-Lys was chosen instead of *L*-Lys, as the side-chain of **Ala1** in H3(1-7)

co-crystal was not pointing towards ARC, and the change of chirality was expected to compensate for this difference. In order to increase the length and flexibility of linker, Gly was added between the AMTH and side-chain of *D*-Lys. Again, for the synthetic rationale, the N- and C-termini of the substrate peptide were acetylated and amidated, respectively. The synthesis of the compound was accomplished as follows: the peptidic part was synthesized on solid phase using the side chain of *D*-Lys with orthogonal protection group ivDde. Then, the linker(s) and AMTH fragment were added through the side chain of *D*-Lys using the traditional SPPS; finally, the finished product was cleaved from the solid phase accompanied by removal of acid-labile protecting groups.

The affinities of the obtained conjugates and reference compounds were characterized by FA displacement assay (Table 4). As the K_d values of the novel ARCs were dramatically higher than those of the starting compounds (**ARC-0902**, **ARC-1141**) or 5-iodotubercidin, thus the further design of counterparts of **ARC-3323** and **ARC-3324** was performed.

Taking into account the possible use of compounds in living cells in the future, the compounds lacking the oligo-arginine sequence may also lack the ability to penetrate the plasma membrane of a cell, and thus oligo-arginine segment was introduced to the structure of novel ARCs. Unfortunately, the further synthesis of Adc-containing compounds (i.e., **ARC-3323** counterparts) was not successful due to the synthetic complications: the addition of peptidic part to the solid phase fragment similarly as in the synthesis of the **ARC-3323** did not occur. The possible solution that will be pursued in the future studies would be the exchange of positioning of Lys(ivDde) and Asp, i.e., the *D*-Asp will be located in the first fragment (Adc-Ahx-dAsp) and the orthogonally protected *L*-Lys in the peptidic part in place of phosphorylatable Thr.

The first step in optimization of **ARC-3324** counterparts was introduction of the oligo-arginine (*D*- or *L*-) sequence to the N-terminus of the H3 part of ARC, resulting in compounds **ARC-3327** and **ARC-3328**, respectively. As expected, the addition of oligo-arginine sequence increased the affinity of conjugates towards basophilic PK Haspin, whereas the affinity was the similar for both isomers (Table 4). In parallel, the biochemical characterization of novel compounds was also performed with PKAc as the target kinase: it is known that the starting compounds **ARC-0902** and **ARC-1141** have high affinity towards basophilic kinase PKAc, and it was thereby important to determine the relative selectivity profile (i.e., Haspin *versus* PKAc) of the novel ARCs. The results indicated that **ARC-3327** and **ARC-3328** had the affinity in the same range towards both Haspin and PKAc; still, the selectivity profile of novel compounds was significantly improved in respect to the starting compound **ARC-1141** which possessed prominently better affinity towards PKAc than

Haspin. Remarkably, of all novel ARCs the compound **ARC-3328** had the highest affinity towards Haspin (Table 4).

Next, the changes in the structure of **ARC-3324** were introduced. First, the N-terminus of H3(1-7) fragment was kept non-acetylated, which was synthetically accomplished as follows: the Fmoc-group was removed from N-terminus of the peptide represented by the backbone amino group of *D*-Lys (side-chain protected with an orthogonal protecting group ivDde). Subsequently, the N-terminus was protected with a Boc-group [72]. Thereafter, similarly to the synthesis of **ARC-3324** the linker(s) and AMTH were added through the side chain of Lys, however, the Boc-group was removed together with the cleavage of finished product, thus retaining the N-terminus of peptide free. The rationale for such structural change was that the N-terminal Ala1 in the co-crystal of Haspin with Histone H3 peptide (1-7) forms a hydrogen bond with **Glu613** of Haspin (Figure 5), and it was presumed that this interaction might contribute significantly to the binding energy of the novel conjugates. The non-acetylated N-terminus increased the affinity for Haspin approximately 18-folds, but for PKAc the affinity remained the same range (compounds **ARC-3324** and **ARC-3342**) (Table 4).

Additionally, variation of the length of the linker between the AMTH and side-chain of *D*-Lys was performed (compounds **ARC-3343** to **ARC-3345**). The replacement of Gly by γ -aminobutyric acid (GABA; compounds **ARC-3342** and **ARC-3343**) increased the flexibility of the linker as well the affinity towards both kinases: affinity for Haspin and PKAc increased about 1.4- and 1.1-folds, respectively. Furthermore, the addition of oligo-arginine sequence (*D*- or *L*-) to the C-terminus of the H3 peptide in combination with the elongation of linker and lack of acetylation at the N-terminus of H3(1-7) fragment (compounds **ARC-3344** and **ARC-3345**) significantly increased the affinity for Haspin compared to analogous compound without oligo-arginine (**ARC-3343**). On the other hand, if compared to the counterparts incorporating the oligo-arginine segment at the N-terminus of peptide (**ARC-3327** and **ARC-3328**), the affinities of **ARC-3344** and **ARC-3345** was nearly the same. Importantly, however, the latter compounds showed the most potent selectivity profile towards Haspin compared to PKAc (4 to 5 times). Overall, the compounds **ARC-3344** and **ARC-3345** had the highest selectivity of all novel ARCs towards Haspin (Table 4).

Table 4. Schematic structures of the starting compounds and novel ARCs, and the displacement constant K_d values calculated from the FA displacement assay with the kinase domain of Haspin or PKAc (N = 2)

Code	Structure	$\log IC_{50}^a$		K_d (μM) ^b		Selectivity Index ^c
		Haspin	PKAc	Haspin	PKAc	
ARC-3323	Adc-Ahx-dLys*-[Ac-lAla-lArg-lAsp*-lLys-lGln-lAla-lAla-NH ₂]	-4,47 [0,06]	-4,58 [0,02]	5,2	1,8	0,4
ARC-3324	AMTH-Gly*-[Ac-dLys*-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂]	over -4	over -4	n.c.	n.c.	n.c.
ARC-3327	AMTH-Gly*-[Ac-dArg ₆ -dLys*-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂]	-6,08 [0,06]	-5,87 [0,02]	0,13	0,099	0,8
ARC-3328	AMTH-Gly*-[Ac-lArg ₆ -dLys*-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂]	-6,27 [0,06]	-6,22 [0,02]	0,082	0,042	0,5
ARC-3342	AMTH-Gly*-[dLys*-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂]	-4,24 [0,05]	-4,06 [0,03]	8,7	6,2	0,7
ARC-3343	AMTH-GABA*-[dLys*-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂]	-4,37 [0,36]	-4,12 [0,03]	6,4	5,4	0,9
ARC-3344	AMTH-GABA*-[dLys*-lArg-lThr-lLys-lGln-lThr-lAla-dR ₆ -NH ₂]	-6,09 [0,03]	-5,06 [0,02]	0,12	0,62	5
ARC-3345	AMTH-GABA*-[dLys*-lArg-lThr-lLys-lGln-lThr-lAla-lR ₆ -NH ₂]	-6,01 [0,03]	-5,08 [0,02]	0,15	0,59	4
ARC-3346	AMTH-Ahx-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	-4,47 [0,06]	-5,30 [0,02]	5,2	0,35	0,1
ARC-3347	AMTH-Aoc-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	over -4	-6,11 [0,02]	n.c.	0,054	n.c.
ARC-3348	AMTH-dPro-Gly-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	over -4	over -4	n.c.	n.c.	n.c.
ARC-3349	AMTH-lPro-Gly-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	over -4	over -4	n.c.	n.c.	n.c.
ARC-3350	AMTH-isonipeconic acid-Gly-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	-4,21 [0,06]	-4,26 [0,10]	9,3	4,0	0,4
ARC-3351	AMTH-tranexamic acid-Gly-dDAP-lArg-lThr-lLys-lGln-lThr-lAla-NH ₂	-4,05 [0,22]	over -4	13	n.c.	n.c.
ARC-0902	Adc-Ahx-dArg ₆ -NH ₂	-7,67 [0,04]	-7,30 [0,13]	0,0031	0,0032 ^d	1,0
ARC-1141	AMTH-Ahx-dAla-dArg ₆ -dLys-Gly	-6,97 [0,07]	-8,25 [0,02]	0,016	0,0001	0,006
	5-iodotubercidin	-8,00 [0,05]	over -4	0,0008	n.c.	n.c.

Red text denotes residues corresponding to peptide of H3, and blue text denotes the residues and fragments that were changed compared to the original structure of H3 peptide. The asterisks indicate attachment of the substituent to the side-chain of AA. **ARC-0902**, **ARC-1141** and 5-Iodotubercidin were used as reference compounds. ^a Logarithms of IC_{50} values measured with FA-assay by displacement of 2 nM fluorescent probe ARC-1081 from its complex with 6 nM Haspin or PKAc. ^b K_d values calculated by online calculator [71]. ^c The index is defined as the ratio of the K_d values of inhibitors in experiments with PKAc *versus* Haspin. ^d The previously published inhibition constant K_i value is in excellent agreement with the obtained data (K_i value of 3.2 nM [48]). n.c. – not calculated.

Finally, in order to simplify the overall synthetic scheme, it was decided to link AMTH directly to the N-terminus of H3(1-7) peptide *via* a hydrophobic linker, i.e., 6-aminohexanoic (Ahx; compound **ARC-3346**) or 8-aminooctanoic acid (Aoc; compound **ARC-3347**). However, to preserve the hydrogen bonding pattern characteristic to binding of N-terminus of histone H3(1-7) to Haspin, 2,3-diaminopropionic acid (dDAP) moiety was substituted in conjugates for Ala1 of H3(1-7) peptide (it was presumed that the side-chain amine of dDAP could develop the analogical charge-reinforced hydrogen bond with **Glu613** of Haspin). The resulting compound with Ahx linker (**ARC-3346**) was so far the most potent AMTH-containing compound of novel ARCs without oligo-arginine (K_d 5.2 μ M; Table 4). Also, since there was an unusual sharp turn of hydrophobic linker moiety following the AMTH in the co-crystal structure of Haspin/ARC-1141 (not seen in co-crystals of ARCs with PKAc; Figure 13), the variation of the linker structure was performed by introduction of a cyclic element to the structure of the linker. It was expected that the cyclic element could facilitate the turn of the linker and should increase the selectivity of compounds towards Haspin over PKAc, as it has been shown that the latter binds preferably compounds with long flexible linkers. *D*-Pro, *L*-Pro, isonipecotic acid moiety, tranexamic acid moiety were used as such cyclic elements (compounds **ARC-3348** to **ARC-3351**), whereas the net length of the linker remained unchanged, as the aromatic or cyclic fragment plus Gly gave approximately the same length as aminohexanoic acid moiety. Importantly, the results showed that the latter change resulted in dramatic drop of the affinity of conjugates towards both Haspin and PKAc, suggesting that despite the different binding modes of the linker moieties of the conjugates observed in Haspin *versus* PKAc co-crystals, the flexibility of the linker is important for binding of conjugate to either PK.

In summary, the following tasks were accomplished within the given thesis:

- Screening of a variety of scaffolds of previously reported ARCs with Haspin was performed;
- Three new diffractive co-crystals of Haspin/ARC were obtained and the data sets were collected with high resolution (1.5 Å);
- Based on analysis of previous co-crystal structures of Haspin with ARCs or histone H3(1-7), 14 novel ARCs were synthesized and characterized in biochemical displacement assays with Haspin and PKAc;
- The highest affinity toward Haspin was obtained for the compound **ARC-3328** (K_d = 82 nM), and the highest selectivity over PKAc for the compounds **ARC-3344** and **ARC-3345** (K_d values for Haspin of 120 and 150 nM, and for PKAc of 620 and 590 nM, respectively; selectivity index value of 4-5).

5.6. Prospectives

The present work focused on the development of bisubstrate inhibitor scaffolds towards mitotic Haspin kinase. 31 previously synthesized adenosine analogue-oligoarginine conjugates were characterized in biochemical assays with Haspin, and 3 candidates were chosen for crystallographic studies. From the available co-crystal structures, rational design of 14 novel compounds was performed which were also characterized in biochemical assays with Haspin and a reference basophilic kinase PKAc. Altogether, this work serves as the strong starting point for further studies of Haspin in the future, though some possible improvement of the developed compounds might be considered.

To start with, synthesis of series of Adc-containing conjugates should be performed to take whole advantage of the available ARC-0902/Haspin co-crystal. Next, it is important to identify other natural (or natural-like) peptides and their analogues targeting the substrate-site of kinase domain of Haspin. For instance, a recent study that investigated the preferred consensus sequences in substrates of Haspin suggested that *L*-Val could be used instead of *L*-Ala as the first AA of N-terminus of histone H3 [24]. Other major changes in the peptidic fragments of novel conjugates leading to the improved affinity and selectivity of the latter can be identified by screening of peptide libraries. In parallel, further characterization of the obtained compounds should be continued, including selectivity profiling in biochemical assays, determination and improvement of the ability of conjugates to penetrate the cell membrane, etc.

As the long-term goal, it is envisioned that the existing ARCs as well as the compounds that are to be developed based on the generated co-crystals and structure-affinity relationship will be applied for detection and quantification of Haspin kinase activity in a variety of biochemical and biological systems, including diagnostic panels.

6. SUMMARY

Protein kinases (PKs) catalyse phosphorylation, a process during which the γ -phosphoryl group is transferred from ATP to a protein or a peptide substrate. The phosphorylation of proteins serves as a molecular switch, increasing or decreasing the ability of the substrate protein to participate in different cellular processes. A flow of these processes is represented by the cell cycle (including mitosis), the timing of which is precisely regulated by crosstalk of different PKs. Hence, the aberrant activity of kinases is associated with several diseases related to abnormal cell proliferation, and protein kinases have accordingly become important research targets. A wide variety of compounds reporting and regulating the abnormally increased activity of PKs (i.e., protein kinase-targeting probes and inhibitors) has been developed during the last two decades.

The present work describes the rational design, synthesis and biochemical characterization of inhibitors targeting the recently discovered mitotic kinase Haspin. After the initial screening of a large panel of ARC-type PK inhibitors towards Haspin, the most potent candidate compounds were co-crystallized with the freshly produced Haspin kinase. The previously obtained co-crystal structures were used for the rational design of the unique compounds, incorporating a fragment targeting the ATP-site of the kinase, and the peptidic part originating from the only known substrate protein of Haspin, histone H3. 14 novel conjugates were synthesized according to the conventional Fmoc solid phase peptide synthesis procedures, and characterized in binding/displacement assay with Haspin and a reference kinase PKAc. The most potent novel compound (**ARC-3328**) possessed two-digit nanomolar affinity towards Haspin. Furthermore, the most selective compounds **ARC-3344** and **ARC-3345** had 4 to 5 times higher affinity for Haspin compared to the basophilic protein kinase PKAc.

To sum up, this work has revealed several important structure-affinity trends for the design of inhibitors targeting mitotic protein kinase Haspin. The affinity and selectivity of the compounds developed within this thesis will enable their application as the Haspin-targeting probes in the subsequent studies in a variety of biological systems, including live cells and tissues. Furthermore, as no bisubstrate inhibitors of Haspin have been reported so far, our compounds represent the unique scaffolds that can provide valuable information for studies of binding modes of natural substrates of Haspin.

7. KOKKUVÕTE

Mitootilise Haspin-kinaasi mõlema substraadi taskusse seonduva inhibiitori konstrueerimine

Proteiinkinaasid katalüüsivad fosforüülimisreaktsiooni, mille käigus toimub γ -fosforüülrühma ülekandmine lähtemolekulilt (ATP) valksubstraadile. Valkude fosforüülimine toimib molekulaarse lülitina, kaasates või lülitades substraatvalku välja olulistest rakulistest protsessidest. Selliste protsesside kogumina võib vaadelda rakutsükli (sh mitoosi), mille ajastus on täpselt reguleeritud erinevate proteiinkinaaside koostöö poolt. Kinaaside häirunud aktiivsust seostatakse erinevate haiguste, sealhulgas rakkude kontrollimatust jagunemisest tingitud vähkkasvajate tekke ja/või arenguga. Seetõttu on proteiinkinaasidest saanud olulised uurimisobjektid, mille anomaalselt kõrgeks aktiivsuse tuvastamiseks ja reguleerimiseks on viimaste aastakümnete jooksul arendatud suur hulk madalmolekulaarseid ühendeid (nt proteiinkinaaside sondid ja inhibiitorid).

Käesolev töö kirjeldab hiljuti avastatud mitootilisele Haspin-kinaasile suunatud inhibiitorite ratsionaalset disaini, sünteesi ja biokeemilist iseloomustamist. Töö käigus teostati esialgne Haspin-kinaasi sõeluuring 31 tundud ARC-tüüpi inhibiitoriga, mille abil valiti välja kõige afiinsemad kandidaatühendid koos-kristallimiseks värselt toodetud Haspiniga. Seejärel kasutati eelnevalt saadud kristallstruktuure uute unikaalsete ühendite ratsionaalseks disainiks. Need ühendid sisaldasid kinaasi ATP-saiti seostuvat osa ühendatuna substraadi-sidumissaiti seostuva peptiidse fragmendiga, mis jäljendas Haspini ainsat teadaolevat substraatvalku histoon H3. Kasutades traditsioonilist tahkefaassünteesi meetodit, valmistati kokku 14 uudet ühendit, mille võimet seostuda Haspini ja võrdluskinaasi PKAc-ga iseloomustati fluorestsentsi anisotroopia mõõtmisel põhineva sidumis-/väljatõrjumismeetodiga. Kõige efektiivsem uudne ühend (**ARC-3328**) omas nanomolaarset afiinsust Haspini suhtes ($K_d = 82$ nM). Lisaks, parima selektiivsusega ühendite **ARC-3344** ja **ARC-3345** afiinsus oli 4 – 5 korda kõrgem Haspini suhtes võrreldes PKAc-ga. Saavutatud afiinsust ja selektiivsust võib pidada silmapaistvaks, kuna antud kinaasile on antud hetkeks suudetud arendada vähe inhibiitoreid, kusjuures puudub info bisubstraatsete inhibiitorite kohta.

