WARSAW UNIVERSITY OF TECHNOLOGY

FIELD OF SCIENCE NATURAL SCIENCES

DISCIPLINE: CHEMICAL SCIENCES

Ph.D. Thesis

Piotr Krzysztof Grześkowiak, M.Sc.

Synthesis and evaluation *in vitro* of novel arginine analogues as potential new components of metabolic anticancer therapy

Supervisors

Dr. Mariola Koszytkowska-Stawińska, PhD, DSc Eng, Assoc. Prof.

Prof. Maria Jolanta Rędowicz, PhD, DSc,

WARSAW 2024

Praca doktorska wykonana w ramach Interdyscyplinarnych studiów doktoranckich TRI-BIO-CHEM

Niniejsza rozprawa doktorska powstała dzięki finansowaniu interdyscyplinarnych studio doktoranckich TRI-BIO-CHEM. Projekt "Od chemii do bioinnowacji dla lepszego życia - interdyscyplinarne studia doktoranckie TRI-BIO-CHEM" realizowany przez trzy jednostki partnerskie (Wydział Chemii Uniwersytetu Warszawskiego, Wydział Chemiczny Politechniki Warszawskiej, Instytut Biologii Doświadczalnej im. M. Nenckiego) w ramach Projektu Operacyjnego Wiedza Edukacja Rozwój 2014-2020, współfinansowany ze środków Europejskiego Funduszu Społecznego, POWR.03.02.00-00-I007/16-00 (POWER 2014-2020).

Projekt realizowany w ramach Programu Operacyjnego Wiedza Edukacja Rozwój 2014-2020 współfinansowanego ze środków Europejskiego Funduszu Społecznego

Instytucja wiodąca: **Wydział Chemii, Politechnika Warszawska**
Instytucja partnerska: **Instytut Biologii Doświadczalnej im. M. Nenckiego PAN**

Acknowledgements

Neither part of this work would happen without unwavering support from my supervisors Dr Mariola Koszytkowska-Stawińska and Prof. Maria Jolanta Rędowicz. I am beyond grateful for their guidance and mentorship over the last six years.

I direct my special thanks to Olena Karatsai for sharing with me not only her insight into molecular biology but also the practical laboratory knowledge. As well as my wholehearted gratitude for entire team of Laboratory of Molecular Basis of Cell Motility PAS for sharing not only their limited lab space with me but also warm words and support during my time there.

I am grateful to my Dad, for sparking in me the curiosity about the world that made me pursue research career. His thirst for knowledge and analytical way of thinking have always been a role model for me.

I am grateful to my Mom, for instilling in me those few bits of tenacity that stuck with me over the years. Without those I would not be able to accomplish anything.

Lastly, to Alicja, for sticking with me through thick and thin.

This six year journey made me grow not only as a researcher but also as a person. It composed of myriad of interactions with dozens of people. Support of each and every person will stick with me through the rest of my life.

Thank you!

Contents

Abstract

Cancer, being a personal tragedy for millions of people every year, is the extremely important problem to tackle itself. With the added impact on global economy, it grows to be the priority target for modern biomedicine. The demand for new, more potent and less toxic chemotherapeutics is ever growing.

Over the last decade, a promising metabolic target has emerged, it is the arginine metabolism. Arginine is one of 20 biogenic amino acids. Apart from its crucial role in protein synthesis and modification, due to its high nitrogen content, it plays a major role in nitrogen metabolism I the human organism. That is why a disruption of arginine levels can be fatal to any cell lacking endogenic arginine. A number of cancer types, including glioblastoma and melanoma, have been shown to lack argininosuccinate synthase 1 (ASS) expression, enzyme crucial for the endogenous arginine synthesis. Such cells being functional arginine auxotrophs, rely solely on the arginine uptake for the proper protein synthesis. Arginine deprivation-based therapeutic strategy has been tested in multiple studies, for example the effect of PEGylated arginine deiminase (a microbial enzyme degrading arginine) on melanoma and hepatocellular carcinoma with a notable cytostatic effect. These findings open a window for a selective therapy via the arginine starvation and/or replacement.

This effect, however significant, is not prominent enough for the successful treatment. Thus, the idea has emerged, of a therapy combining arginine starvation and natural arginine analogue, canavanine. Canavanine has been tested previously as an anticancer agent on its own. Its efficacy, however, was not high enough compared to its toxicity, leaving a therapeutic window to narrow for medical standards. Canavanine in combination therapy with arginine starvation exhibited strong synergistic effect in low micromolar doses. Though potent, canavanine, due to its short half-life in the cell matrix, might not be a best candidate for therapeutic purpose. This leaves open window for novel arginine analogues. Moreover, its mode of action is not yet fully understood. I hope to that my research will shed more light into this area with our newly designed compounds. Thus, research proposed in this project is desired and justified. The herein proposed methodology provides novel arginine analogues with a potentially high bioactivity.

The main goal of this project was the design, synthesis and testing of potential arginine analogues that, in combination with arginine deprivation conditions, could provide the strong cytostatic or even cytotoxic effect selectively towards the examined glioblastoma cell lines. A secondary aim was to gather a knowledge on the structure–activity relationship for arginine binding pocket in arginyl–tRNA synthetase for the future therapeutics development.

Keywords

Amidine, arginine, cancer, click chemistry, glioblastoma, metabolic therapy, Pinner synthesis

Streszczenie

Nowotwory, stanowiąc osobistą tragedię dla milionów ludzi każdego roku są ważnym problemem oczekującym rozwiązania. Zachorowania na choroby nowotoworowe mają też wpływ na globalną ekonomię. Z tego względu choroby nowotworowe stają się jednym z najważniejszych problemów nowoczesnej medycyny. Z tego względu istnieje coraz większa potrzeba nowych chemioterapeutyków o mniejszych skutkach ubocznych.

W ostatniej dekadzie metabolizm argininy okazał się interesującym celem terapeutycznym. Arginina jest jednym z dwudziestu aminokwasów biogennych. Oprócz istotnego udziału w syntezie i modyfikacji potranslacyjnej białek, aminokwas ten odgrywa kluczową rolę w obiegu azotu w ludzkim organizmie. Z tego względu zaburzenie poziomu argininy może być śmiertelne dla komórek pozbawionych endogennej argininy. Wybrane typy nowotworów (m.in. glejaki lub czerniaki) wykazują zaburzenia w produkcji syntazy argininobursztynianowej 1 (ASS1). Enzym ten jest niezbędny do endogennej syntezy argininy. Takie komórki są auksotrofami argininy i polegają jedynie na pozakomórkowych źródłach argininy. Strategia terapeutyczna polegająca na rozkładzie argininy w miejscu objętym zmianą nowotworową była przedmiotem kilku badań. Wykazano, że wprowadzenie PEGylowanej deiminazy argininowej (bakteryjnego enzymu rozkładającego argininę) w otoczenie czerniaka lub raka wątrobowokomórkowego, wykazuje istotny efekt cytostatyczny. Odkrycia te otwierają nowe możliwości w selektywnej terapii nowotworów poprzez pozbawianie argininy lub jej zastępowanie.

Obserwowany efekt enzymów rozkładających argininę nie jest jednak wystarczający aby terapia ta była skuteczna. Z tego względu narodził się nowy koncept terapii kombinowanej tj. łączącej ograniczenie podaży endogennej argininy (tzw. głodzenie argininowe ang. Argininę starvation) z podawaniem naturalnego analogu argininy – kanawaniny. Kanawanina była wcześniej testowana w roli chemioterapeutyku, jednak jej skuteczność była zbyt niska w porównaniu z toksycznością wobec zdrowych komórek. Zastosowanie kanawaniny w terapii skojarzonej wykorzystującej głodzenie argininowe wykazało interesujący efekt synergistyczny. Potencjał terapeutyczny kanawaniny jest jednak ograniczony, ponieważ jest ona relatywnie szybko rozkładana przez enzymy trawiące argininę. Z tego względu, istnieje potrzeba opracowania nowych analogów argininy charakteryzujących się potencjalną aktywnością kanawaniny ale zwiększoną stabilnością metaboliczną.

Głównym celem mojej pracy było zaprojektowanie, synteza oraz ewaluacja *in vitro* nowych analogów argininy charakteryzujących się wysokim potencjałem terapeutycznym w terapii wykorzystującej głodzenie argininowe. Kolejnym celem było zdobycie informacji na temat relacji struktura–aktywność miejsca aktywnego wiążącego argininę w syntazie arginylowej, które mogą być przydatne w opracowaniu potencjalnych chemioterapii

Słowa kluczowe

Amidyna, arginina, nowotwór, click chemistry, glejak, terapia metaboliczna, synteza Pinnera

List of abbreviations

AcOEt	–	ethyl acetate
AFM	–	arginine free medium
APT	–	attached proton test
Arg	–	L-Arginine
ArgRS	–	arginyl-tRNA synthetase
ARG1	–	arginase 1
ARG2	–	arginase 2
ASL	–	argininosuccinate lyase
Asn	–	asparagine
Asp	–	aspartate
ASS	–	argininosuccinate synthetase 1
ATCC	–	American Type Culture Collection
ATE1	–	arginyl transferase 1
ATP	–	adenosine triphosphate
Bn	–	benzyl
Boc	–	*tert*-butyloxycarbonyl
Cav	–	canavanine
Cbz	–	benzyloxycarbonyl
CM	–	complete medium
CNS	–	central nervous system
cPARP1	–	cleaved (inactive) poly ADP-ribose polymerase 1
CuAAC	–	copper assisted azide-alkyne cycloaddition
C3	–	caspase 3

C9	–	caspase 9
DBU	–	1,8-diazabicyclo[5.4.0]undec-7-ene
DCM	–	dichloromethane
DMEM	–	Dulbecco's Modified Eagle medium
DMF	–	dimethylformamide
DMPM	–	3,4-dimethoxybenzyl
DNA	–	deoxyribonucleic acid
FAK	–	focal adhesion kinase
FBS	–	fetal bovine serum
Fmoc	–	fluorenylmethyloxycarbonyl
GPU	–	graphic processing unit
His	–	histidine
HTS	–	high throughput screening
Isp	–	indospicine
LM	–	light microscopy
MD	–	molecular docking
MeOH	–	methanol
Moz	–	*p*-methoxybenzyloxycarbonyl
MsCl	–	mesyl chloride
MTS	–	(3-(4,5-dimethylthiazol-2-yl)-5-(3-carboxymethoxyphenyl)-2-(4-sulfophenyl)-2*H*-tetrazolium)
MW	–	molecular weight
NADPH	–	nicotinamide adenine dinucleotide phosphate
NMR	–	nuclear magnetic resonance
NO	–	nitric oxide

NOS	–	nitric oxide synthetase
OBn	–	oxybenzyl
PARP1	–	poly ADP-ribose polymerase 1
PDB	–	protein data bank
PEG	–	polyethylene glycol
PG	–	protecting group
PMB	–	*p*-methoxybenzyl
P3	–	procaspase 3
P9	–	procaspase 9
RDD	–	rational drug design
RM	–	reactive moiety
RT	–	room temperature
TEA	–	triethylamine
TFA	–	trifluoroacetic acid
THF	–	tetrahydrofuran
TLC	–	thin layer chromatography
tRNA	–	transporter ribonucleic acid
$tRNA^{Arg}$	–	arginyl-tRNA
USD	–	United States dollars
U87MG	–	Uppsala 87 Malignant Glioma
U251MG	–	Uppsala 251 Malignant Glioma
WB	–	western blot
WHO	–	World Health Organization

1. Introduction

1.1 Arginine as a protein building block

L-Arginine is one of the 20 biogenic amino acids, meaning it is a necessary component for protein synthesis in every living cell (**Fig 1**, Arg)[1]. Arg is also classified as a conditionally essential amino acid since it cannot be synthesized endogenously under specific conditions (infancy, disease, etc.) Apart from being a protein building block, it is also involved in numerous metabolic pathways and serves as a source for this element due to the high nitrogen content.

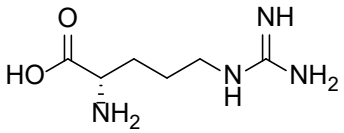


Fig 1 L-Arginine (Arg) structure

By definition, Arg as amino acid is composed of a "backbone" that includes amine and carboxylic acid group, and three carbon atom long sidechain ended with a guanidine group. In physiological conditions, the guanidine group is protonated, what results in Arg having a net positive charge *in vivo*[1].

The crucial role of Arg in protein synthesis is emphasized by the fact this is one of the three naturally occurring amino acids introducing a positive charge into a protein molecule. Charge distribution in a peptide chain is vital for its proper folding into a tertiary structure[2]. Moreover, protein-protein and protein-environment interactions also depend on the correct charge location on the protein surface. The replacement of a positively charged amine group in the sidechain of the lysine to slightly more charged guanidine group of Arg can impact protein functionality[3].

Incorporation of Arg into the polypeptide backbone is not the only way that alters the protein structure. Arginylation is a common post-translational modification in which side chains or *N*--terminal amino group of a protein are altered by bonding with carboxylic group of Arg. This subtle addition to the structure has a significant impact on the protein conformation, activity, function and its lifecycle[4].

First, every living cell depends on protein recycling for its wellbeing. The primary system responsible for protein degradation in human cells is proteasome (**Fig 2**)[5]. Proteasomes are complexes of multiple enzymes breaking down peptide bonds. Access to the proteasome

is controlled by a lid subunit (**Fig 2**) that recognizes proteins tagged with ubiquitin. Thus, any protein that receives the ubiquitin tag is deemed unnecessary by mechanisms governing the cell metabolism and is sent to the proteasome to recycle its building blocks[6]. According to the *N*-end rule, the likelihood of a protein being ubiquitinylated is directly affected by the type of an amino acid located at its *N*-terminus[7]. For example, post-translational arginylation of ornithine decarboxylase, more than halves its half-life in the cultured rat hepatocytes[8].

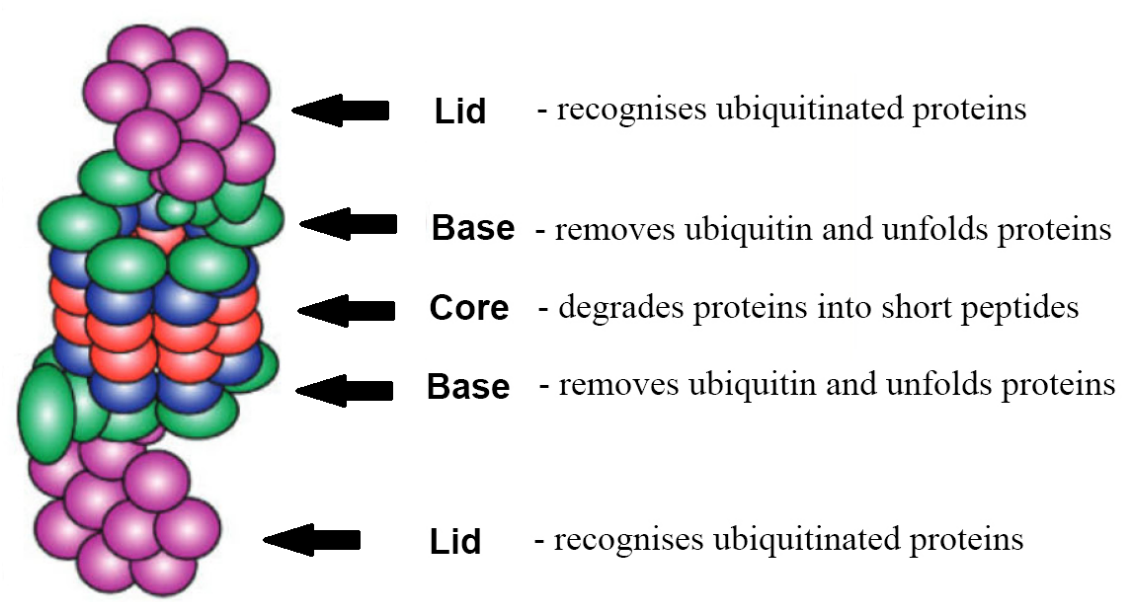


Fig 2 Schematic structure of proteasome (adapted from[5])

Secondly, arginylation also modulates a peptide affinity towards receptors. It has been found that addition of Arg to the sidechain of glutamate in neurotensin improves its affinity towards known neurotensin receptors.[9] Neurotensin (**Fig 3**) is a peptide composed of 13 amino acids. Its primary function is neuromodulation, *i.e.* regulation of the central nervous system activity by governing the release of lutropin and prolactin, important reproductive system hormone.[10]