Kokkuvõttes näitas antud töö mitmeid olulisi struktuur-afiinsus sõltuvuse suundumusi bisubstraatsete inhibiitorite arendamisel mitootilisi kinaasi Haspini jaoks. Uudseid ühendeid on võimalik kasutada Haspini sondidena järgnevates uuringutes erinevates bioloogilistes süsteemides (nt elusrakud ja koed), selgitamaks selle väheuuritud kinaasi rolle ja partnereid rakusisestes signaaliradades.

8. ACKNOWLEDGEMENTS

I would like to thank my supervisors Darja Lavõgina and Asko Uri for comprehensive guidance and support.

Also, I wish to express my gratitude to SA Archimedes for the opportunity to work the lab of Structural Genomic Consortium in Oxford. My great thankfulness to Prof Stefan Knapp, Dr Alex Bullock, and Sarah Dixon-Clark for the instruction and successful collaboration.

In addition, I would like to thank my co-workers in our lab, especially Gerda Raidaru for chromatographic work.

Finally, my special thanks for Ants and Maria Silvere and Sigfried Pant Memorial Foundation for supporting my studies.

9. REFERENCES

- [1] Shugar, D. The NTP phosphate donor in kinase reactions: Is ATP a monopolist? *Acta Biochim. Pol.* **1996**, *43*, 9–24.
- [2] Adams, J. Kinetic and catalytic mechanisms of protein kinases. *Chem. Rev.* **2001**, *101*, 2271–2290.
- [3] Kinnings, S.L.; Jackson, R.M. Binding site similarity analysis for the functional classification of the protein kinase family. *J. Chem. Inf. Model.* **2009**, *49*, 318–329.
- [4] Pearce, L.R.; Komander, D.; Alessi, D.R. The nuts and bolts of AGC protein kinases. *Nat. Rev. Mol. Cell Biol.* **2010**, *11*, 9–22.
- [5] Fu, J.; Jiang, Q.; Zhang, C. Collaboration of Mitotic Kinases in Cell Cycle Control. *Nat. Ed.* **2010**, *3*, 82.
- [6] Bayliss, R.; Fry, A.; Haq, T.; Yeoh, S. On the Molecular Mechanism of Mitotic Kinase Activation. *Open. Biol.* **2012**, *2*, 120–136.
- [7] Higgins, J.M.G. Haspin: a newly discovered regulator of mitotic chromosome behavior. *Chomosome.* **2010**, *119*, 137–147.
- [8] Lavogina, D.; Kalind, K.; Bredihina, J.; Hurt, M.; Vaasa, A.; Kasari, M.; Enkvist, E.; Raidaru, G.; Uri, A. Conjugates of 5-isoquinolinesulfonylamides and oligo-D-arginine possess high affinity and selectivity towards Rho kinase (ROCK). *Bioorg. Med. Chem. Lett.* **2012**, *22*, 3425–3430.
- [9] Ekambaram, R.; Enkvist, E.; Vaasa, A.; Kasari, M.; Raidaru, G.; Knapp, S.; Uri, A. Selective Bisubstrate Inhibitors with Sub-nanomolar Affinity for Protein Kinase Pim-1. *ChemMedChem.* **2013**, *8*, 909 - 913.
- [10] Enkvist, E.; Viht, K.; Bischoff, N.; Vahter, J.; Saaver, S.; Raidaru, G.; Issinger, O.G.; Niefind, K.; Uri, A. A subnanomolar fluorescent probe for protein kinase CK2 interaction studies. *Org. Biomol. Chem.* **2012**, *10*, 8645–8653.
- [11] Enkvist, E.; Vaasa, A.; Kasari, M.; Kriisa, M.; Ivan, T.; Ligi, K.; Raidaru, G.; Uri, A. Protein-induced long lifetime luminescence of nonmetal probes. *ACS Chem. Biol.* **2011**, *6*, 1052–1062.
- [12] Holland, J.F.; Frei, E. *Cancer Medicine*, 5th ed.; BC Decker Inc, Hamilton, 2000.
- [13] Scitable by Nature Education, <http://www.nature.com/scitable/topicpage/mitosis-14046258>, Last updated on 20th March 2014
- [14] SparkNotes Editors, <http://www.sparknotes.com/biology/cellreproduction/mitosis/>, Last updated on 21th March 2014
- [15] Cyclacer Pharmaceuticals, http://www.cyclacer.com/research_science_cell-cycle.shtml, Last updated on 20th March 2014
- [16] Ma, H. T.; Poon, R. Y. C. How protein kinases co-ordinate mitosis in animal cells. *Biochem. J.* **2011**, *435*, 17–31.
- [17] Tanaka, H.; Yoshimura, Y.; Nishina, Y.; Nozaki, M.; Nojima, H.; Nishimune, Y. Isolation and characterization of cDNA clones specifically expressed in testicular germ cells. *FEBS Lett.* **1994**, *335*, 4–10.
- [18] Higgins, J. M. G. Haspin-like proteins: A new family of evolutionarily conserved putative eukaryotic protein kinases. *Protein Sci.* **2001**, *10*, 1677–1684.
- [19] Eswaran, J.; Patnaik, D.; Filippakopoulos, P.; Wang, F.; Stein, R.L.; Murray, J.W.; Higgins, J.M.G.; Knapp, S. Structure and functional characterization of the atypical human kinase haspin. *Proc. Natl. Acad. Sci.* **2009**, *106* (48), 20198–20203.
- [20] Dai, J.; Sultan, S.; Taylor, S. S.; Higgins, J.M.G. The kinase haspin is required for mitotic histone H3 Thr3 phosphorylation and normal metaphase chromosome alignment. *Genes Dev.* **2005**, *19*, 472–488.
- [21] Higgins, J. M. G. Structure, function and evolution of haspin and haspin-related proteins, a distinctive group of eukaryotic protein kinases. *Cell. Mol. Life Sci.* **2003**, *60*, 446–462.
- [22] Johnson, D.A.; Akamine, P.; Radzio-Andzelm, E.; Madhusudan, M.; Taylor, S. S. Dynamics of cAMP-dependent protein kinase. *Chem. Rev.* **2001**, *101*, 2243–2270.
- [23] Maiolica, A.; Redondo, M. D. M.; Schoof, E. M.; Chaikuad, A.; Villa, F.; Gatti, M.; Penego, L.; Jeganatgan, S.; Lou, H. J.; Novy, K.; Hauri, S.; Toprak, U. H.; Herzog, F.; Meraldi, P.; Benjamin, E. T.; Knapp, S.; Linding, R.; Aebersold, R. Modulation of the chromatin phosphoproteome by the Haspin protein kinase. *ASBMB*. (in press, online from 14.04.2014, <http://www.mcponline.org/content/early/2014/04/14/mcp.M113.034819.abstract>
- [24] Patnaik, D.; Xin, J.; Glicksman, M. A.; Cuny, G. D.; Stein, R. L.; Higgins, J. M. G. Identification of Small Molecule Inhibitors of the Mitotic Kinase Haspin by High-Throughput Screening Using a Homogeneous Time-Resolved Fluorescence. *J. Biomol. Screen.* **2008**, *13*, 1025–1034.
- [25] Ghenoiu, C.; Wheelock, M.S.; Funabiki, H. Autoinhibition and Polo-Dependent Multisite Phosphorylation Restrict Activity of the Histone H3 Kinase Haspin to Mitosis. *Mol. Cell.* **2013**, *52*, 734–745.

- [26] Parang, K.; Sun, G. Design strategies for protein kinase inhibitors. *Curr. Opin. Drug. Di. De.* **2004**, *7*, 617–629.
- [27] Fischer, P.M. The design of drug candidate molecules as selective inhibitors of therapeutically relevant protein kinases. *Curr. Med. Chem.* **2004**, *11*, 1563–1583.
- [28] Scapin, G. Structural biology in drug design: selective protein kinase inhibitors. *Drug Discov. Today.* **2002**, *7*, 601–611.
- [29] Balzano, D.; Santaguida, S.; Musacchio, A.; Villa, F. A general framework for inhibitor resistance in protein kinases. *Chem. Biol.* **2011**, *18*, 966–975.
- [30] Parkinson, F.E.; Geiger, J.D. Effects of iodotubercidin on adenosine kinase activity and nucleoside transport in DDT1 MF-2 smooth muscle cells. *J. Pharmacol. Exp. Ther.* **1996**, *277*, 1397–1401.
- [31] Cuny, G. D.; Robin, M.; Ulyanova, N. P.; Patnaik, D.; Pique, V.; Casano, G.; Liu, J.; Lin, X.; Xian, J.; Glicksman, M. A.; Stein, R. L.; Higgins, J. M. G. Structure–activity relationship study of acridine analogs as haspin and DYRK2 kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 3491–3494.
- [32] Cuny, G. D.; Robin, M.; Ulyanova, N. P.; Patnaik, D.; Pique, V.; Casano, G.; Liu, J.; Lin, X.; Auerbach, K.; Ray, S. S.; Xian, J.; Glicksman, M. A.; Stein, R. L.; Higgins, J. M. G. Structure–activity relationship study of beta-carboline derivatives analogs as haspin kinase inhibitors. *Bioorg. Med. Chem. Lett.* **2012**, *22*, 2015–2019.
- [33] Eldar-Finkelman, H.; Eisenstein, M. Peptide inhibitors targeting protein kinases. *Curr. Pharm. Design.* **2009**, *15*, 2463–2470.
- [34] Masterson, L.R.; Mascioni, A.; Traaseth, N.J.; Taylor, S.S.; Veglia, G. Allosteric cooperativity in protein kinase A. *PNAS.* **2008**, *105*, 506–511.
- [35] Parang, K.; Sun, G. Design strategies for protein kinase inhibitors. *Curr. Opin. Drug Di. De.* **2004**, *7*, 617–629.
- [36] Bogoyevitch, M.A.; Barr, R.K.; Ketterman, A.J. Peptide inhibitors of protein kinases - discovery, characterisation and use. *Biochim. Biophys. Acta.* **2005**, *1754*, 79–99.
- [37] Guichard, G.; Benkirane, N.; Zeder-Lutz, G.; van Regenmortel, M.H.; Briand, J.P.; Muller, S. Antigenic mimicry of natural L-peptides with retro-inverso-peptidomimetics. *PNAS.* **1994**, *91*, 9765–9769.
- [38] Wolfenden, R. Analog approaches to the structure of the transition state in enzyme reactions. **1972**, *5*, 10–18.
- [39] Lienhard, G.E.; Secemski, I.I. P₁P₅-Di(adenosine-5')pentaphosphate, a potent multisubstrate inhibitor of adenylate kinase. *J. Biol. Chem.* **1973**, *248*, 1121–1123.
- [40] Lavogina, D.; Enkvist, E.; Uri, A. Bisubstrate inhibitors of protein kinases: from principle to practical applications. *ChemMedChem.* **2010**, *5*, 23–34.
- [41] Saxty, G.; Woodhead, S.J.; Berdini, V.; Davies, T.G.; Verdonk, M.L.; Wyatt, P.G.; Boyle, R.G.; Barford, D.; Downham, R.; Garrett, M.D.; Carr, R.A. Identification of inhibitors of protein kinase B using fragment-based lead discovery. *J. Med. Chem.* **2007**, *50*, 2293–2296.
- [42] Kane, R.S. Thermodynamics of multivalent interactions: influence of the linker. *Langmuir.* **2010**, *26*, 8636–8640.
- [43] Vaasa, A.; Lust, M.; Terrin, A.; Uri, A.; Zaccolo, M. Small-molecule FRET probes for protein kinase activity monitoring in living cells. *Biochem. Bioph. Res. Co.* **2010**, *397*, 750–755.
- [44] Uri, A.; Lust, M.; Vaasa, A.; Lavogina, D.; Viht, K.; Enkvist, E. Bisubstrate fluorescent probes and biosensors in binding assays for HTS of protein kinase inhibitors. *Biochim. Biophys. Acta.* **2010**, *1804*, 541–546.
- [45] Räägel, H.; Lust, M.; Uri, A.; Pooga, M. Adenosine-oligoarginine conjugate, a novel bisubstrate inhibitor, effectively dissociates the actin cytoskeleton. *FEBS J.* **2008**, *275*, 3608–3624.
- [46] Nakase, I.; Takeuchi, T.; Tanaka, G.; Futaki, S. Methodological and cellular aspects that govern the internalization mechanisms of arginine-rich cell-penetrating peptides. *Adv. Drug Deliver. Rev.* **2008**, *60*, 598–607.
- [47] Vaasa, A.; Ligi, K.; Mohandessi, S.; Enkvist, E.; Uri, A.; Miller, L.W. Time-gated luminescence microscopy with responsive nonmetal probes for mapping activity of protein kinases in living cells. *Chem. Commun.* **2012**, *48*, 8595–8597.
- [48] Enkvist, E.; Lavogina, D.; Raidaru, G.; Vaasa, A.; Viil, I.; Lust, M.; Viht, K.; Uri, A. Conjugation of adenosine and hexa-(D-arginine) leads to a nanomolar bisubstrateanalog inhibitor of basophilic protein kinases. *J. Med. Chem.* **2006**, *49*, 7150–7159.
- [49] Enkvist, E.; Kriisa, M.; Roben, M.; Kadak, G.; Raidaru, G.; Uri, A. Effect of the structure of adenosine mimic of bisubstrate-analog inhibitors on their activity towards basophilic protein kinases. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 6098–6101.

- [50] Vaasa, A.; Viil, I.; Enkvist, E.; Viht, K.; Raidaru, G.; Lavogina, D.; Uri, A. High affinity bisubstrate probe for fluorescence anisotropy binding/displacement assays with protein kinases PKA and ROCK. *Anal. Biochem.* **2009**, *385*, 85–93.
- [51] Kasari, M.; Padrik, P.; Vaasa, A.; Saar, K.; Leppik, K.; Soplepmann, J.; Uri, A. Timegated luminescence assay using nonmetal probes for determination of protein kinase activity-based disease markers. *Anal. Biochem.* **2012**, *422*, 79–88.
- [52] Lavogina, D.; Lust, M.; Viil, I.; König, N.; Raidaru, G.; Rogozina, J.; Enkvist, E.; Uri, A.; Bossemeyer, D. Structural analysis of ARC-type inhibitor (ARC-1034) binding to protein kinase A catalytic subunit and rational design of bisubstrate analogue inhibitors of basophilic protein kinases. *J. Med. Chem.* **2009**, *52*, 308–321.
- [53] Pflug, A.; Rogozina, J.; Lavogina, D.; Enkvist, E.; Uri, A.; Engh, R.A.; Bossemeyer, D. Diversity of bisubstrate binding modes of adenosine analogue-oligoarginine conjugates in protein kinase a and implications for protein substrate interactions. *J. Mol. Biol.* **2010**, *403*, 66–77.
- [54] Lavogina, D.; Nickl, C. K.; Enkvist, E.; Raidaru, G.; Lust, M.; Vaasa, A.; Uri, A.; Dostmann, W. R. Adenosine analogue-oligo-arginine conjugates (ARCs) serve as high-affinity inhibitors and fluorescence probes of type I cGMP-dependent protein kinase (PKGIalpha). *Biochim. Biophys. Acta.* **2010**, *1804*, 1857 - 1868.
- [55] Merrifield, R.B. Solid Phase Peptide Synthesis. I. The Synthesis of tetrapeptide. *J. Am.Chem. Soc.* **1963**, *85*, 2149–2154.
- [56] Nelson, D.L.; Cox, M.M. *Lehninger Principles of Biochemistry*, 5th ed; Freeman: New York, 2005, pp. 104–106.
- [57] Attardi, M.E.; Falchi, A.; Taddei, M. A sensitive visual test for detection of OH groups on resin. *Tetrahedron Lett.* **2000**, *41*, 7395–7399.
- [58] Mařík, J.; Song, A.; Lam, K.S. Detection of primary aromatic amines on solid phase. *Tetrahedron Lett.* **2003**, *44*, 4319–4320.
- [59] Kaiser, E.; Colescott, R.L.; Bossinger, C.D.; Cook, P.I. Color test for detection of free terminal amino groups in the solid-phase synthesis of peptides. *Anal. Biochem.* **1970**, *34*, 595–598.
- [60] Niesen, F.H.; Berglund, H.; Vedadi, M. The use of differential scanning fluorimetry to detect ligand interactions that promote protein stability. *Nat Protoc.* **2007**, *2*, 2212 – 2221.
- [61] Lea, W.A.; Simeonov, A. Differential Scanning Fluorometry as Indicators of Enzyme Inhibitor Mode of Action: Case Study of Glutathione S-Transferase. *PloS ONE.* **2013**, *7*, 1-9.
- [62] Bullock, A.N.; Debreczeni, J.E.; Fedorov, O.Y.; Nelson, A.; Marsden, B.D.; Knapp, S. Structural basis of inhibitor specificity of the human protooncogene proviral insertion site in moloney murine leukemia virus (PIM-1) kinase. *J. Med. Chem.* **2005**, *48*, 7604–7614.
- [63] Zubriene, A.; Matulienė, J.; Baranauskienė, L.; Jachno, J.; Torresan, J.; Michailovienė, V.; Cimmerman, P.; Matulis, D. Measurement of nanomolar dissociation constants by titration calorimetry and thermal shift assay - radicicol binding to Hsp90 and ethoxzolamide binding to CAII. *Int. J. Mol. Sci.* **2009**, *10*, 2662–2680.
- [64] Filippakopoulos, P.; Qi, J.; Picaud, S.; Shen, Y.; Smith, W. B.; Fedorov, O.; Morse, E.M.; Keates, T.; Hickman, T. T.; Felletar, I.; Philpott, M.; Munro, S; McKeown, M. R.; Wang, Y.; Christie, A.L.; West, N.; Cameron, M. J.; Schwartz, B.; Heightman, T. D.; La Thangue, N.; French, C. A.; Wiest, O.; Kung, A.L.; Knapp, S.; Bradner, J. E. Selective inhibition of BET bromodomains. *Nature.* **2010**, *468*, 1067-1073.
- [65] <http://www.proteinstructures.com/Experimental/Experimental/protein-crystallography.html>, Last updated on 3rd March 2014
- [66] Drenth, J. *Principles of Protein X-Ray Crystallography*, 3rd, ed.; Springer, New York, 2007.
- [67] <http://people.ds.cam.ac.uk/ml527/publications/assets/leunissen-literaturerecherche.pdf>, Last updated on 3rd March 2014
- [68] Flewitt, P. E.; Wild, R. K. *Physical methods for materials characterization*, 2nd, ed.; Institute of Physics Publishing: Bristol and Philadelphia, 2003.
- [69] Wlodawer, A.; Minor, W.; Dauter, Z.; Jaskolski, M. Protein crystallography for non-crystallographers, or how to get the best (but not more) from published macromolecular structures. *FEBS J.* **2008**, *275*, 1-21.
- [70] Chan, W.C.; White, P.D. *Fmoc Solid Phase Peptide Synthesis*, Oxford University Press, 2000.
- [71] Nikolovska-Coleska, Z.; Wang, R.; Fang, X.; Pan, H.; Tomita, Y.; Li, P.; Roller, P.P.; Krajewski, K.; Saito, N.G.; Stuckey, J.A.; Wang, S. Development and optimization of a binding assay for the XIAP BIR3 domain using fluorescence polarisation. *Anal. Biochem.* **2004**, *332*, 261-273.
- [72] Song, A.; Wang, X.; Zhang, J.; Marik, J.; Lebrilla, C. B.; Lam, K. S. Synthesis of hydrophilic and flexible linkers for peptide derivatization in solid phase. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 161-165.