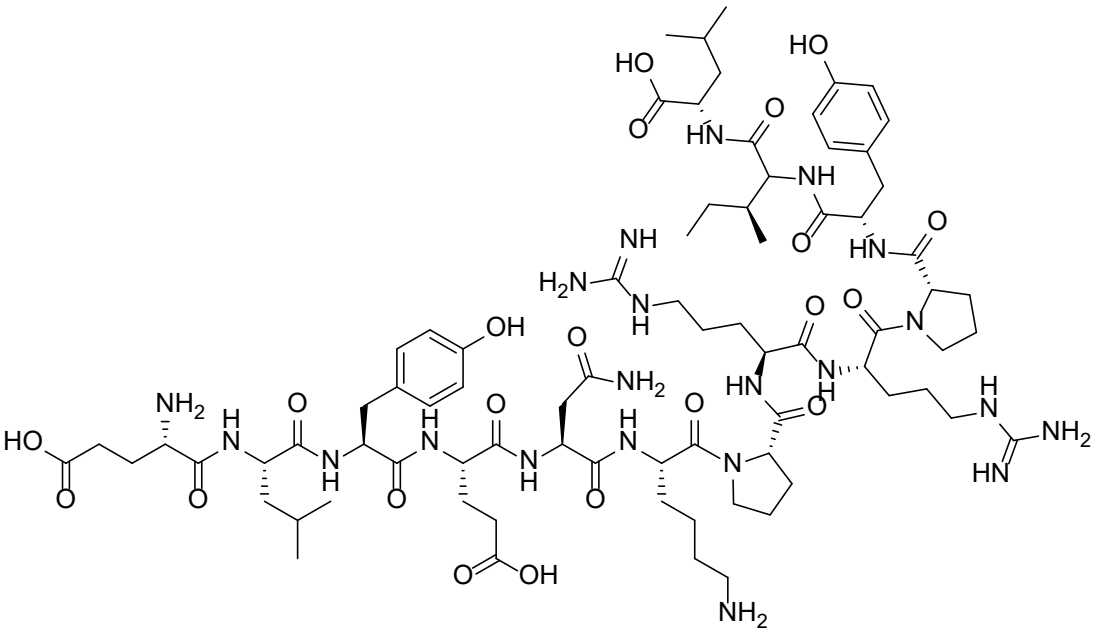


Fig 3 Neurotensin structure

Thirdly, post-translational introduction of Arg into some proteins is crucial for their proper function. One of them is *β*-actin (**Fig 4**), which is a cytoplasmic actin isoform serving as a major component of the cytoskeleton.[11] Assembly of actin ~42 kDa monomers results in formation of filaments with a generally negatively charged surface[11] However, around 20% of actin monomers are being arginylated, this results in actin filaments being "coated" with positively charged Arg sidechains (**Fig 4 B**). This charge distribution prevents formation of actin filaments bundles, important structures for cell motile properties.[11]

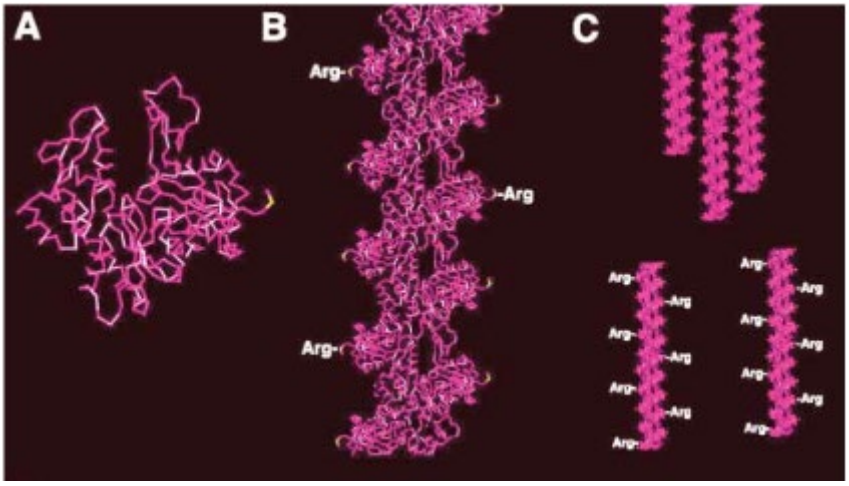


Fig 4 Schematic representation of effect of arginylation on actin organization. A – monomeric *β*-actin subunit. B – actin filament assembled with arginylation sites marked. C – comparison of a behavior of arginylated and non-arginylated actin filaments (Adapted from [11])

1.2 Arginine metabolism

One of the pathways of endogenous Arg synthesis is the urea cycle[12]. The urea cycle is a system of biochemical reactions that mammalian cells employ to neutralize toxic ammonia. Its net product is urea. Urea is inert enough to be safely excreted out of organisms through the skin, lungs or renal system. Arg synthesis is a two-step reaction (**Fig 5** marked in green). First, argininosuccinate intermediate is produced from citrulline and aspartate. This process is assisted by the argininosuccinate synthetase 1 (ASS). In the next step, argininosuccinate lyase (ASL) breaks arginosuccinate into Arg and fumarate. Resulting Arg is used up or further incorporated into the urea cycle by being hydrolyzed to ornithine by arginase 1 (ARG1)[12,13] (**Fig 5,** marked in blue).

Freebase Arg is not recognized by enzymes responsible for protein synthesis and arginylation. For Arg to become their substrate, it needs to be attached to Arg specific tRNA tag. The process of attaching Arg to tRNA is chaperoned by arginyl-tRNA synthetase (ArgRS). The produced arginyl-tRNA ($tRNA^{Arg}$) can be used by rybosomes in protein synthesis, or for arginylation by enzymes such as arginyl transferase 1 (ATE1)[14].

Arg is also a precursor of important biomolecules that take part in energy production, regulation of DNA expression and tissue remodeling, including ornithine, polyamines, proline, creatine and nitric oxide[15–18]

Ornithine is a major Arg metabolite (**Fig 5**). Ornithine is a precursor of multiple important molecules such as proline and putrescine. Ornithine also takes a direct part in energetic metabolism as it is a major source for ammonia for transamination process. Transamination is an important part of citric acid cycle. Thus, ornithine plays a big part in energy production from amino acid catabolism.[15,17]

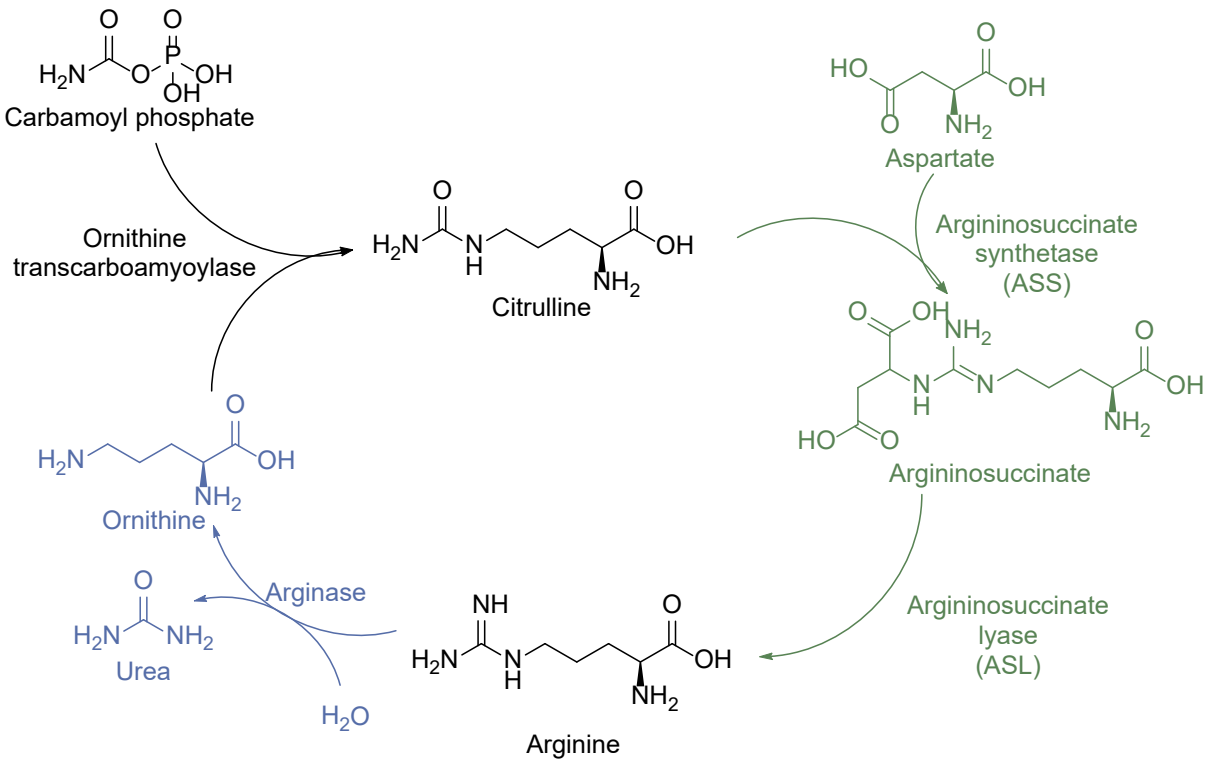


Fig 5 Urea cycle. Arg synthesis in green, Arg hydrolysis in blue

Polyamines are a class of organic molecules with multiple amine groups. *In vivo,* polyamines are directly synthesized from putrescine (**Fig 6**), an aliphatic diamine derivative of ornithine. Polyamine function and mode of action are not well understood, but they influence the activity of ion channels and regulate gene expression. This renders ornithine (and indirectly Arg) crucial in mammal fetal development. Knocking out genes for polyamine synthesizing enzymes arrest mice fetal development in early stages.[18]

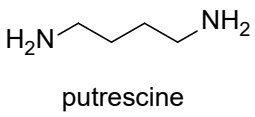
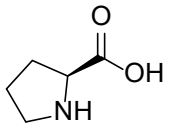


Fig 6 Structure of putrescine

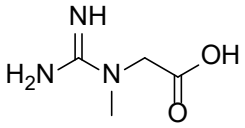
Arg/ornithine pathway is a major internal source of proline (**Fig 7**)[17]. Proline is the only proteinogenic amino acid with a cyclic structure. Proline composes as much as a quarter of the amino acid sequence in collagen[16]. Collagen, a protein which fills voids in extracellular matrix, is essential for maintaining tissue structure and wound healing.



proline

Fig 7 Structure of proline

Another molecule that can be traced to Arg is creatine (**Fig 8**)[19]. Creatine is a small organic molecule containing guanidine group that undergoes phosphorylation readily. Creatine phosphate serves as a phosphate source for ATP synthesis. ATP recirculation is crucial for proper function of any cell but particularly for muscle cells.[20]



creatine

Fig 8 Structure of creatine

There are two major Arg breakdown pathways. Primarily Arg is hydrolyzed by arginases I and II to produce ornithine and urea. This reaction is a part of the urea cycle and is important for managing nitrogen levels in the body (**Fig 5**). Urea formed in this reaction can be excreted via kidneys or perspiration to remove excess of nitrogen.

The secondary breakdown pathway of Arg is performed by nitric oxide synthetases (NOS). This reaction is the main endogenic source of nitric oxide (NO)[16]. NO is an endothelium-derived relaxing factor, a hormone important in blood flow regulation. Because of its gaseous form, NO has high membrane permeability. Additionally, NO is highly reactive, what makes it an ideal paracrine hormone for local signaling. Endothelial cells lining blood vessels employ NO to relax surrounding smooth muscles, leading to an increase of the blood flow.[13]

Maintaining cellular concentration of Arg is also important in the immune response. Multiple findings show that proper function of T-cell receptors is lost in conditions of Arg deprivation[21]. This is a mechanism by which certain types of cancer cells avoid immune response.

NOS activity combined with an influence of Arg on immune system as well as collagen and polyamine production render it crucial in wound healing. Mice starved of Arg for 3-4 weeks, show difficulty in closing even small lacerations[16].

1.3 Role of Arginine in cancer development

Cancer is not only a personal tragedy for almost 20 million new patients every year[22]. It also has a high negative impact on the lives of their relatives and friends. Each year life of dozens of millions of people worldwide is affected by oncologic diseases. Moreover, cancer-related productivity loss for year 2020, sums up to over 300 billion USD in the United States alone [23]. This makes fighting cancer illnesses a highly important problem to tackle. According to WHO data (**Fig 9**), cancer of the central nervous system (CNS) constituted only 1.6% of new cases in 2020, making it 19th most prevalent type of cancer[22]. Over 250 thousand deaths every year and a five-year survival rate of 36% point to lack of a successful therapeutic method.

Cancer is a result of deregulated cell proliferation. Since neurons themselves undergo division sporadically, the main source of CNS cancer are glial cells[24,25]. Glia is a tissue surrounding neuronal cells. Although full function of glia inside CNS is still being discovered, its main purpose is to provide ideal environment for neurons to perform their functions. Amongst all glial cells are responsible for providing neurons with nutrients as well as removing metabolic waste. Apart from controlling the environmental conditions in the CNS, glia performs functions of the immune system - it controls inflammatory response and counteracts pathogens. Additionally, glial cells are crucial for proper signaling in neuronal pathways. They support synaptic signaling by maintaining proper ion gradients and regulating neurotransmitter concentrations. Glial sheath is also crucial for signaling along axons as it insulates parts of the axon membrane speeding up signaling along axons.[26]

Glial cancer is one of the most difficult to treat or even impede its development. Its location and the way how it intertwines with a healthy brain tissue, renders surgical treatment practically impossible.[27] Because of that, a pharmacological approach, especially one that targets cancerous cells selectively, is often the only viable possibility.

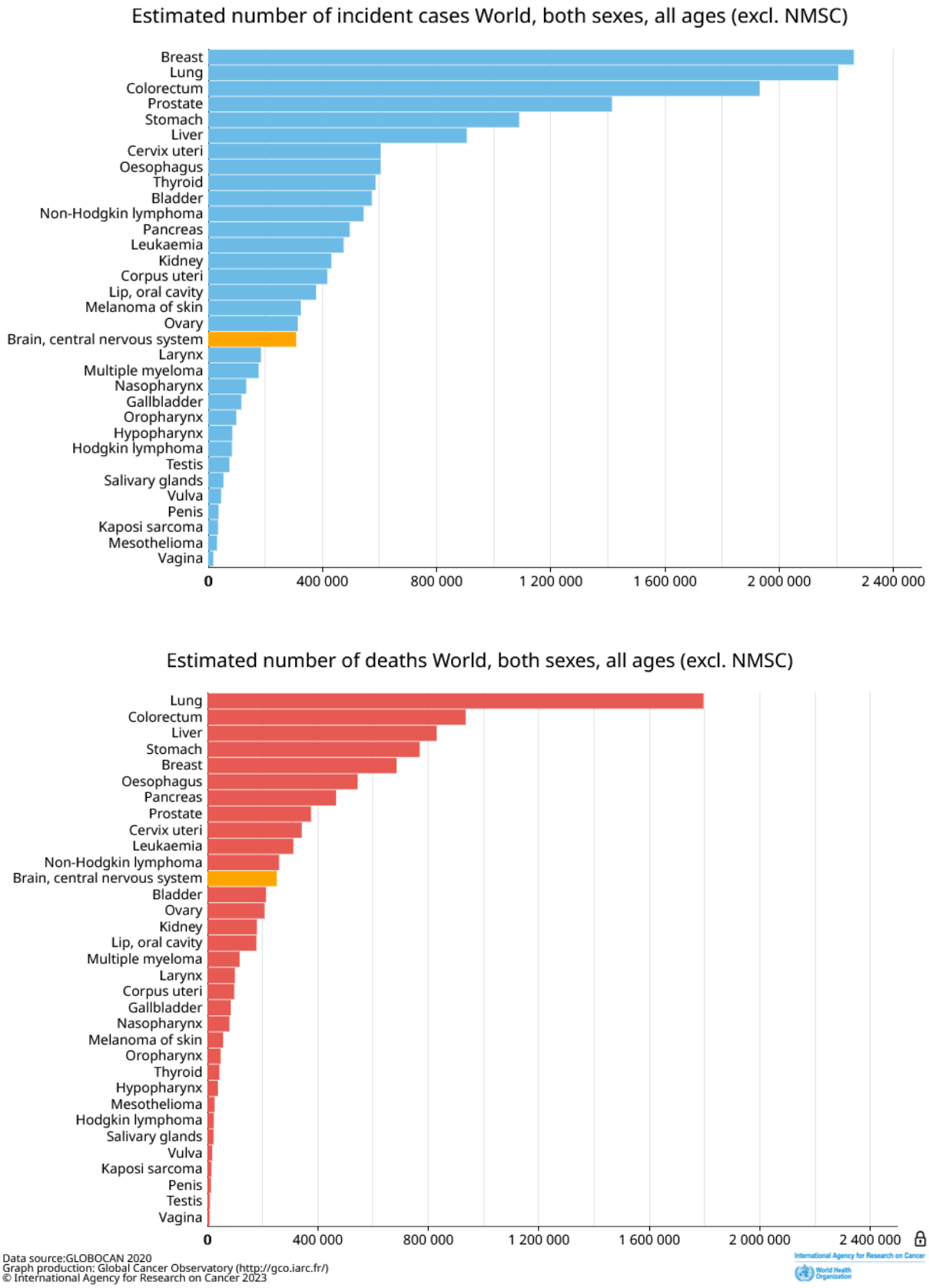


Fig 9 Contribution of CNS cancer in global cancer incidence [22]