10. APPENDICES

Appendix 1. Amino Acid Sequence of Full-Length Haspin and Recombinant His-tagged Haspin Kinase Domain; Calculated and Experimental Molecular Weight and Extinction Coefficient

Appendix 2. Gel-Filtration Chromatogram of Haspin Purification

Appendix 3. Images of SDS-PAGE Gels Stained with Coomassie

Appendix 4. Precipitating Solutions of Coarse Screen

Appendix 5. Precipitating Solutions of Fine Screen

Appendix 6. Pictures of Haspin/ARC Co-Crystals Under 3.3x Magnification

Appendix 7. Structural Factors of Co-Crystal Structures Obtained in This Work or Available in Protein Data Bank (PDB)

Appendix 8. Composition of Buffers and Reagents

Appendix 9. Structures and Codes of ARCs

Appendix 10. Chromatograms of Purification, Photodiode Array Spectrum, and ESI-MS Data of Novel ARCs

Appendix 11. Experimental Molecular Weights, Absorption Maxima, Extinction Coefficients (ϵ) of Novel ARCs

Appendix 12. Parameters Characterizing the Affinity of the Inhibitor

Appendix 13. The Examples of Binding and Displacement Curves Obtained by FA Method

Appendix 1. Amino Acid Sequence of Full-Length Haspin and Recombinant His-tagged Haspin Kinase Domain; Calculated and Experimental Molecular Weight and Extinction Coefficient

Full length PK Kinase domain with His-tag	10	20	30	40	50	60
	MAASLPGPGS	RLFRTYGAAD	GRRQRRPGRE	AAQWFPPQDR	RRFFNSSGSS	DASIGDPSQS
	70	80	90	100	110	120
	DDPDDPDDPD	FPGSPVRRRR	RCPGGRVPKD	RPSLTVTPKR	WKLRARPSLT	VTFRRLGLRA
	130	240	150	160	170	180
	RPPQKCSTPC	GPLRLPPFPS	RDSGRLSPDL	SVCGQPRDGD	ELGISASLFS	SLASPCPGSP
	190	200	210	220	230	240
	TPRDSVISIG	TSACLVAASA	VPSDLHLPEV	SLDRASLPCS	QEEATGGAKD	TRMVHQTRAS
	250	260	270	280	290	300
	LRSVLFGLMN	SGTPEDSEFR	ADGKNMRESC	CKRKLVVGNQ	PEGPGLSSTG	KRRATGQDSC
	310	320	330	340	350	360
	QERGLQEAVR	REHQEASVPK	GRIVPRGTDR	LERTRSSRES	KHQEATETSL	LHSHRFKKGQ
	370	380	390	400	410	420
	KLKDSFPTQ	DLTPLQNACF	WTKTRASFSF	HKKKIVTDVS	EVCSIYTTAT	SLSGSLLSEC
	430	440	450	460	470	480
	SNRPVMNRTS	GAPSSWHSSS	MYLLSPLNTL	SISNKKASDA	EKVYGECSQK	GPVPFSSHCLP
	490	500	510	520	530	540
	TEKLQRCEKI	GEGVFGEVFQ	TIADHTPVAI	KIIAIEGPDL	VNGSHQKTFF	EILPEIIISK
	TEKLQRCEKI	GEGVFGEVFQ	TIADHTPVAI	KIIAIEGPDL	VNGSHQKTFF	EILPEIIISK
	550	560	570	580	590	600
	ELSLLSGEVC	NRTEGFIGLN	SVHCVQGSYP	PLLLKAWDHY	NSTKGSANDR	PDDFKDDQLE
	ELSLLSGEVC	NRTEGFIGLN	SVHCVQGSYP	PLLLKAWDHY	NSTKGSANDR	PDDFKDDQLE
	610	620	630	640	650	660
	IVLEFEFFGGI	DLEQMRTKLS	SLATAKSILH	QLTASLAVAE	ASLRFEHRDL	HWGNVLLKKT
	IVLEFEFFGGI	DLEQMRTKLS	SLATAKSILH	QLTASLAVAE	ASLRFEHRDL	HWGNVLLKKT
	670	680	690	700	710	720
	SLKKLHYTLN	GKSSTIPSCG	LQVSIIDYTL	SRLERDGIIV	FCDVSMDEDL	FTGDGDYQFD
	SLKKLHYTLN	GKSSTIPSCG	LQVSIIDYTL	SRLERDGIIV	FCDVSMDEDL	FTGDGDYQFD
	730	740	750	760	770	780
	IYRLMKKENN	NRWGEYHPYS	NVLWLHYLTD	KMLKQMTFKT	KCNTPAMKQI	KRKIQEFHRT
	IYRLMKKENN	NRWGEYHPYS	NVLWLHYLTD	KMLKQMTFKT	KCNTPAMKQI	KRKIQEFHRT
	790					
	MLNFSSATDL	LCQHSFLFK				
	MLNFSSATDL	LCQHSFLFK				

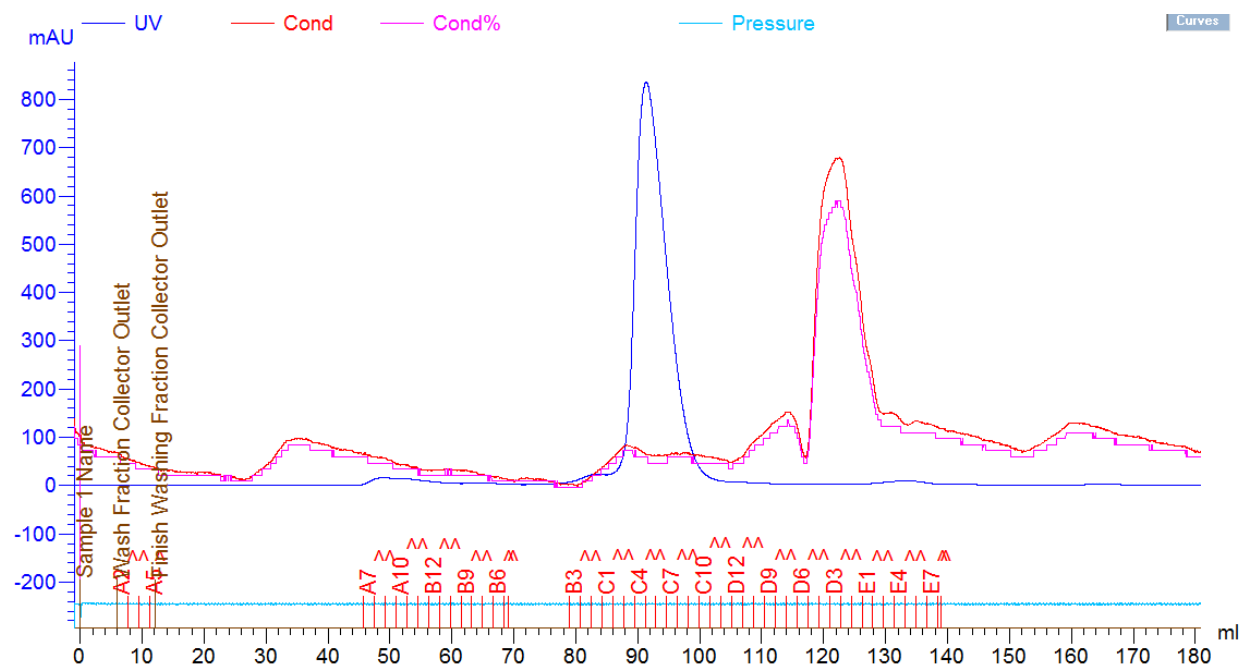
Calculated molecular weight of kinase domain (*tagged*) = 40654.6 Da

Experimental molecular weight of kinase domain (*tagged*) = 40657.8 Da

Calculated extinction coefficient of kinase domain (*tagged*) = 36900

Kinase domain is marked with yellow or green; His-tag is marked with red.

Appendix 2. Gel-Filtration Chromatogram of Haspin Purification



UV – UV-signal detection at 280 nm (mAU)

Cond – conductivity (mS/cm)

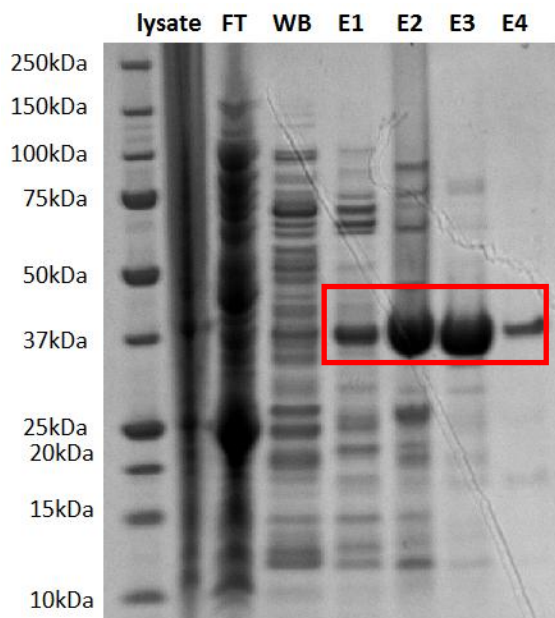
Cond% - relative conductivity signal (mS/cm)

Pressure – pressure

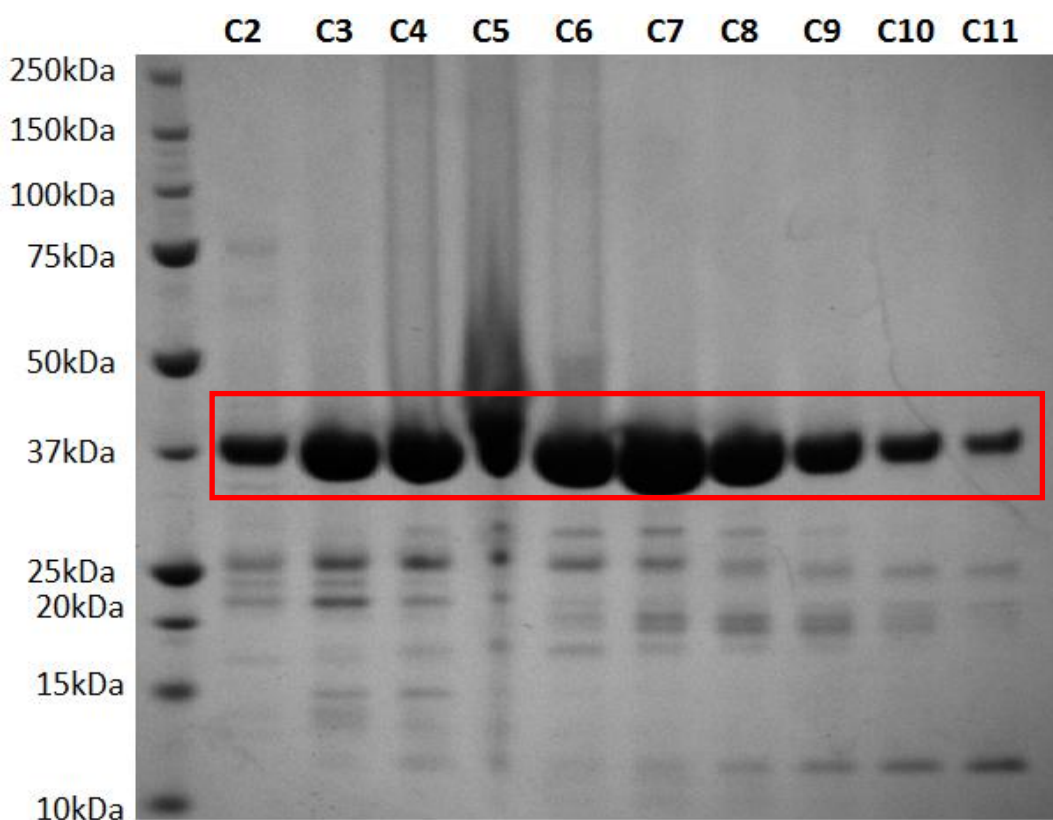
A1 – E7 – collected fractions

Appendix 3. Images of SDS-PAGE Gels Stained with Coomassie

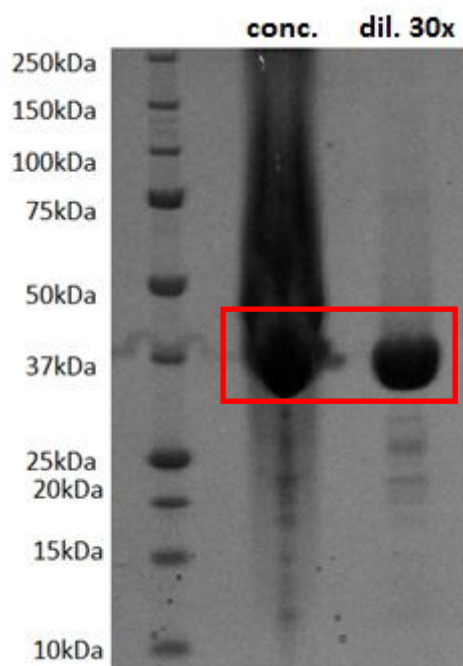
1. After Ni-NTA column (lysate: lysate after sonification; FT: flow through; WB: wash buffer aliquote; E1: elution buffer I aliquote; E2: elution buffer II aliquote; E3: elution buffer III aliquote; E4: elution buffer IV aliquote). The bands containing Haspin are surrounded by red box.



2. After gel-filtration chromatography (fractions C2-C11 are shown which contained the protein). The bands containing Haspin are surrounded by red box.



3. Protein used for crystallization trials (conc. – concentrated protein sample; dil. 30x - 30-fold diluted protein sample). The bands containing Haspin are surrounded by red box.