Arg being so tightly interwoven into human metabolism is a valuable target in chemotherapeutic research. Two opposing approaches emerged in the last decade, both proved promising depending on right target selection. This puts emphasis on the fact that cancer is a common name for a large variety of diseases with different metabolic footprint. For that reason, proper diagnostics and accurate cancer type identification is as vital as a potent chemotherapeutic to achieve success in treatment.

One of the proposed treatments, targets the cancer cells that exhibit immunosuppressive traits[21,28]. Those cancer types show elevated transcription of arginase 1 and 2 genes. Activity of ARG1 and ARG2 depletes Arg available for immune system, leading to limited T-cell proliferation. T-cell response is necessary for inducing inflammation and apoptosis. A selection of novel arginase inhibitors has been proposed and tested by the Molecure company (former *OncoArendi* Therapeutics S.A.). Their detrimental effect on ARG1 and ARG2 activity proved to be enough for their formulations to enter second phase of clinical trials in treatment of sarcoidosis[29].

The second approach is based on an opposing strategy. It aims to disable rogue cancer cells by depriving them of their sole Arg source. Several cancer types including melanoma, hepatocellular carcinoma, mesotheliomas, renal cell carcinomas, and glioblastomas have been shown to lack expression of the gene encoding ASS, the enzyme crucial for endogenous Arg synthesis [30,31]. Such cells being functional Arg auxotrophs, rely solely on the Arg uptake from extracellular matrix. This makes Arg replacement a promising therapeutic goal. Disrupting intracellular Arg levels induces system-wide stress in cancerous cells. It has detrimental effect on cell morphology, adhesion and motility, as well as negative impact on function of mitochondria and endoplasmic reticulum.

Thus, the idea of local Arg deprivation therapy emerges, where PEGylated Arg degrading enzymes are introduced into tumor surroundings to cut external influx of Arg into cancer cells. Therapies based on this idea currently undergo phase I/II clinical trials in various malignant tumor cases[32–35].

As mentioned above, Arg is crucial for the proper actin filaments assembly. Actin filaments (termed *in vivo* as microfilaments) are long threads resulted from polymerization (assembly) of actin monomers[11]. Multiple threads of microfilaments create the backbone of the cytoskeleton. The cytoskeleton is a dynamic structure that creates scaffolding inside the cell. This scaffolding allows it to maintain cell shape and movement in a desired direction. The

cytoskeleton serves also as an anchor for proteins and organelles as well as a track for the motor protein-based intracellular transport. Due to the fact that arginylation affects actin filament organization, disruption in Arg levels leads to microfilament network disassembly, impairing cell morphology and thus motility[36]. Loss of ability of improperly arginylated actin to form filament bundles affects cell shape and corrupts function of multiple metabolic systems inside the cell. For example, poorly anchored endoplasmic reticulum is unable to properly fold produced proteins. Also, delocalized mitochondria are not functioning to their full capabilities. Furthermore, low Arg concentration also results in impairment of NO and polyamine synthesis. Lack of those signaling molecules impacts a plethora of different pathways. This accumulation of stressors can lead to cell death.[34,36,37]

Secondary impact of corrupted actin assembly is the improper function of focal adhesion kinase (FAK). FAK is a protein responsible for creating adhesion contacts that enable cells to interact with extracellular matrix. Cells utilize them to anchor themselves in their environment thus having impact on their motility, and in the case of cancers cells - on their invasiveness into surrounding healthy tissue.[38,39]

Despite being promising, Arg deprivation alone might not be sufficient to combat glioblastoma. Experience from *in vitro* research shows that Arg deprived glioblastoma cells can return to their full function as soon as 3 hours after Arg resupplementation[37]. Additionally, results from two-dimensional cell cultures do not translate linearly onto tests *in vivo* and on three-dimensional cell cultures[40].

Recently an idea of a therapy combining Arg deprivation with supplementing a compound mimicking Arg has emerged. Two natural nonproteinogenic plant amino acids (canavanine and indospicine) showed promising efficacy on cancer cell lines *in vitro*[37,40].

Canavanine (**Fig 10.**, Cav) is a non-biogenic amino acid produced by a number of leguminous plants (*i.e.* clover, alfalfa)[14]. It is accumulated mostly in the seeds of those plants to serve as a deterrent against the herbivores. It is toxic in high quantities due to its similarity to Arg. Being nitrogen rich compound, Cav also serves as a source of reactive nitrogen for the emerging embryo. Cav can constitute up to 10% dry weight of seed. Additionally, it is metabolized by arginases into canaline (**Fig 10.**), a potent insecticide. Structurally Cav is fairly similar to Arg, having an oxygen atom in place of the carbon at the δ-position of the sidechain. This miniscule difference allows Cav to maintain affinity to enzymes recognizing Arg but severely

disrupts charge distribution in the guanidine group. For that reason, Cav is incorporated into proteins in place of Arg, impairing their function[37].

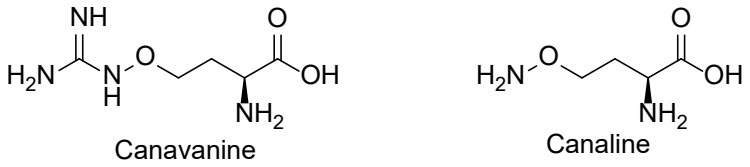


Fig 10 Structure of canavanine (Cav) and its metabolite – canaline

Cav has been previously tested as an anticancer agent on its own. The results have shown that its efficacy was too low, compared to its toxicity. For that reason, Cav's therapeutic window is too narrow for medical standards[41].

Cav impact on glioblastoma cell lines cultivated in Arg-free environment has been exhaustively studied by Karatsai et al.[37]. Proteomic studies proved incorporation of Cav into multiple proteins of Arg deprived glioblastoma cells, including actin. Disruption of actin filament assembly leads to multiple systemwide defects that can lead to the cell death. Influx of Cav into cancer cells also impacts the activity of endoplasmic reticulum and mitochondria. Moreover, damage to cell morphology, motility and function has been shown to be irreversible even upon Arg re-supplementation. Additionally, assays conducted on healthy rat glial cells showed limited negative impact, proving this therapeutic method to be selective towards cancerous cells[37].

Indospicine (**Fig 11**, Isp) is produced by a number of plants of the *Indigofera* species. Structurally, it differs from Arg**,** replacing the guanidine group of the sidechain with an amidine moiety. This subtle difference allows Isp to interact with all proteins recognizing Arg while being resistant to all enzymes degrading it.

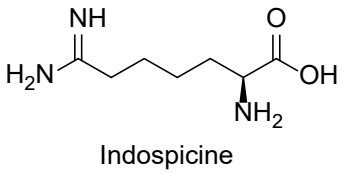


Fig 11 Structure of indospicine (Isp)

Recently, the research has been published on the Isp influence on colorectal cancer lines lacking ASS expression[40]. It includes *in vitro* tests on cells cultured in Arg free media, supplemented with Isp. In those conditions, Isp exhibits similar behavior to that of Cav, meaning Isp induces stress in endoplasmic reticulum and gets incorporated into proteins in the place of Arg. The major problem with Isp-based therapy is that Isp is inert to enzymes degrading Arg. For that reason, Isp can be toxic to healthy cells *via* its accumulation, as mammals are unable to decompose Isp. For that reason, prolonged exposure to Isp can lead to hepatotoxicity. It has been shown that dogs consuming meat of animals grazing on *Indigofera* die of liver failure induced by Isp.[42]

1.4 Cell death

Cell death is not necessarily negative outcome in multicellular organisms. In most of the cases it is evolutionary conserved process that maintains control over growth and regulates the numbers of the old and damaged cells. Loss of this function is underlying cause for most types of cancer as well as a number of neurodegenerative disorders[43].

Two major types of cellular death can be distinguished: necrosis and apoptosis. Each is characterized by specific signaling pathways being activated, however recent findings show that these processes can overlap or even occur simultaneously[44].

Necrosis occurs when stress induced on the cell overwhelms mechanisms responsible for repair and regulation. This deregulation results in cell losing its ability to maintain ionic gradients across the membranes. The resulting influx of water expands the organelles and ruptures the cellular membrane[43]. Teared cell leaks intracellular contents inducing inflammatory response in the surrounding tissue, resulting in damage of healthy surrounding cells[45].

Apoptosis is the most orderly pathway of cellular death. It induces severe changes in cell morphology (**Fig 12**). Firstly, cells undergo blebbing – they lose their shape, forming a series of protrusions in membrane (blebs). Later, individual blebs tear off the main cell body forming apoptotic bodies. In the meantime, organelles get dissolved including the fragmentation of the nucleus. This series of biochemical and morphologic changes that cells undergo during apoptosis allows for cleanup of toxic and inflammatory cell contents. For that reason, cell and its contents get removed without damaging or putting stress on surrounding tissue[43,46,47].

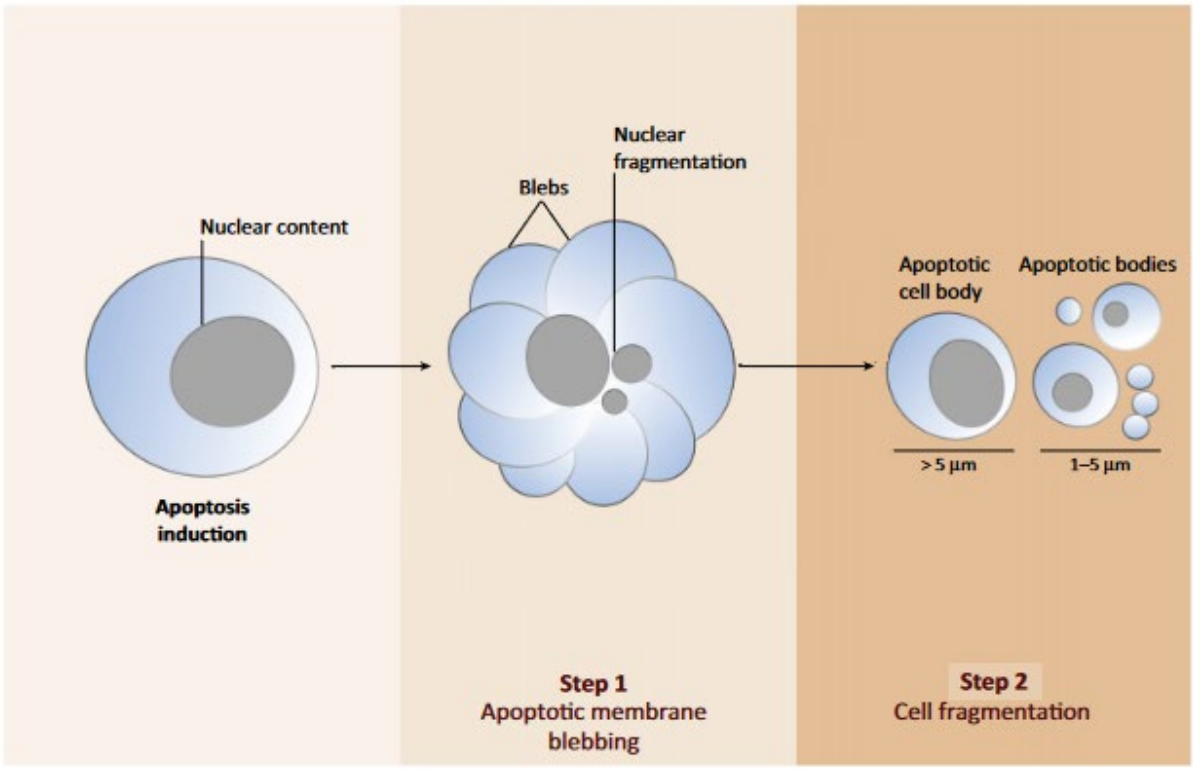


Fig 12 Changes to cell morphology during apoptosis (adapted from [47])

A cascade activation of cysteine-aspartic proteases (caspases) is a cornerstone pathway for apoptosis. They have been found upon activation to cleave over a thousand different proteins. This proteolytic activity of caspases sends the cell into apoptosis.[48]

From the caspase family of proteins two are important in the scope of this work – caspase 9(C9) and caspase 3(C3). Caspase 9 is an initiator caspase, its main function is to activate other proteins in pathway. C9 normally exists as a procaspase 9 monomer. It is activated by dimerization and subsequent autocatalytic cleavage that stabilizes the dimer. Caspase 3 is an executioner caspase, its activity influences various other proteins that result in apoptosis. In contrast, to C9, C3 naturally exists in a form of inactive dimer – procaspase 3. Cleavage by initiator caspase or other active C3 results in conformational changes that bring two active sites together allowing proteolytic function[48]. One of proteins cleaved by C3 is poly ADP-ribose polymerase 1 (PARP1). PARP1 function is DNA repair and inhibition of its function (by a cleavage) is necessary for apoptosis. The presence of cleaved PARP1 (cPARP1) is one of the known markers of apoptosis[49].

2. Tools

2.1 Target molecule design

The modern approach to drug design diverts into two distinct pathways: high throughput screening (HTS) and rational drug design (RDD)[50]. The former is more experimental, based on sifting through a vast number of known compounds to find structures that show activity. The latter employs comprehension of the biological mechanism underlying the disease and applies it to modelling *in silico*.

HTS is an older but still commonly used method in drug design. HTS relies on physically testing vast number of small organic molecules on a biological model to find ones with high efficacy measuring against the desired effect. HTS's major advantage is process simplicity. HTS does not demand full comprehension of the biology underlying the disease. *In vitro* and *in vivo* models of disease as well as proteins of unknown structure can be used as targets for screening. HTS's main disadvantage is the necessity to create and maintain an enormous library of compounds. Obtaining and proper storage of thousands of organic molecules is labor- and cost-intensive. The process of screening itself is also resource-exhausting - running tens of thousands of tests with minimal error count - and demands well-equipped facilities with properly trained staff. It remains however, the only viable method for diseases with an unknown molecular mechanism.

The method of RDD relies tightly on computational chemistry and biology. In RDD a three-dimensional model of a target protein is used in the molecular docking (MD) process to simulate interactions of this protein with a ligand. Ligands can be not only organic molecules but also proteins and peptides. Comprehensive knowledge of a molecular mechanism underlying the disease is needed for a successful RDD application. This knowledge allows for a selection of the target protein, which activity or structure can be modified to achieve the desired therapeutic effect. The advantage of this method is that the sifting process is conducted *in silico*. This cuts not only costs and time, but also makes this type of design process available for research units with limited resources. Moreover, the only limiting factor to the size of a tested library is the time allocated for the computer simulations. On the other hand, RDD is limited to target proteins with a known structure. Only a small percentage of proteins have been analyzed *via* X-ray crystallography, giving a proper target for MD. Although partial or even homologue models are often used due to the lack of a better target *per se*. Using partial

models introduces additional deviations from the real protein structure. Those deviations introduce additional degrees of freedom into an already complex model. Often false positive and false negative results are to be expected whenever *in silico* modelling is applied.

In this research I chose RDD as a design strategy. I had access to a powerful computer with a graphic processor that sped up calculations. Moreover, ATE1 enzyme is a promising docking target with a known structure (PDB entry: 4Q2T).

2.2 Molecular docking

Molecular docking (MD) is a process of an *in silico* modelling of interactions between organic compound and its protein target[50]. This iterative process simulates different molecule conformations and applies them onto a protein model surface. Possible modes of interactions are scored by various algorithms according to their probability and a number of interactions with the binding site. Over the last two decades a number of computational tools emerged with the application for MD. Three of the most popular are Schrödinger, DOCK and Autodock.