Appendix 4. Precipitating Solutions of Coarse Screen

LFS4

	A	B	C	D	E	F	G	H
1	30% PEG1k 0.1M SPG pH 6	30% PEG1k 0.1M MIB pH 6	30% PEG1k 0.1M PCB pH 6	30% PEG1k 0.1M MMT pH 6	20% PEG3.3K 10% EtGly 0.2M NaF	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaF	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaF	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaF
2	30% PEG1k 0.1M SPG pH 7	30% PEG1k 0.1M MIB pH 7	30% PEG1k 0.1M PCB pH 7	30% PEG1k 0.1M MMT pH 7	20% PEG3.3K 10% EtGly 0.2M NaBr	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaBr	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaBr	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaBr
3	30% PEG1k 0.1M SPG pH 8	30% PEG1k 0.1M MIB pH 8	30% PEG1k 0.1M PCB pH 8	30% PEG1k 0.1M MMT pH 8	20% PEG3.3K 10% EtGly 0.2M NaI	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaI	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaI	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaI
4	60% MPD 0.1M SPG pH 6	60% MPD 0.1M MIB pH 6	60% MPD 0.1M PCB pH 6	60% MPD 0.1M MMT pH 6	20% PEG3.3K 10% EtGly 0.2M KSCN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M KSCN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M KSCN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M KSCN
5	60% MPD 0.1M SPG pH 7	60% MPD 0.1M MIB pH 7	60% MPD 0.1M PCB pH 7	60% MPD 0.1M MMT pH 7	20% PEG3.3K 10% EtGly 0.2M NaNO ₃	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaNO ₃	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaNO ₃	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaNO ₃
6	60% MPD 0.1M SPG pH 8	60% MPD 0.1M MIB pH 8	60% MPD 0.1M PCB pH 8	60% MPD 0.1M MMT pH 8	20% PEG3.3K 10% EtGly 0.2M NaForm	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaForm	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaForm	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaForm
7	20% PEG6k 10% EtGly 0.2M NaCl	20% PEG6k 10% EtGly 0.1M MES pH 6 0.2M NaCl	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.2M NaCl	20% PEG6k 10% EtGly 0.1M tris pH 8 0.2M NaCl	20% PEG3.3K 10% EtGly 0.2M NaAc	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaAc	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaAc	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaAc
8	20% PEG6k 10% EtGly 0.2M AmCl	20% PEG6k 10% EtGly 0.1M MES pH 6 0.2M AmCl	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.2M AmCl	20% PEG6k 10% EtGly 0.1M tris pH 8 0.2M AmCl	20% PEG3.3K 10% EtGly 0.2M Na ₂ SO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M Na ₂ SO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M Na ₂ SO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M Na ₂ SO ₄
9	20% PEG6k 10% EtGly 0.2M LiCl	20% PEG6k 10% EtGly 0.1M MES pH 6 0.2M LiCl	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.2M LiCl	20% PEG6k 10% EtGly 0.1M tris pH 8 0.2M LiCl	20% PEG3.3K 10% EtGly 0.2M NaKTART	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaKTART	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaKTART	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaKTART
10	20% PEG6k 10% EtGly 0.1M MgCl ₂	20% PEG6k 10% EtGly 0.1M MES pH 6 0.1M MgCl ₂	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.1M MgCl ₂	20% PEG6k 10% EtGly 0.1M tris pH 8 0.1M MgCl ₂	20% PEG3.3K 10% EtGly 0.2M Na/KPO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M Na/KPO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M Na/KPO ₄	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M Na/KPO ₄
11	20% PEG6k 10% EtGly 0.1M CaCl ₂	20% PEG6k 10% EtGly 0.1M MES pH 6 0.1M CaCl ₂	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.1M CaCl ₂	20% PEG6k 10% EtGly 0.1M tris pH 8 0.1M CaCl ₂	20% PEG3.3K 10% EtGly 0.2M KCit	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M KCit	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M KCit	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M KCit
12	20% PEG6k 10% EtGly 0.01M ZnCl ₂	20% PEG6k 10% EtGly 0.1M MES pH 6 0.01M ZnCl ₂	20% PEG6k 10% EtGly 0.1M HEPES pH 7 0.01M ZnCl ₂	20% PEG6k 10% EtGly 0.1M tris pH 8 0.01M ZnCl ₂	20% PEG3.3K 10% EtGly 0.2M NaMLN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 6.5 0.2M NaMLN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 7.5 0.2M NaMLN	20% PEG3.3K 10% EtGly 0.1M bis-tris- propane pH 8.5 0.2M NaMLN

Abbreviations:

AmCl – ammonium chloride

EtGly – ethylene glycole

KCit – potassium citrate

MIB – malonate:imidazole:borate

MMT – malate:MES:TRIS

MPD - 2-methyl-2,4-pentanediol

NaAc – sodium acetate

NaForm - sodium formate

NaKTRAT – sodium potassium tartrate

NaMLN – sodium malonate

PCB – propionate:cacodylate:bis-tris-propane

PEG – polyethylene glycole

SPG – succinate:phosphate:glycine

Appendix 5. Precipitating Solutions of Fine Screen

LFS4-MPD-sp

	A	B	C	D	E	F	G	H
1	51% MPD 0.1M SPG pH 6	51% MPD 0.1M SPG pH 6.5	51% MPD 0.1M SPG pH 7	51% MPD 0.1M SPG pH 7.5	51% MPD 0.1M PCB pH 6	51% MPD 0.1M PCB pH 6.5	51% MPD 0.1M PCB pH 7	51% MPD 0.1M PCB pH 7.5
2	54% MPD 0.1M SPG pH 6	54% MPD 0.1M SPG pH 6.5	54% MPD 0.1M SPG pH 7	54% MPD 0.1M SPG pH 7.5	54% MPD 0.1M PCB pH 6	54% MPD 0.1M PCB pH 6.5	54% MPD 0.1M PCB pH 7	54% MPD 0.1M PCB pH 7.5
3	57% MPD 0.1M SPG pH 6	57% MPD 0.1M SPG pH 6.5	57% MPD 0.1M SPG pH 7	57% MPD 0.1M SPG pH 7.5	57% MPD 0.1M PCB pH 6	57% MPD 0.1M PCB pH 6.5	57% MPD 0.1M PCB pH 7	57% MPD 0.1M PCB pH 7.5
4	60% MPD 0.1M SPG pH 6	60% MPD 0.1M SPG pH 6.5	60% MPD 0.1M SPG pH 7	60% MPD 0.1M SPG pH 7.5	60% MPD 0.1M PCB pH 6	60% MPD 0.1M PCB pH 6.5	60% MPD 0.1M PCB pH 7	60% MPD 0.1M PCB pH 7.5
5	63% MPD 0.1M SPG pH 6	63% MPD 0.1M SPG pH 6.5	63% MPD 0.1M SPG pH 7	63% MPD 0.1M SPG pH 7.5	63% MPD 0.1M PCB pH 6	63% MPD 0.1M PCB pH 6.5	63% MPD 0.1M PCB pH 7	63% MPD 0.1M PCB pH 7.5
6	54% MPD 0.1M SPG pH 8	57% MPD 0.1M SPG pH 8	60% MPD 0.1M SPG pH 8	63% MPD 0.1M SPG pH 8	54% MPD 0.1M PCB pH 8	57% MPD 0.1M PCB pH 8	60% MPD 0.1M PCB pH 8	63% MPD 0.1M PCB pH 8
7	51% MPD 0.1M MIB pH 6	51% MPD 0.1M MIB pH 6.5	51% MPD 0.1M MIB pH 7	51% MPD 0.1M MIB pH 7.5	51% MPD 0.1M MMT pH 6	51% MPD 0.1M MMT pH 6.5	51% MPD 0.1M MMT pH 7	51% MPD 0.1M MMT pH 7.5
8	54% MPD 0.1M MIB pH 6	54% MPD 0.1M MIB pH 6.5	54% MPD 0.1M MIB pH 7	54% MPD 0.1M MIB pH 7.5	54% MPD 0.1M MMT pH 6	54% MPD 0.1M MMT pH 6.5	54% MPD 0.1M MMT pH 7	54% MPD 0.1M MMT pH 7.5
9	57% MPD 0.1M MIB pH 6	57% MPD 0.1M MIB pH 6.5	57% MPD 0.1M MIB pH 7	57% MPD 0.1M MIB pH 7.5	57% MPD 0.1M MMT pH 6	57% MPD 0.1M MMT pH 6.5	57% MPD 0.1M MMT pH 7	57% MPD 0.1M MMT pH 7.5
10	60% MPD 0.1M MIB pH 6	60% MPD 0.1M MIB pH 6.5	60% MPD 0.1M MIB pH 7	60% MPD 0.1M MIB pH 7.5	60% MPD 0.1M MMT pH 6	60% MPD 0.1M MMT pH 6.5	60% MPD 0.1M MMT pH 7	60% MPD 0.1M MMT pH 7.5
11	63% MPD 0.1M MIB pH 6	63% MPD 0.1M MIB pH 6.5	63% MPD 0.1M MIB pH 7	63% MPD 0.1M MIB pH 7.5	63% MPD 0.1M MMT pH 6	63% MPD 0.1M MMT pH 6.5	63% MPD 0.1M MMT pH 7	63% MPD 0.1M MMT pH 7.5
12	54% MPD 0.1M MIB pH 8	57% MPD 0.1M MIB pH 8	60% MPD 0.1M MIB pH 8	63% MPD 0.1M MIB pH 8	54% MPD 0.1M MMT pH 8	57% MPD 0.1M MMT pH 8	60% MPD 0.1M MMT pH 8	63% MPD 0.1M MMT pH 8

Abbreviations:

MIB – malonate:imidazole:borate

MMT – malate:MES:TRIS

MPD - 2-methyl-2,4-pentanediol

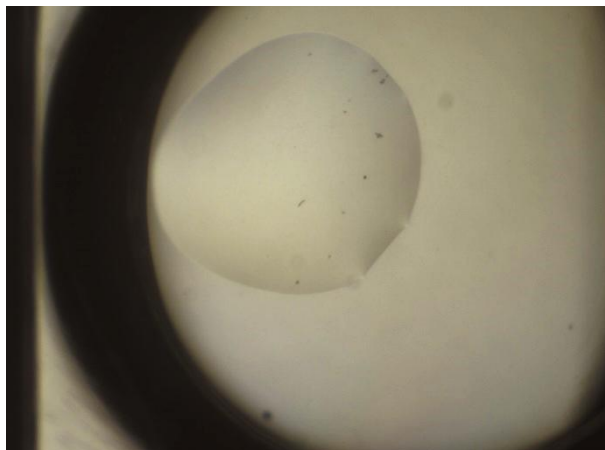
PCB – propionate:cacodylate:bis-tris-propane

SPG – succinate:phosphate:glycine

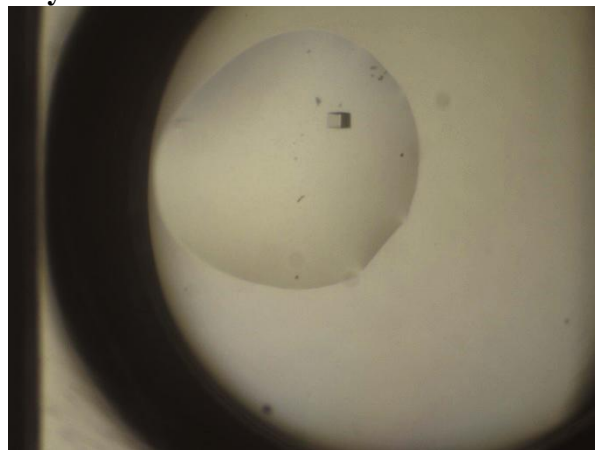
Appendix 6. Pictures of Haspin/ARC Co-Crystals Under 3.3x Magnification

Haspin/ARC-1411 co-crystal

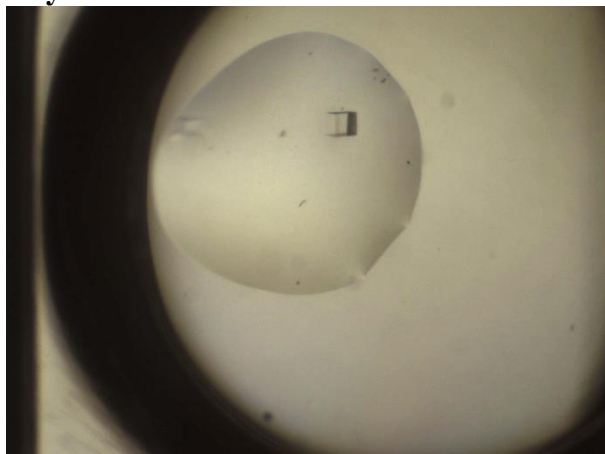
11 hours



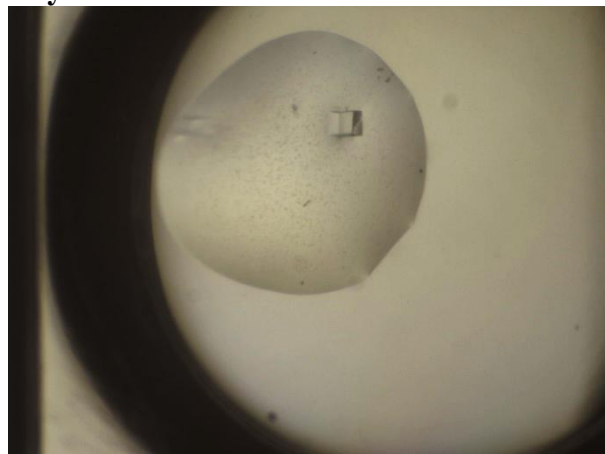
Day 2



Day 4



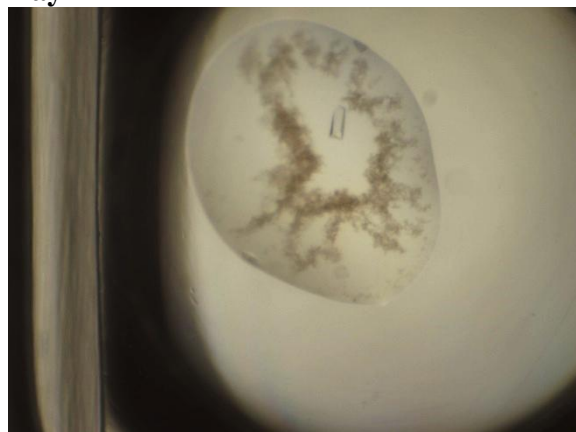
Day 9



11 hours



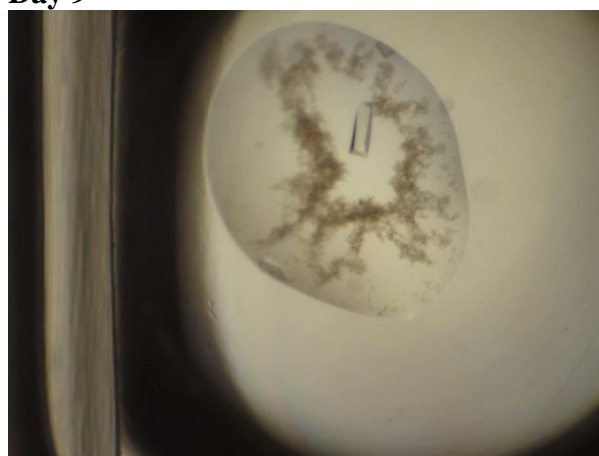
Day 2



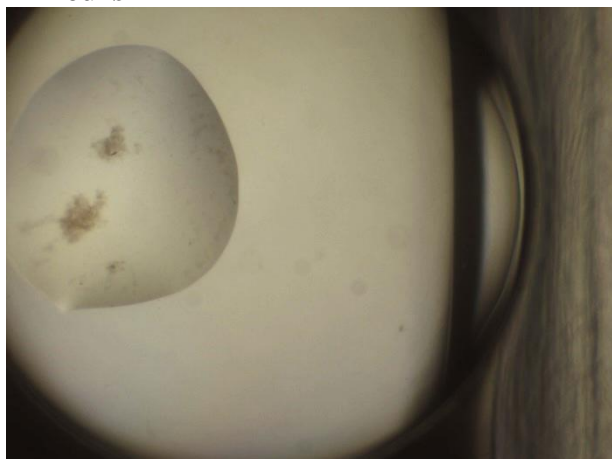
Day 4



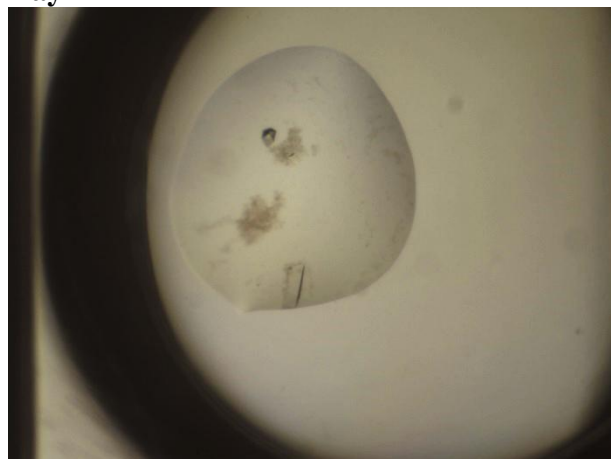
Day 9



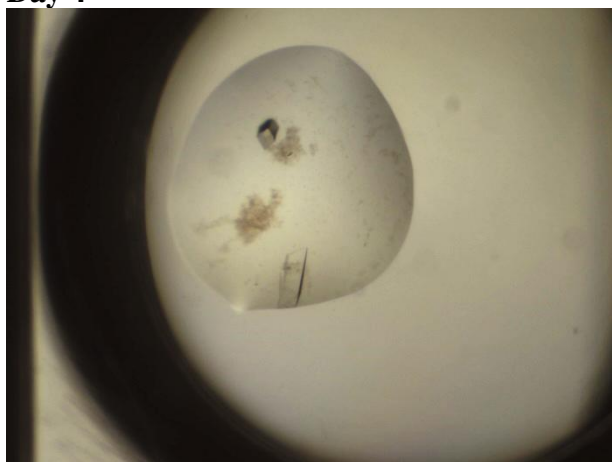
11 hours



Day 2



Day 4



Day 9



Appendix 7. Structural Factors of Co-Crystal Structures Obtained in This Work or Available in Protein Data Bank (PDB)