Schrödinger is a chemical simulation software suite developed by the Schrödinger company. It is a commercial tool that finds application not only in medicinal chemistry but also in molecular physics and material science. Schrödinger molecular docking tools are well-developed, mature software. Their major advantage is the ease of use. Schrödinger is user-friendly and does not demand intricate IT knowledge to produce accurate predictions. It is renowned for its accuracy and repeatability. However, Schrödinger is a sophisticated commercial product. It demands hardware with high computational power and the price for a user license is high. The cost of license and hardware to run this software limits its availability to research teams with extensive resources.

DOCK is an open license software for academic researchers[51]. DOCK has been developed since 1982 by Kuntz et al and is historically the first software designed for molecular docking[52]. It applies a geometric matching algorithm to find probable interaction sites. The program searches the surface of the protein to find the best match for the shape of the ligand. Over 40 years of development, gives DOCK a number of advantages:

- it is a mature software with a limited probability of bugs and errors,
- there is a huge background of successful DOCK applications in molecule design,
- employs the scoring algorithm updated with progress of the state-of-the art knowledge.

Autodock is also an open-source software developed by researchers at Scripps Research Institute[53]. It has been developed since 1989 and is the most widely used molecular docking program. Autodock is the docking engine used by World Community Grid [54], a project that allows people worldwide to voluntarily devote the processing power of their PCs to solve scientific problems. Being as well developed as DOCK, Autodock has a couple of advantages over its predecessor. Firstly, the Autodock scoring algorithm takes into account the ligand geometry and possible interactions between atoms. Secondly, Autodock has been updated to take advantage of modern multithread processors and graphic processing units (GPU). This shortens calculation times significantly.

As I have access to a PC with multithread processor and a powerful GPU, I have decided to use Autodock Vina software for my research.

2.3 Synthetic strategy

The chemistry of amino acids is a well-researched area. Being so prevalent in living organisms, amino acids are interesting synthetic targets for various applications. A common problem in the synthesis of amino acids and their derivatives is the accumulation of various reactive groups in a carbon chain of limited lengths. This often results in unwanted side reactions, rearrangements and can result in loss of chiral purity.

2.3.1 Protecting groups

All of the desired compounds in this research retain a general amino acid structure – an amine group, carboxylic acid group and a sidechain. These function groups are fairly reactive and require to be protected to avoid degradation and unwanted byproducts during multistep syntheses. For that reason, protecting groups (**Fig 13**, PG) are introduced into the structure during synthesis. Fortunately, the synthesis of peptides and amino acids is a well-researched area and a large number of PGs both for amine and carboxylic moiety has been identified[55]. Selecting an appropriate PG for designed synthesis is a tricky task and requires good insight into synthetic organic chemistry. The good PG must not only be inert in the reaction, but also allow for a clean removal in conditions not affecting the rest of the molecule.

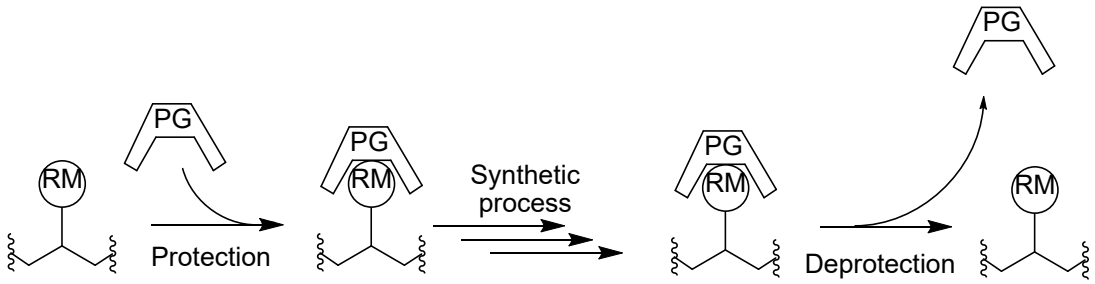


Fig 13 Schematic representation of the synthetic process including the protection-deprotection steps. RM – reactive moiety, PG – protecting group

Two of the most often used PG for amine moiety are *tert*-butyloxycarbonyl (**Fig 14 B**, Boc) and fluorenylmethyloxycarbonyl (**Fig 14 A**, Fmoc). These PGs require fairly common reagents and relatively mild conditions for both protection and deprotection process. Those two PGs differ in the stability window. Boc is inert in basic pH ranges and labile under acidic conditions. In the opposite Fmoc is inert in acidic pH and labile in basic conditions. For that reason, Boc and Fmoc cover most of the possible reaction conditions and are most widely used. Hundreds of protocols are available for syntheses employing those two PGs.

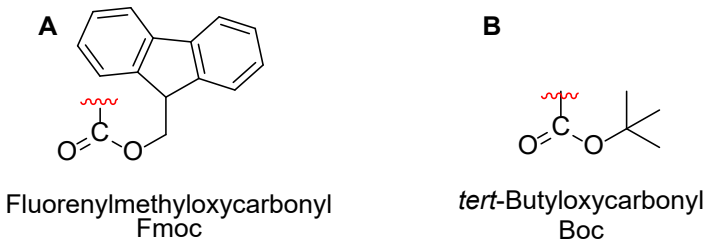


Fig 14 Structures of: **A**- fluorenylmethyloxycarbonyl and **B**- *tert*-butyloxycarbonyl

Methods of protecting and deprotecting amines with Fmoc most prevalent in literature are represented schematically in **Fig 15**. Worth noting is extremely simple protection step, it involves just Fmoc chloride and a protected compound, ethanol is added whenever solubility is an issue. As mentioned before, deprotection occurs under basic conditions. A wide variety of bases can be used depending on the application.

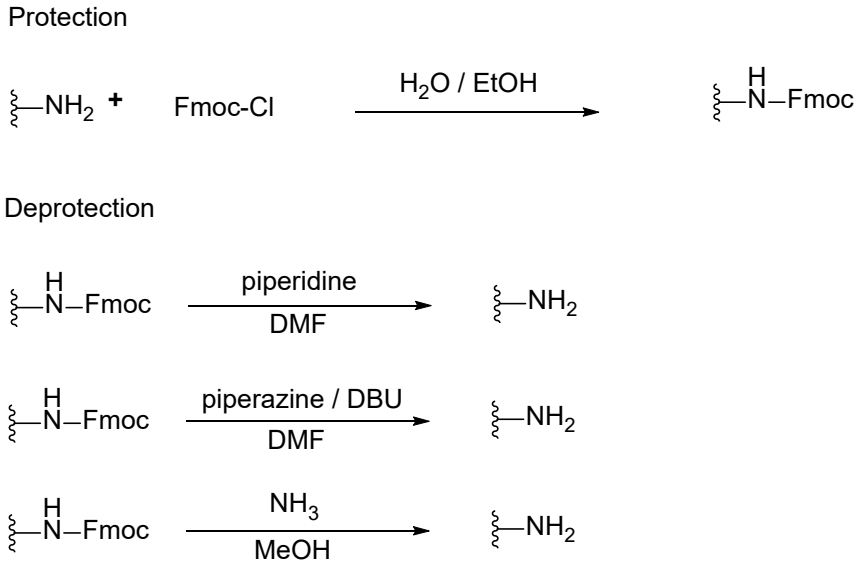


Fig 15 Fmoc protection and removal conditions[55]

Methods of Boc-protecting and deprotecting amines most prevalent in literature are represented schematically in **Fig 16**. The protection step involves the reaction of Boc anhydride with protected amine moiety in the presence of sodium or potassium ions. Deprotection calls for strong acidic conditions of *tri*-phosphoacetic acid in high concentrations (25%+) or chloric acid in high temperatures.