Characteristics	ARC-0902/ Haspin	ARC-1141/ Haspin	5-iodo- tubercidin/ Haspin	H3(1-7)/ Haspin	AMP/ Haspin	ARC-1034/ PKAc	ATP+PKI/ PKAc
PDB code	not published	not published	3IQ7	4OUC	3DLZ	3BWJ	1ATP
Resolution (Å)	1.70	1.87	2.00	1.90	1.85	2.30	2.20
R-value	0.153	0.160	0.174	0.166	0.167	0.201	0.177
Space group	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁	P 2 ₁ 2 ₁ 2 ₁

Appendix 8. Composition of Buffers and Reagents

Binding buffer: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 5mM imidazole

Wash buffer: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 30 mM imidazole

Elution buffer I: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 50 mM imidazole

Elution buffer II: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 100 mM imidazole

Elution buffer III: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 150 mM imidazole

Elution buffer IV: 500 mM NaCl, 50 mM HEPES pH 7.5, 5% glycerol, 250 mM imidazole

Gel-filtration buffer: 300 mM NaCl, 50 mM HEPES pH 7.5, 1 mM TCEP

FA buffer: 50 mM Hepes (pH 7.5), 150 mM NaCl, 5 mM DTT (7.5 μ M BSA)

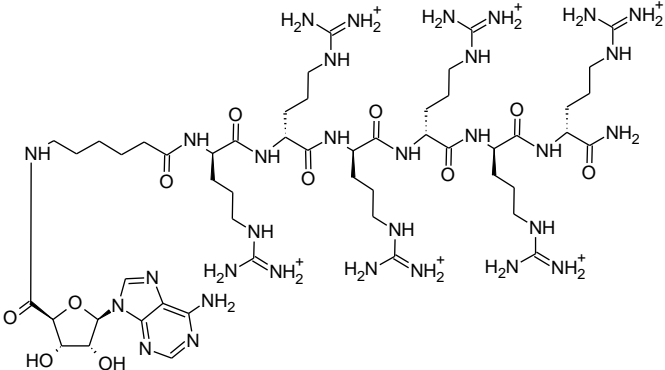
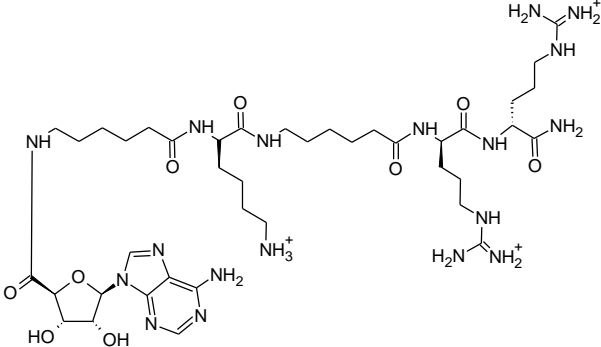
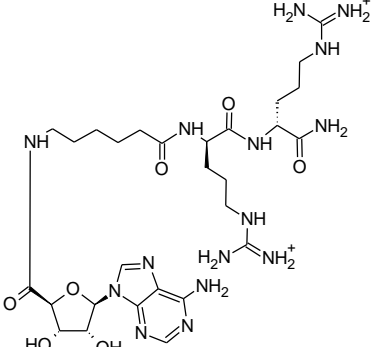
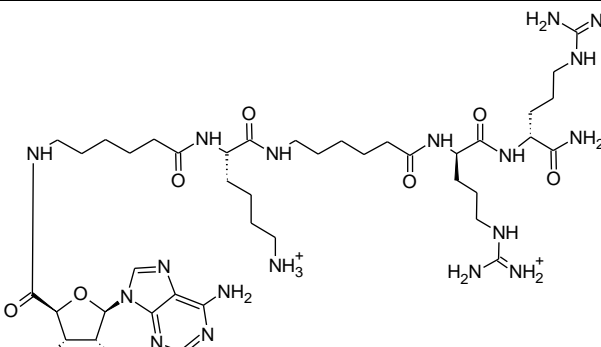
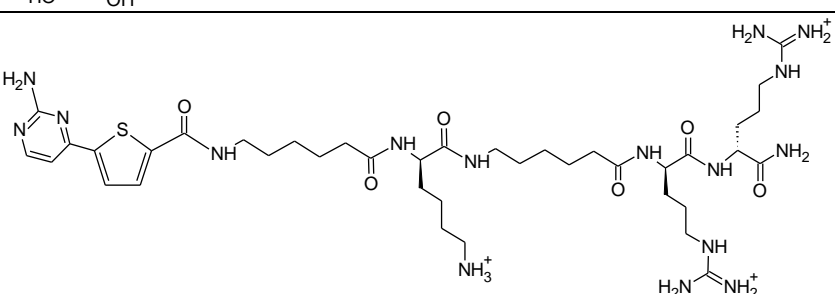
Thermal shift assay buffer: 10 mM Hepes pH 7.5, 500 mM NaCl

Loading buffer with reducing agent: 0.25 M Tris-HCl pH 6.8, 10% sodium dodecyl sulfate, 50% glycerol, 0.5 M DTT. 0.25% bromophenol blue

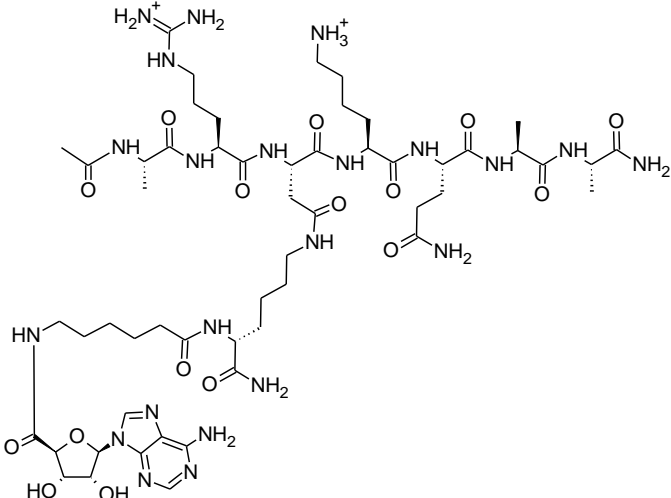
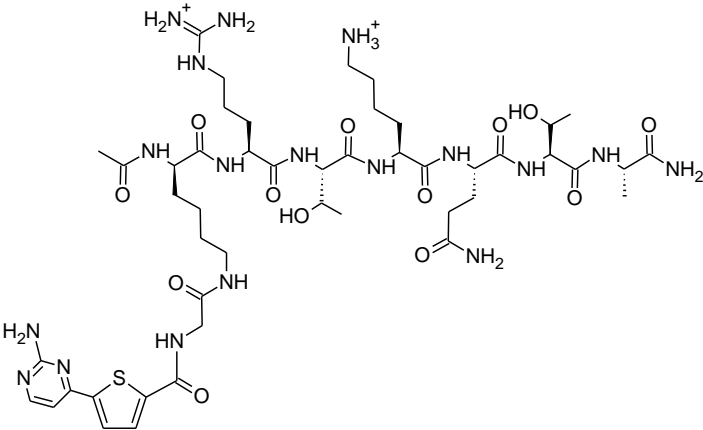
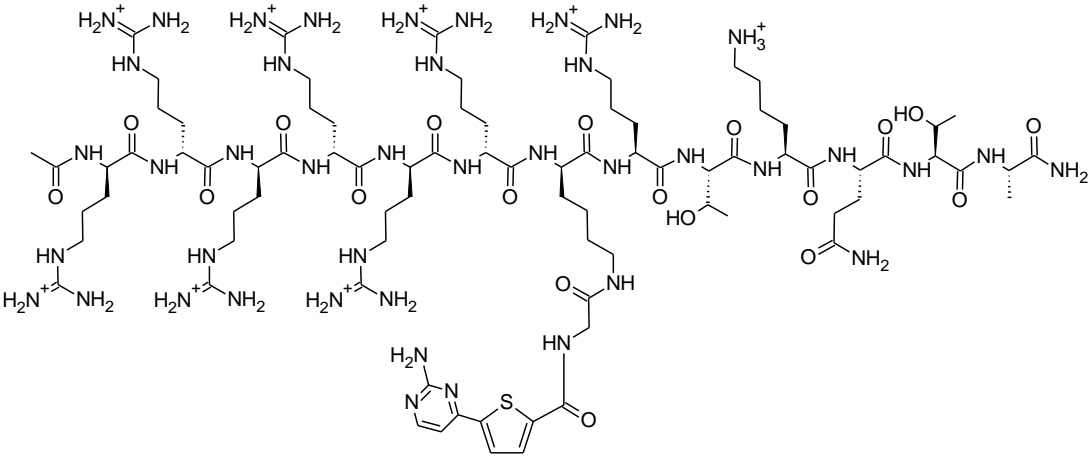
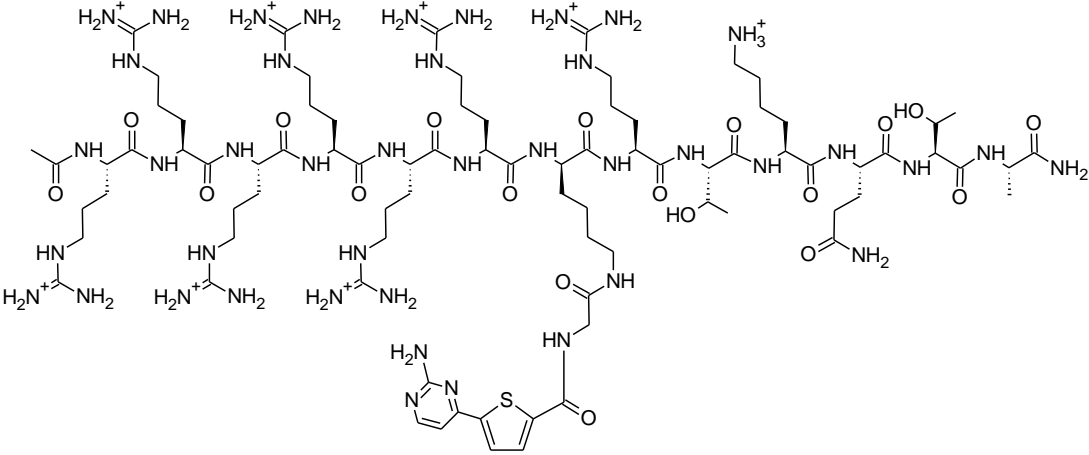
LB Broth, Miller (Fisher Scientific): 10 g tryptone, 5 g yeast extract and 10 g sodium chloride per 1 L of water

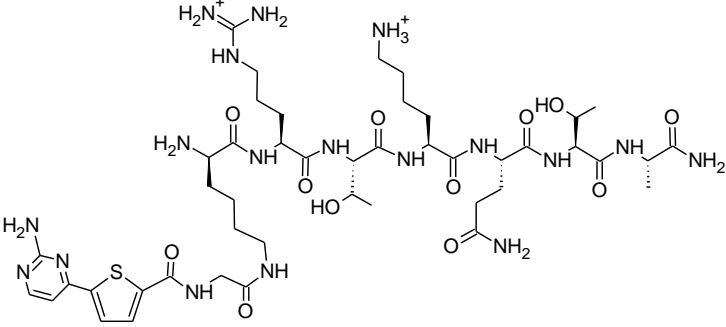
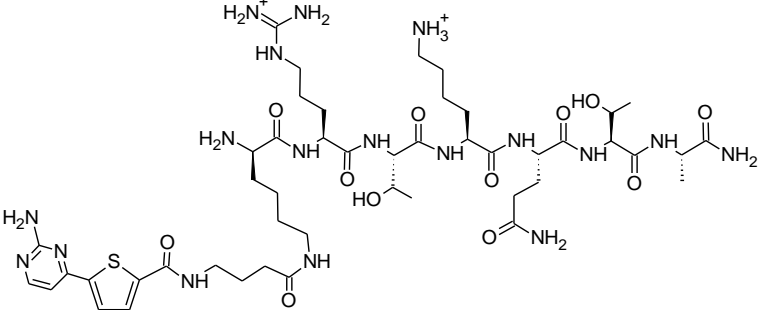
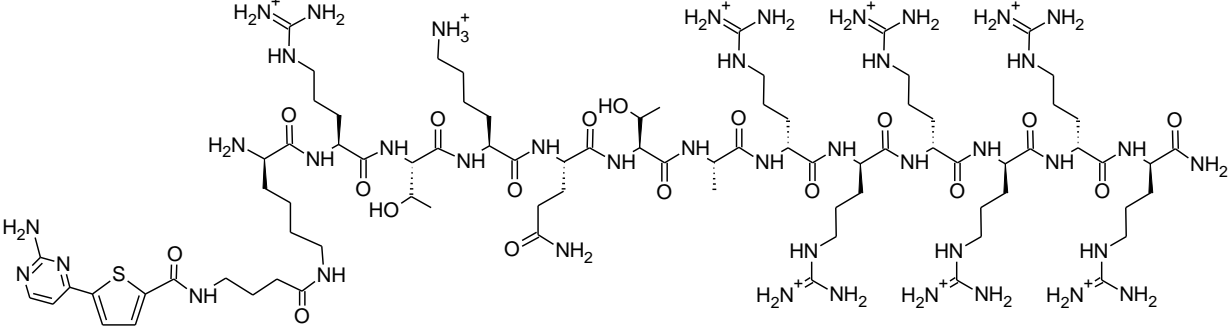
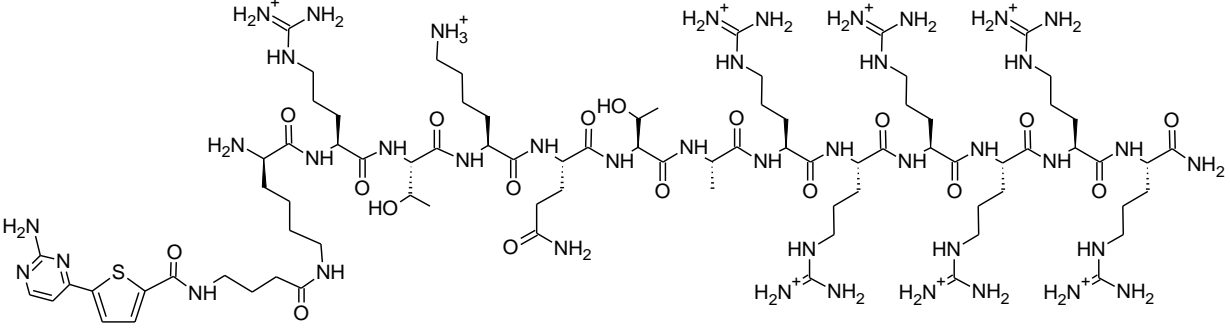
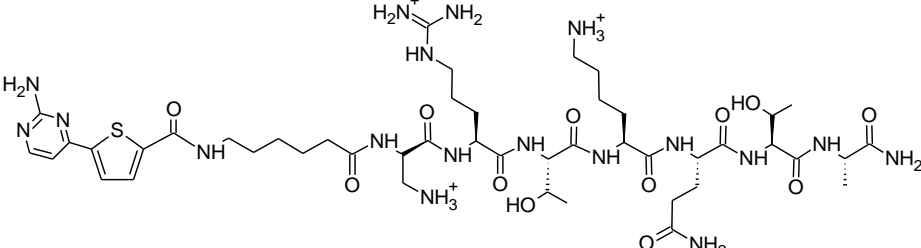
Appendix 9. Structures and Codes of ARCs

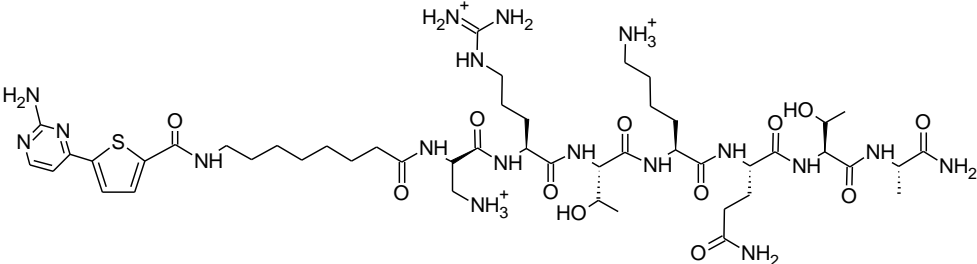
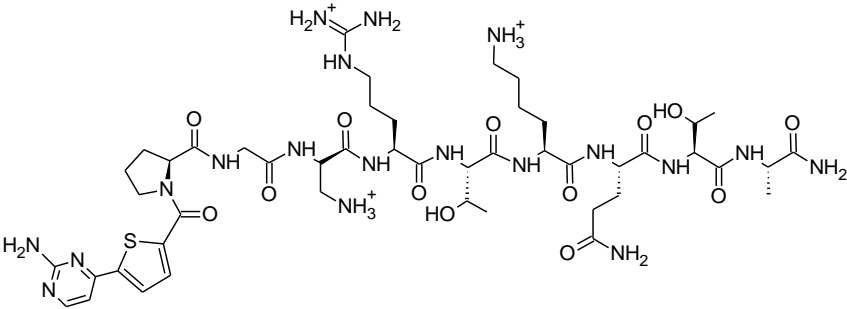
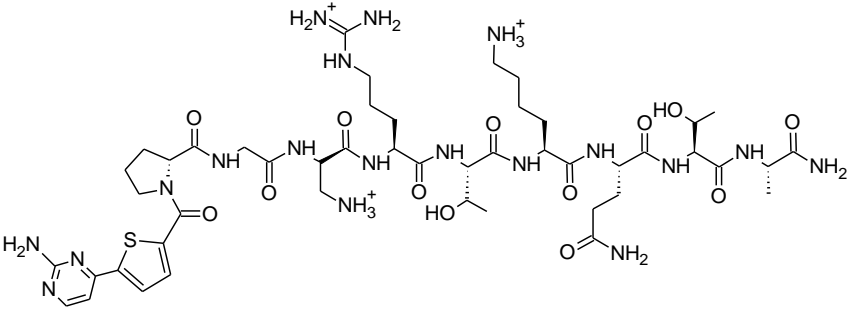
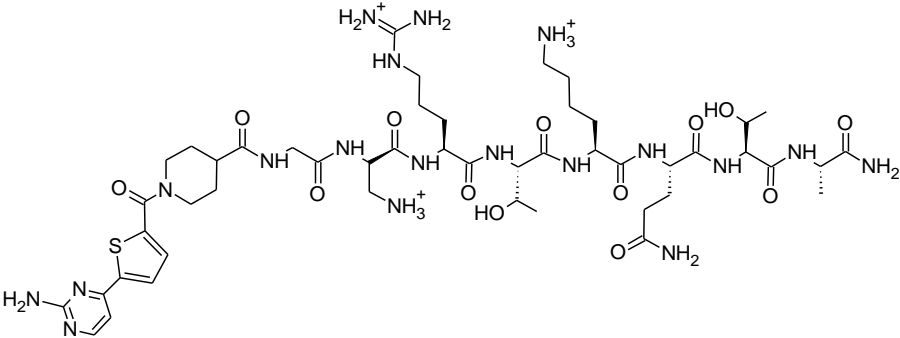
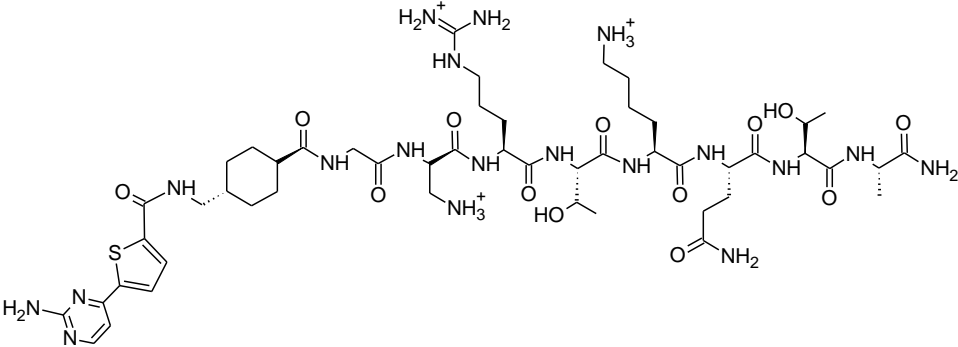
<i>Kood</i>	<i>Struktuur</i>
ARC-0341	
ARC-0668	
ARC-0684	
ARC-0685	

ARC-0902	
ARC-1012	
ARC-1034	
ARC-1038	
ARC-1102	

ARC-1141	
ARC-1408	
ARC-1411	
ARC-3009	
ARC-3010	
ARC-3125	

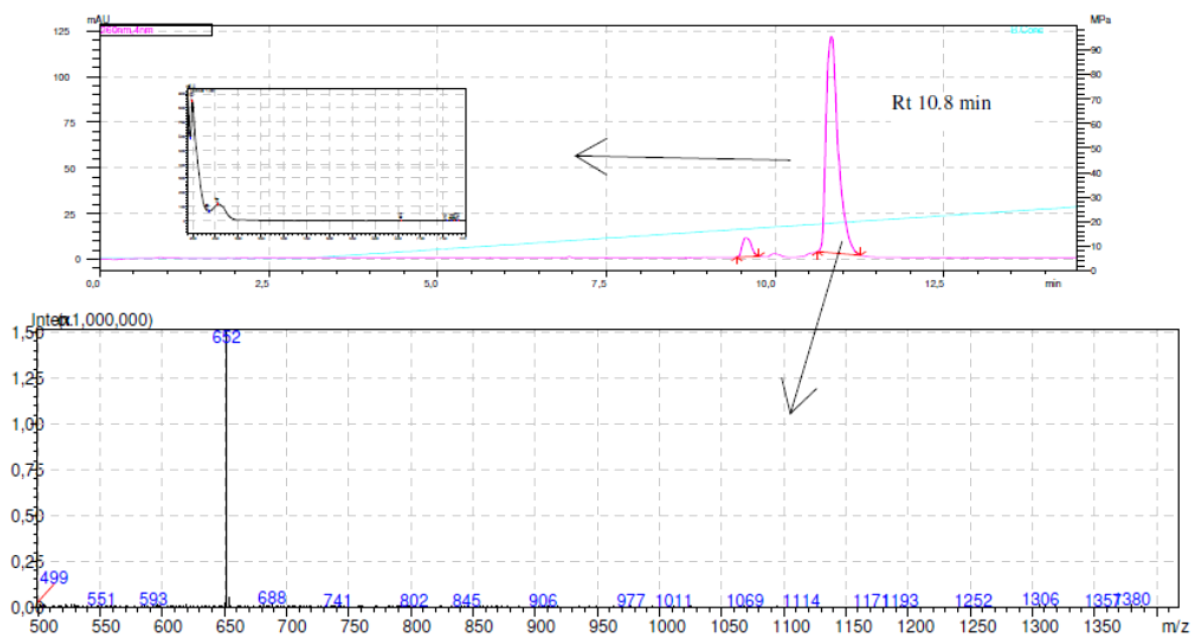
<p>ARC-3323</p>	
<p>ARC-3324</p>	
<p>ARC-3327</p>	
<p>ARC-3328</p>	

ARC-3342	
ARC-3343	
ARC-3344	
ARC-3345	
ARC-3346	

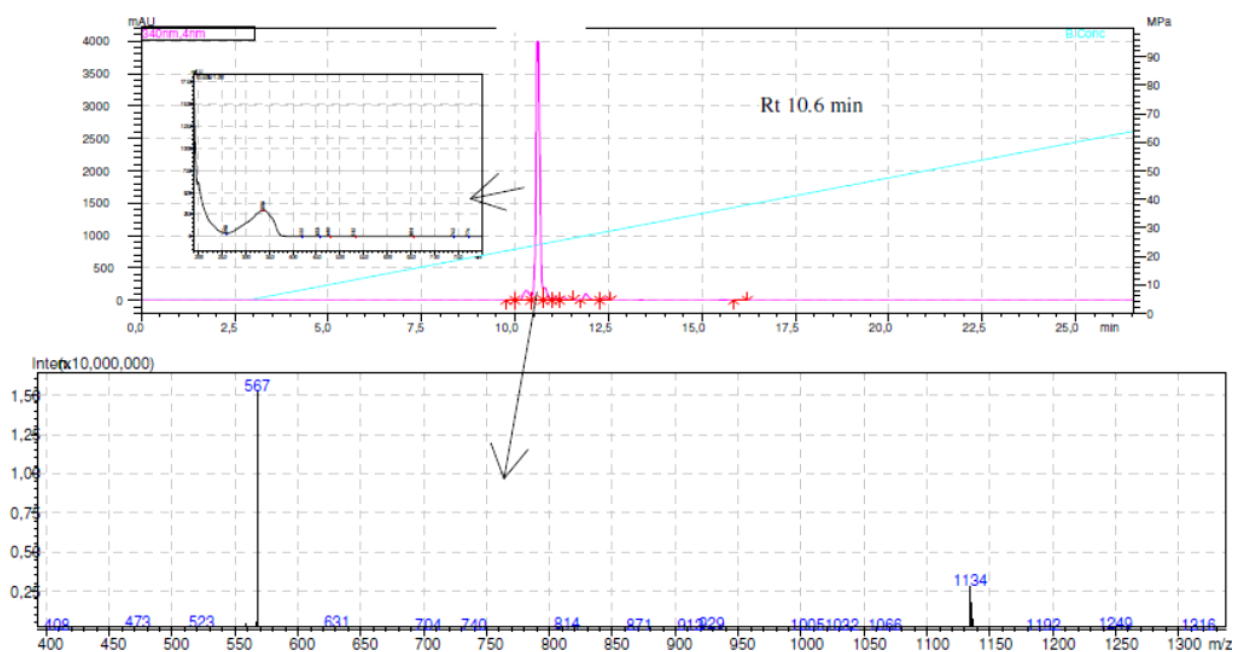
ARC-3347	
ARC-3348	
ARC-3349	
ARC-3350	
ARC-3351	

Appendix 10. Chromatograms of Purification, Photodiode Array Spectrum, and ESI-MS Data of Novel ARCs

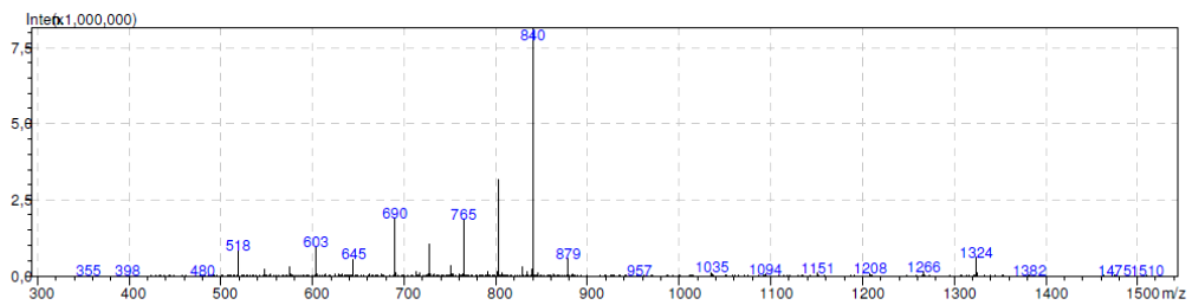
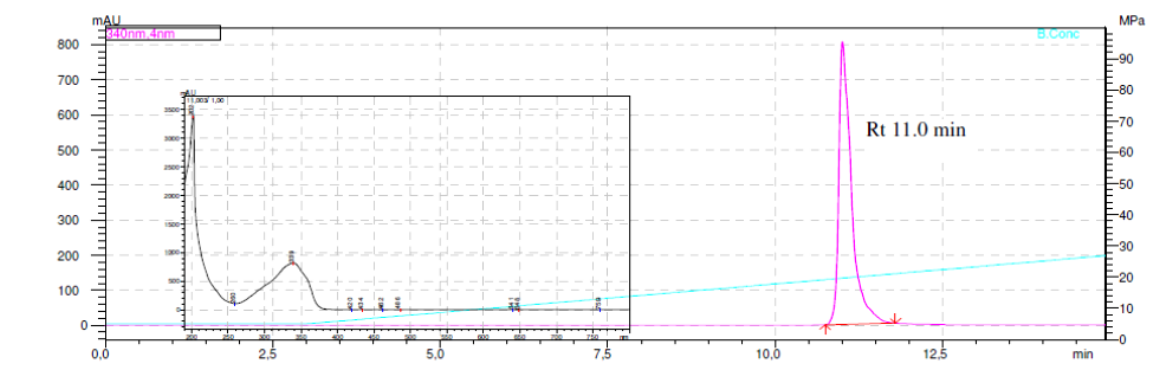
ARC-3323:



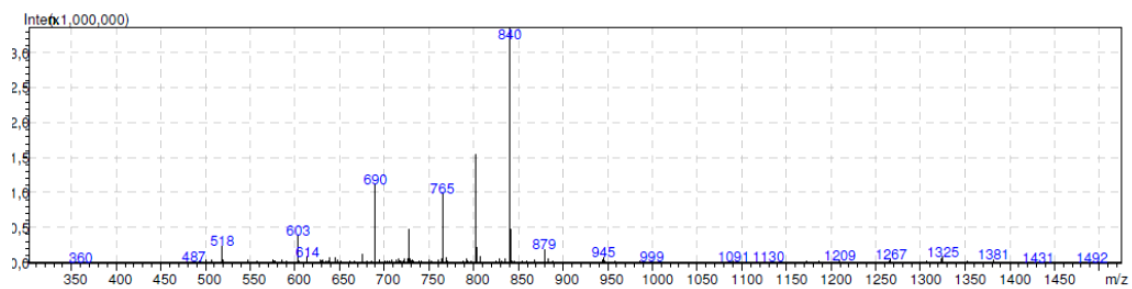
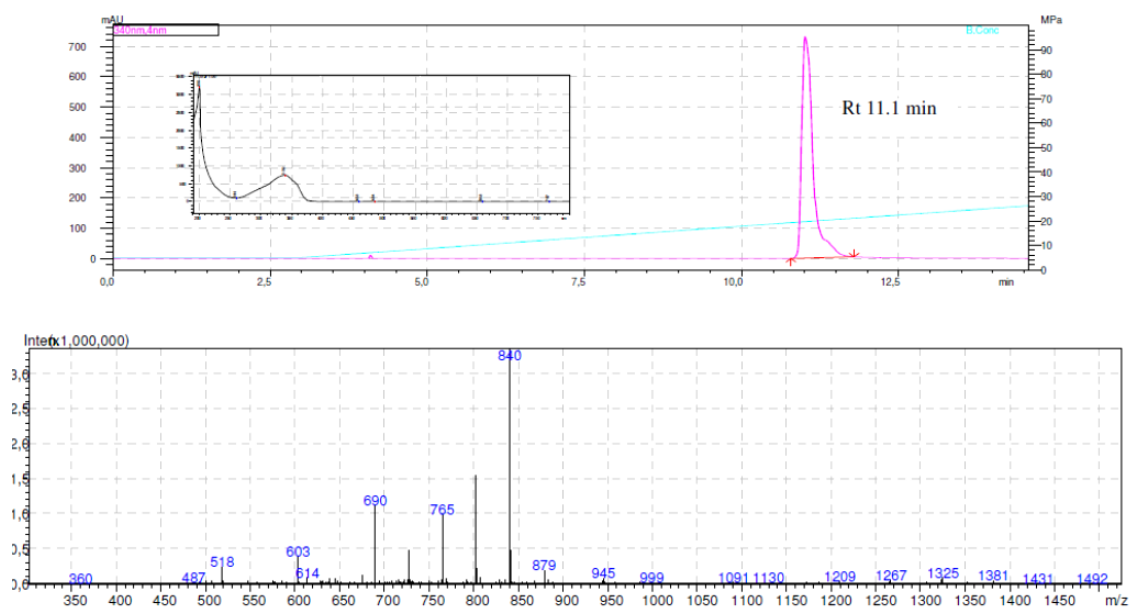
ARC-3324:



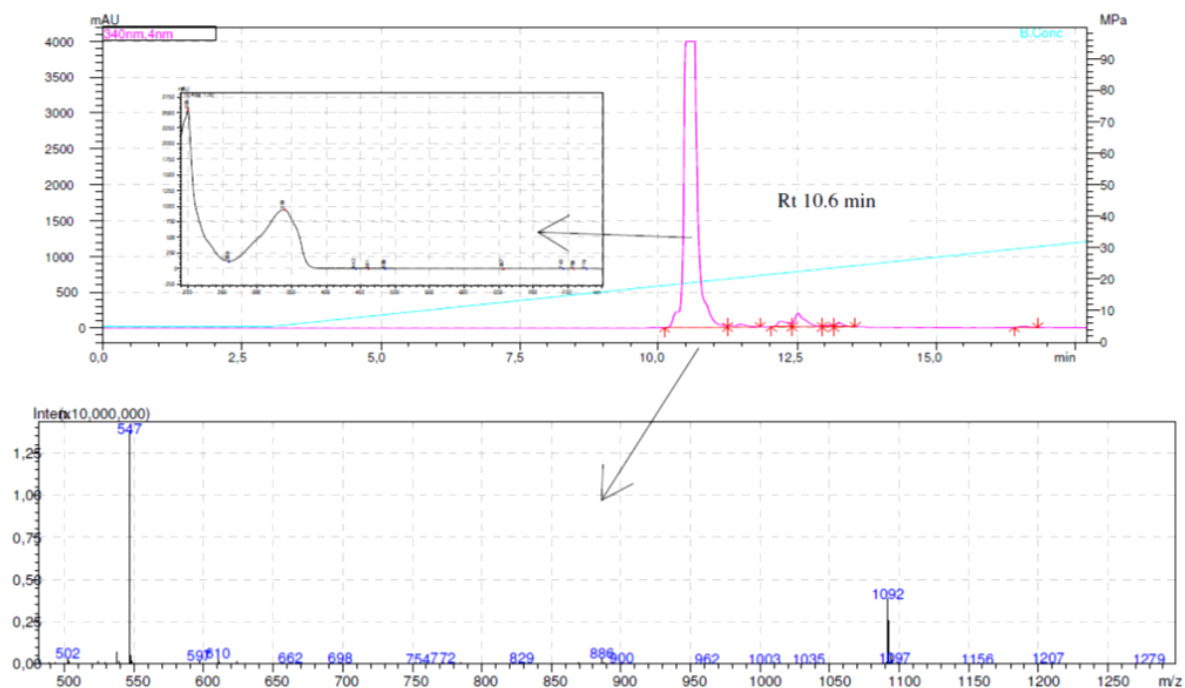
ARC-3327:



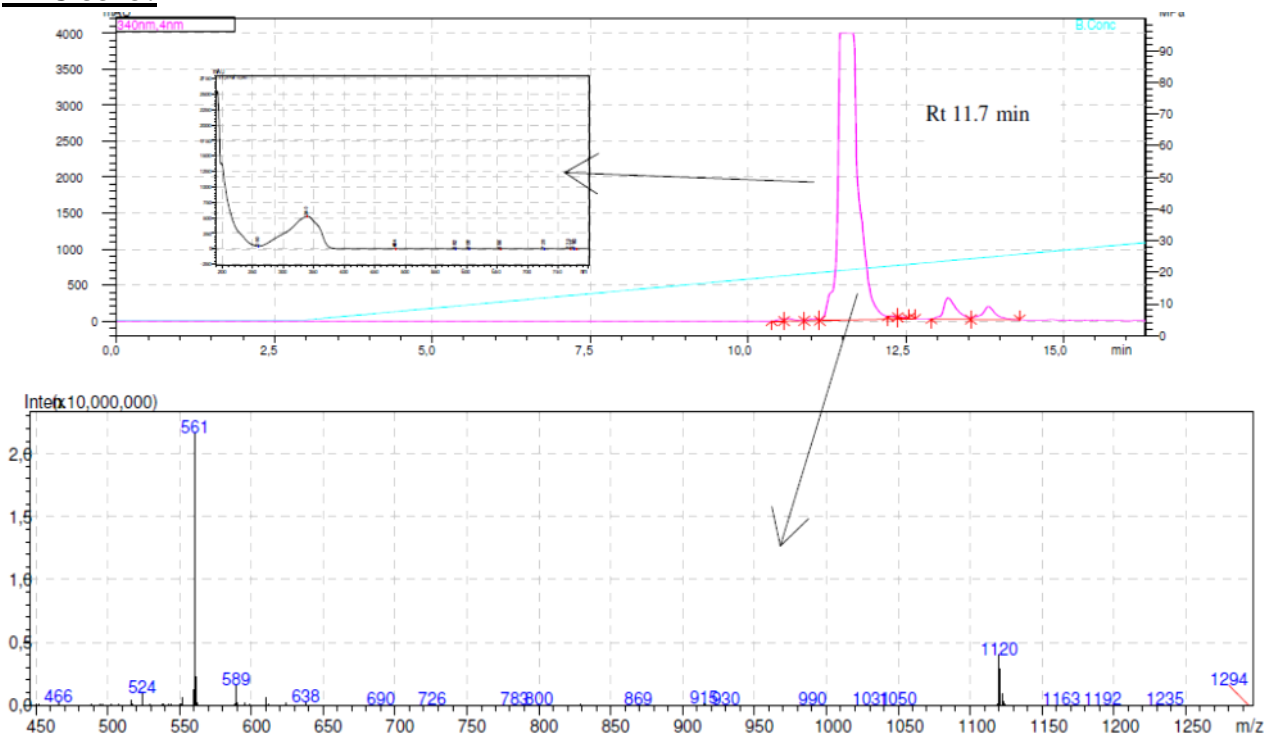
ARC-3328:



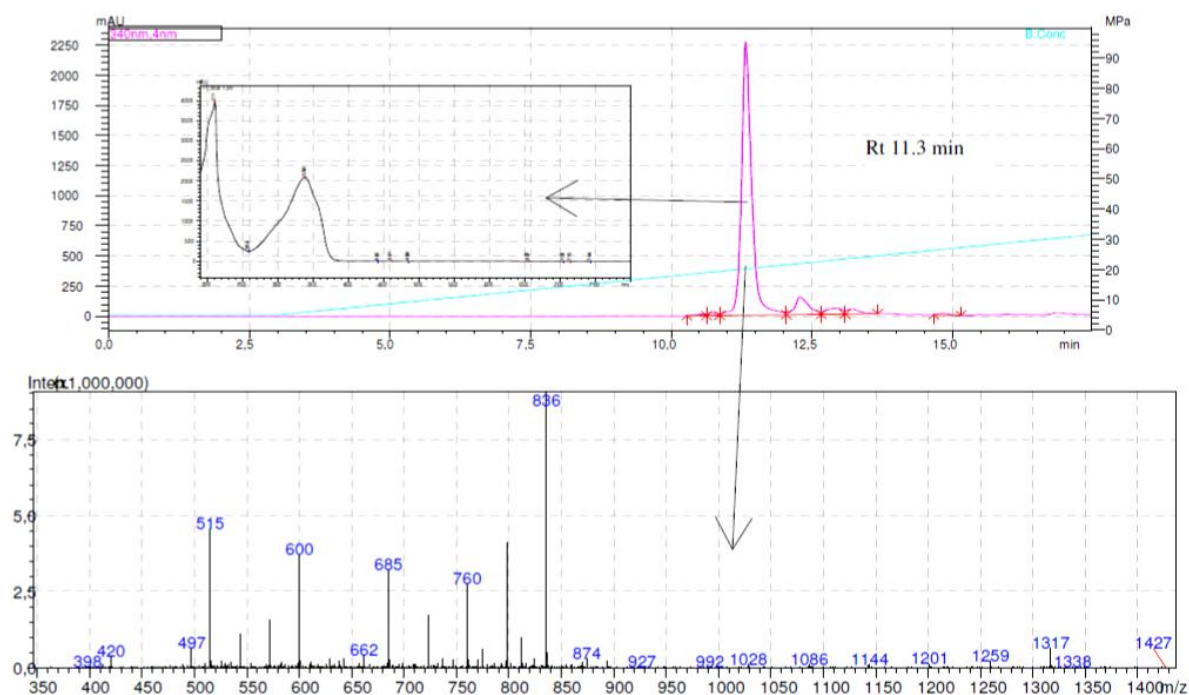
ARC-3342:



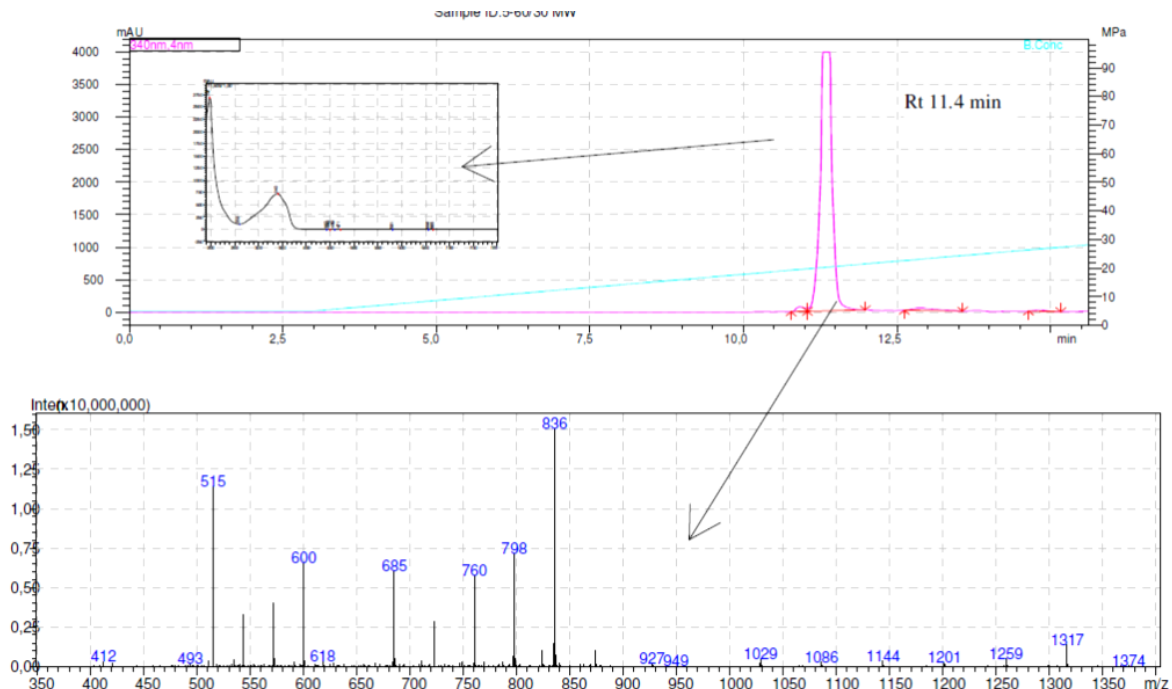
ARC-3343:



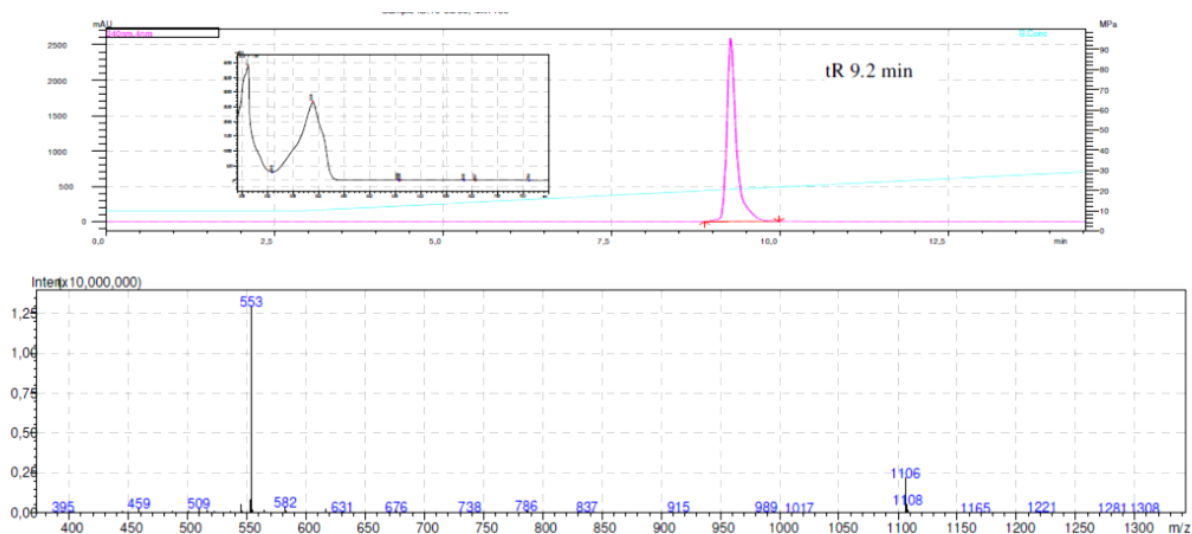
ARC-3344:



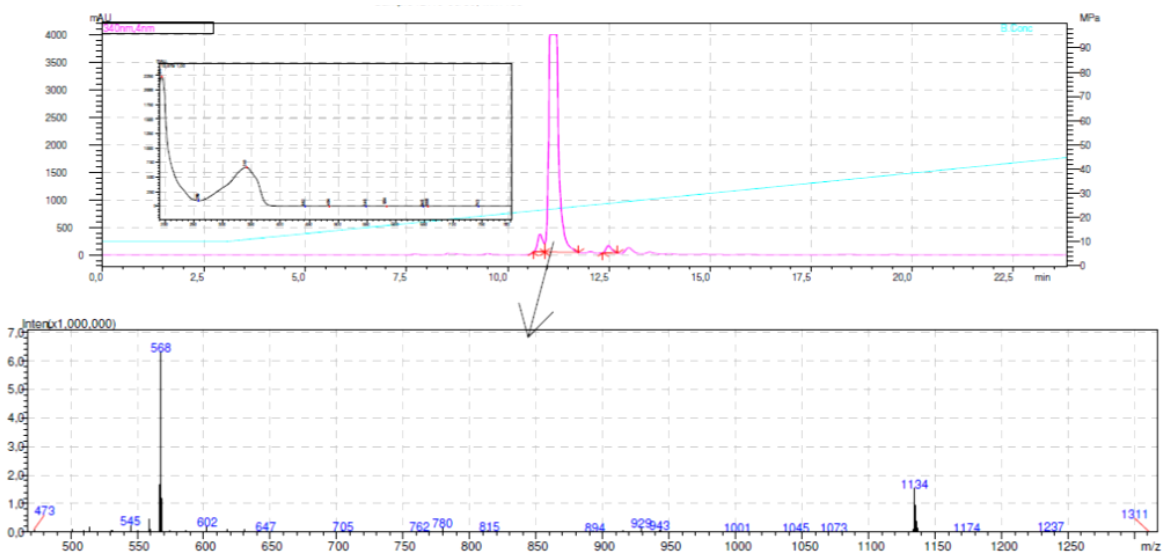
ARC-3345:



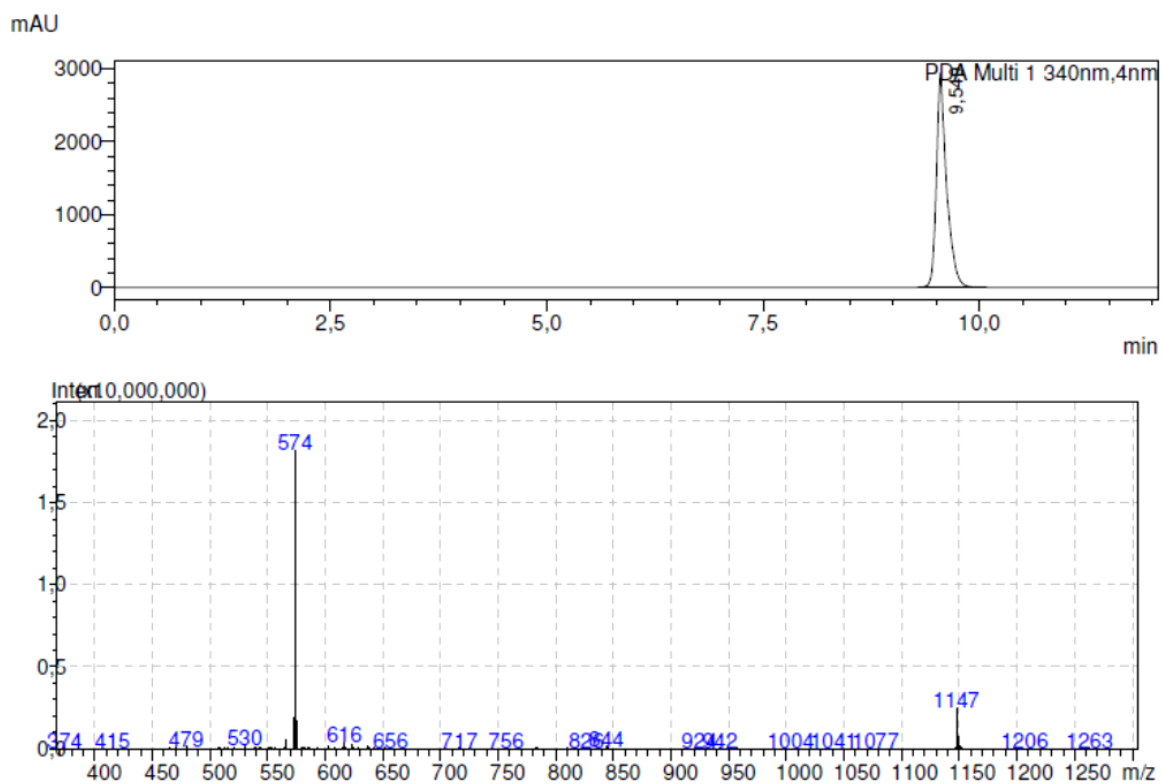
ARC-3346:



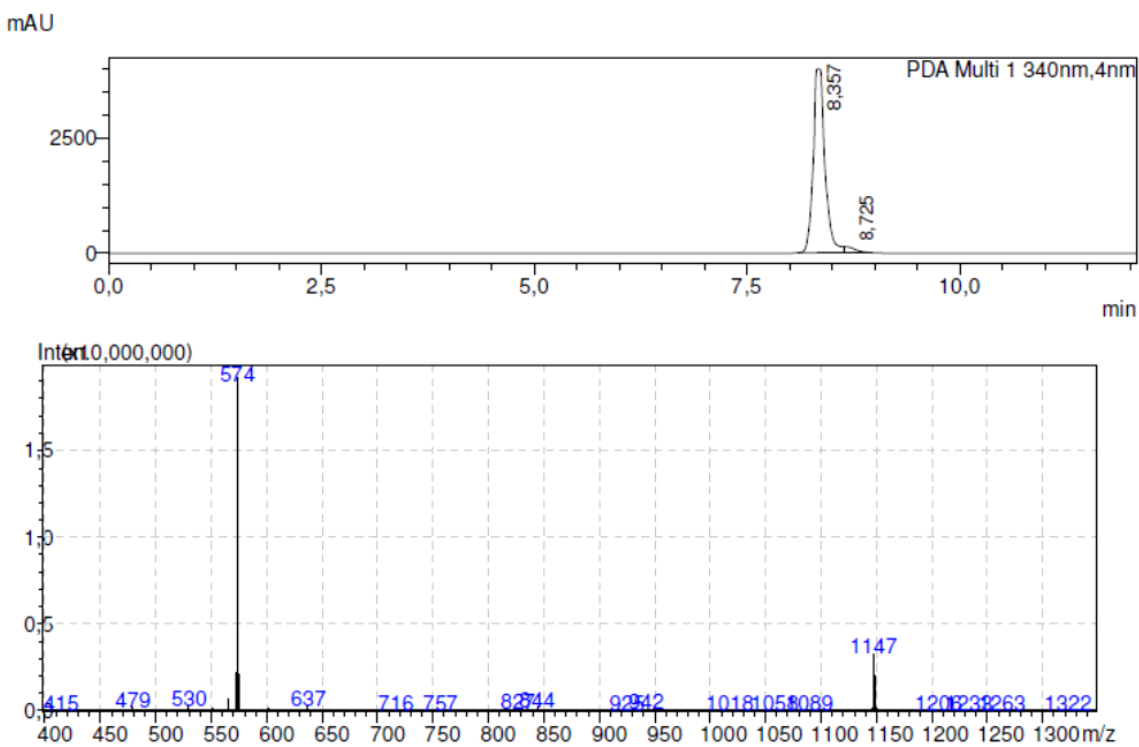
ARC-3347:



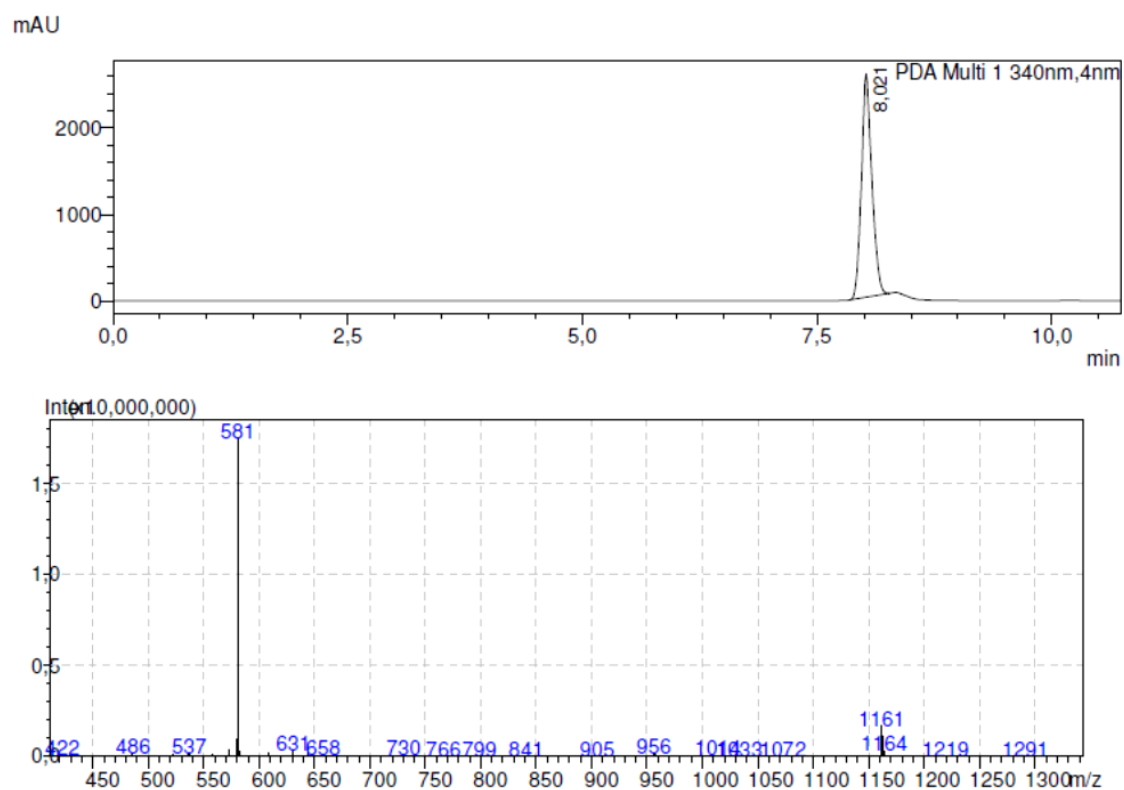
ARC-3348:



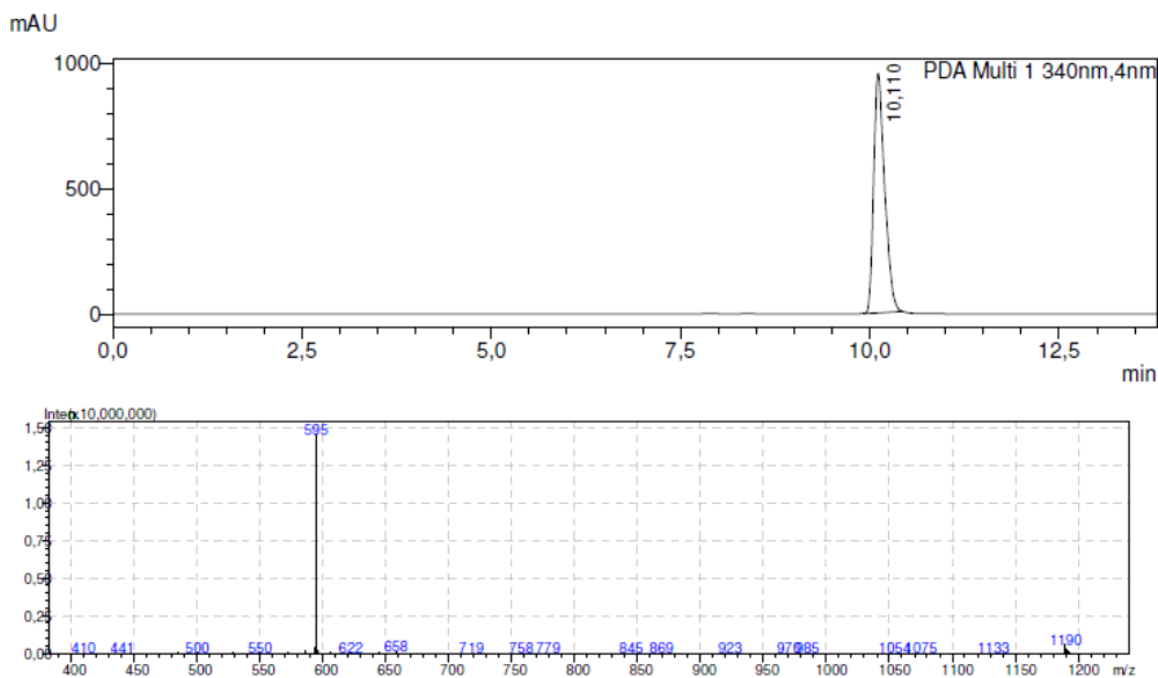
ARC-3349:



ARC-3350:



ARC-3351:



Appendix 11. Experimental Molecular Weights, Absorption Maxima, Extinction Coefficients (ϵ) of Novel ARCs

<i>Code</i>	<i>Molecular Formula</i>	<i>Calculated Monoisotopic Mass (Da)</i>	<i>Experimental Deconvoluted HRMS (Da)</i>	<i>ESI-MS</i>	λ_{max} (nm)	ϵ ($M^{-1}cm^{-1}$)
ARC-3323	C ₅₄ H ₉₀ N ₂₂ O ₁₆	1302.6905	1302.6928	652 (z=2)	259	15000
ARC-3324	C ₄₇ H ₇₆ N ₁₈ O ₁₃ S	1132.5560	1132.5571	1134 (z=1) 567 (z=2)	340	15000
ARC-3327	C ₈₃ H ₁₄₈ N ₄₂ O ₁₉ S	2069.1627	2069.1671	690 (z=3) 518 (z=4)	340	15000
ARC-3328	C ₈₃ H ₁₄₈ N ₄₂ O ₁₉ S	2069.1627	2069.1689	690 (z=3) 518 (z=4)	340	15000
ARC-3342	C ₄₅ H ₇₄ N ₁₈ O ₁₂ S	1090.5454	n.d.	1092 (z=1) 547 (z=2)	340	15000
ARC-3343	C ₄₇ H ₇₈ N ₁₈ O ₁₂ S	1118.5767	n.d.	1120 (z=1) 561 (z=2)	340	15000
ARC-3344	C ₈₃ H ₁₅₀ N ₄₂ O ₁₈ S	2055.1834	n.d.	685 (z=3) 515 (z=4)	340	15000
ARC-3345	C ₈₃ H ₁₅₀ N ₄₂ O ₁₈ S	2055.1834	n.d.	685 (z=3) 515 (z=4)	340	15000
ARC-3346	C ₄₆ H ₇₆ N ₁₈ O ₁₂ S	1104.5611	n.d.	1106 (z=1) 553 (z=2)	340	15000
ARC-3347	C ₄₈ H ₈₀ N ₁₈ O ₁₂ S	1132.5924	n.d.	1134 (z=1) 568 (z=2)	340	15000
ARC-3348	C ₄₇ H ₇₅ N ₁₉ O ₁₃ S	1145.5512	n.d.	1147 (z=1) 574 (z=2)	340	15000
ARC-3349	C ₄₇ H ₇₅ N ₁₉ O ₁₃ S	1145.5512	n.d.	1147 (z=1) 574 (z=2)	340	15000
ARC-3350	C ₄₈ H ₇₇ N ₁₉ O ₁₃ S	1159.5669	n.d.	1161 (z=1) 581 (z=2)	340	15000
ARC-3351	C ₅₀ H ₈₁ N ₁₉ O ₁₃ S	1187.5982	n.d.	1190 (z=1) 595 (z=2)	340	15000

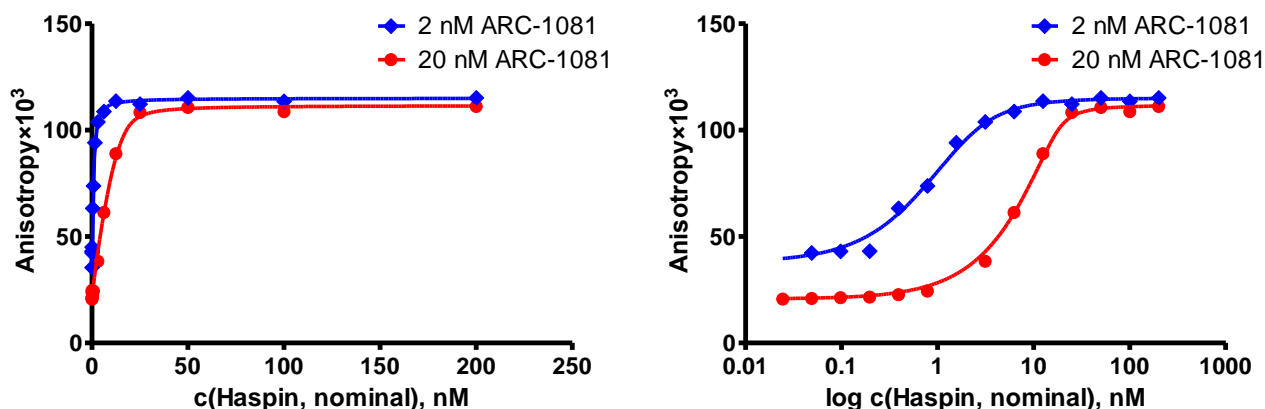
n.d. – not determined

Appendix 12. Parameters Characterizing the Affinity of the Inhibitor

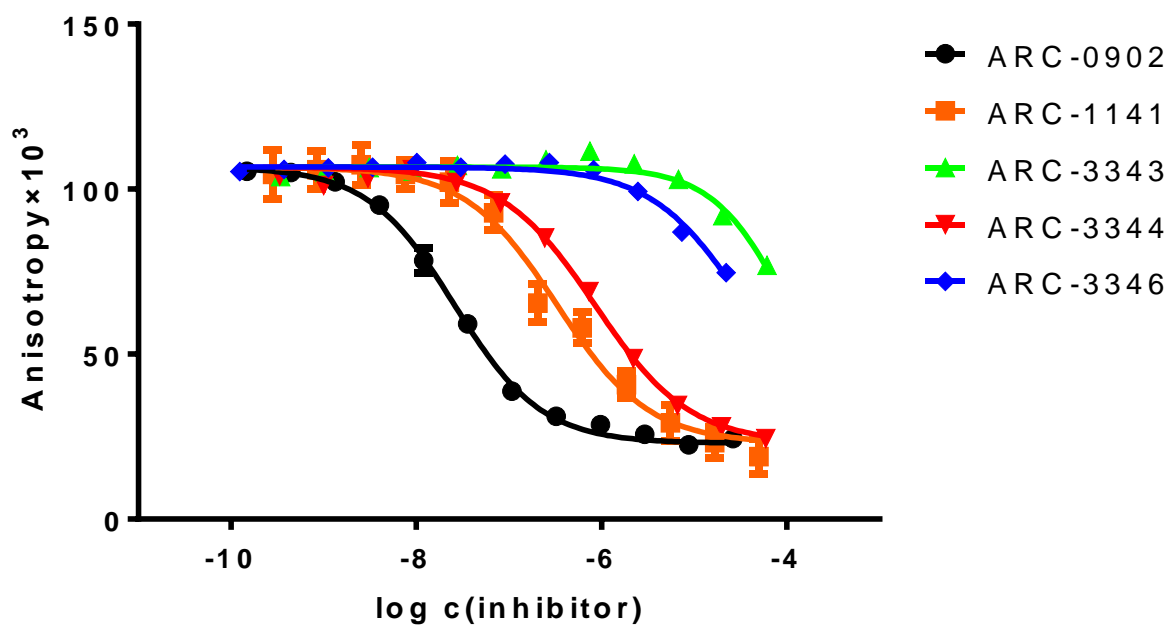
- The absolute inhibition constant (K_i) for competitive inhibition reflects the dissociation of the enzyme-inhibitor complex: $K_i = \frac{[E][I]}{[EI]}$ ($[E]$, $[I]$, and $[EI]$ correspond to equilibrium concentrations of the enzyme, inhibitor and enzyme-inhibitor complex). K_i value can only be established for inhibition assays (i.e., not binding assays) and is frequently calculated by converting the IC_{50} value to K_i using the Cheng-Prusoff equation (see below).
- Dissociation constant (K_D) is the equilibrium constant that characterizes dissociation process of a ligand from its complex with an enzyme: $K_D = \frac{[E][L]}{[EL]}$ ($[E]$, $[L]$, and $[EL]$ correspond to equilibrium concentrations of the enzyme, ligand and enzyme-ligand complex). In case of competitive inhibitors, K_D thus represents an analogue of K_i with the difference that K_D can be established from direct binding assays. In the current work, K_D is used for characterization of affinities of fluorescent probes towards the target PKs determined in binding assays.
- IC_{50} value corresponds to (A) the concentration of inhibitor at which the amount of the product phosphorylation reaction is half maximum (in inhibition assays), or (B) the concentration of competitor at which the amount of complex of fluorescence probe and kinase is half maximum (in binding assays). In inhibition assay, the IC_{50} value is dependent on the concentration and the K_M value of the competing substrate (*via* Cheng-Prusoff equation: $K_i = \frac{IC_{50}}{1 + \frac{[S]}{K_M}}$ (K_i is the inhibitory constant of the inhibitor, $[S]$ is the concentration of the competing substrate, and K_M is the concentration of substrate at which the activity of the enzyme is half the maximum). In binding assay, the IC_{50} value is dependent on the concentration and the K_D value of the competing ligand (in case when the concentrations of enzyme and ligand are in the same range, the IC_{50} of competitor is also dependent on the concentration of enzyme [71]).
- Displacement constant K_d is the equilibrium constant that is analogical to K_D ; in the current work, K_d is used for characterization of affinities of fluorescently non-labelled compounds towards the target PKs determined in the displacement assays. Note that K_d values were not measured directly, but calculated using the IC_{50} values of displacing compounds, the K_D values of fluorescent probes, and the total concentration of protein kinase and fluorescent probe in the assay.

Although the thermodynamic constants as K_i , K_D , and K_d should be dimensionless, the molar dimension is nearly always used in biochemical literature to emphasize the fact that these values are characterizing the dissociation processes.

Appendix 13. The Examples of Binding and Displacement Curves Obtained by FA Method



Determination of the active concentration of Haspin and K_D value of ARC-1081 by FA binding assay. The titration was performed with two different concentrations of fluorescence probe ARC-1081 (final total concentration 2 nM or 20 nM). The X-axis of the left graph is in linear scale and of the right graph in logarithmic scale.



Representative displacement curves obtained in FA displacement assay. Fluorescence probe ARC-1081 (2 nM) was displaced from its complex with Haspin (6 nM) by concentration series of non-fluorescent inhibitors (N=2).

Lihtlitsents lõputöö reprodutseerimiseks ja lõputöö üldsusele kättesaadavaks tegemiseks

Mina, _____ KATRIN KALIND _____,
(*autori nimi*)

1. annan Tartu Ülikoolile tasuta loa (lihtlitsentsi) enda loodud teose
_____ TARGETING BOTH SUBSTRATE-BINDING SITES OF MITOTIC HASPIN _____
_____ KINASE WITH A SINGLE INHIBITOR _____,
(*lõputöö pealkiri*)

mille juhendaja on _____ Dr Darja Lavõgina, Dr Asko Uri _____,
(*juhendaja nimi*)

1.1.reprodutseerimiseks säilitamise ja üldsusele kättesaadavaks tegemise eesmärgil, sealhulgas digitaalarhiivi DSpace-is lisamise eesmärgil kuni autoriõiguse kehtivuse tähtaja lõppemiseni;

1.2.üldsusele kättesaadavaks tegemiseks Tartu Ülikooli veebikeskkonna kaudu, sealhulgas digitaalarhiivi DSpace'i kaudu alates **01.06.2017** kuni autoriõiguse kehtivuse tähtaja lõppemiseni.

2. olen teadlik, et nimetatud õigused jäävad alles ka autorile.

3. kinnitan, et lihtlitsentsi andmisega ei rikuta teiste isikute intellektuaalomandi ega isikuandmete kaitse seadusest tulenevaid õigusi.

Tartus, **22.05.2014**