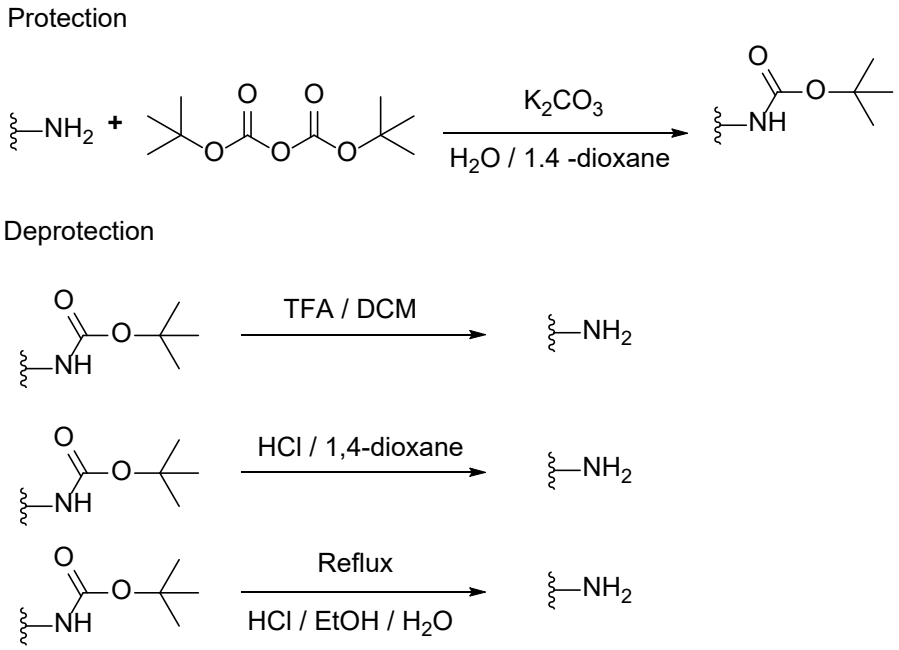


Fig 16 Boc protection and removal conditions[55]

Although Boc and Fmoc are the most common amine PGs found in the literature, a number of other PGs can be applied depending on the use case. Worth noting are the groups based on benzyl skeleton **(Fig 17**): benzyloxycarbonyl (Cbz), *p*-methoxybenzyloxycarbonyl (Moz), benzyl (Bn), *p*-methoxybenzyl (PMB) and 3,4-dimethoxybenzyl (DMPM) calls for hydrogenolysis. These groups are removed by hydrogenolysis, however each of them differs in liability. Benzyl-based PGs are used together with Boc/Fmoc when two or more amine groups exist in the synthesized molecule, and they need to be differentiated during the synthetic process.

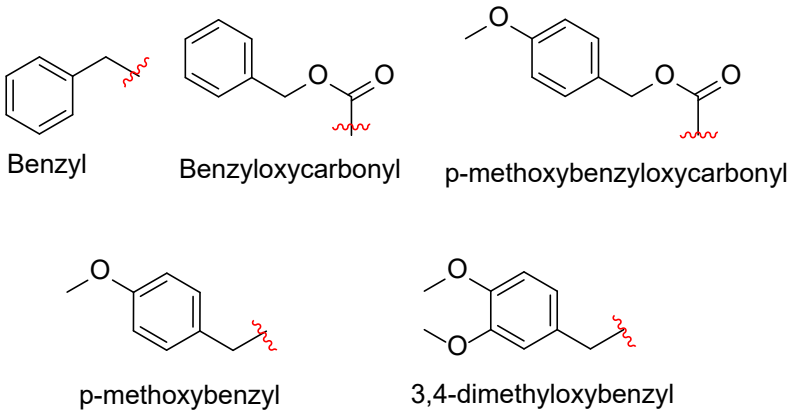


Fig 17 Structure of benzyl based protecting groups[55]

Methods of benzyl group protecting and deprotecting amines most prevalent in literature are represented schematically in **Fig 18**. Deprotection step with use of palladium catalyst is a very reliable method leading to pure products.

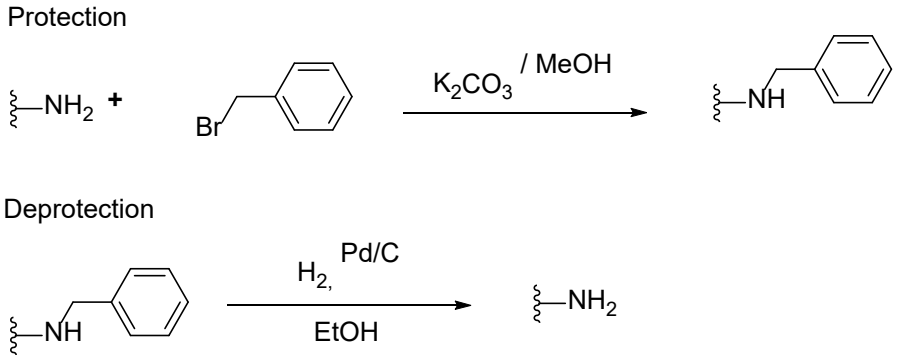


Fig 18 Exemplary protection and deprotection conditions of benzyl based PGs[55]

Regarding the carboxylic group protection, methyl and ethyl esters are usually stable enough for most common applications. Deprotection process sometimes employs lithium hydroxide in mild oxidative conditions (*i.e*. hydrogen peroxide). Whenever the application of lithium hydroxide is undesirable (*i.e.* the final product is soluble in water and lithium removal is difficult) benzyl esters are used. Methods of protecting and deprotecting carboxylic acids with the benzyl group most prevalent in literature are represented schematically in **Fig 19**. Worth noting is the similarity of the reactions with the ones applied to amine moiety protection.

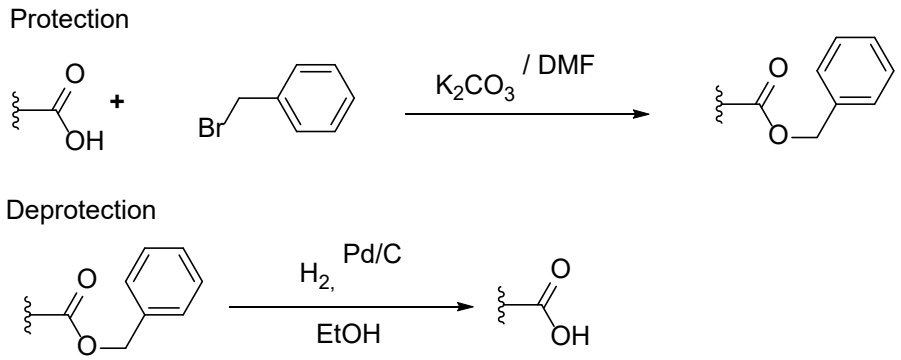


Fig 19 Benzyl ester protection and removal conditions[55]

2.3.2 Cyclic sulfamidates

Cyclic sulfamidates (**Fig 20**) are a class of small heterocyclic compounds with an oxathiazolidine ring. Cyclic sulfamidates are handy precursors of α-mono- or α,α-disubstituted amino acids. The oxathiazolidine ring can be opened by any nucleophile displacing sulfamidate. This reaction shows exceptional regioselectivity and retention of any stereogenic centers.

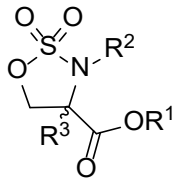


Fig 20 Structure of cyclic sulfamidate backbone

2.3.3 Click chemistry

The pharmaceutical and material chemistry industry is in constant pursuit of a method to generate vast libraries of compounds for screening purposes. The fastest way to build such a library is by combining smaller building blocks into different larger structures. This interest has sparked an invention of multiple reactions that are wide in scope and reliable. K.B. Sharpless defined the term "click chemistry" to describe such reactions[56]. To qualify as a "click" reaction must meet multiple criteria:

- works both in the small- and large-scale applications,
- is very high yielding,
- its potential byproducts are inert and can be removed by nonchromatographic methods,
- the reaction process is insensitive to presence of oxygen and water,
- has high thermodynamic driving force,
- its products are stereospecific,
- uses no solvent or benign solvent that is easily removed.

This image of universal reaction is unlikely to ever fit perfectly with every application. Despite that, a number of reactions have been found to adhere to those conditions (**Fig 21**). Notable examples are the Diels-Alder reaction (**Fig 21 A**) and nucleophilic substitution to aziridines (**Fig 21 B**). However, over the years one reaction has proven its versatile application to the degree that it is commonly named "the click reaction"[57] (**Fig 21 C**). Copper assisted azide-alkyne cycloaddition (CuAAC, **Fig 21 C**) is a variant of the Huisgen 1,3-dipolar cycloaddition catalyzed by copper(I) ions. The heart of CuAAC is the reaction between the terminal alkyne and azide that results in the formation of 1,2,3-triazole. Compared with classic Huisgen 1,3-dipolar cycloaddition, CuAAC undergoes at room temperature and maintains complete regioselectivity. CuAAC has been widely adapted to multiple applications and conditions. It finds use in small molecule synthesis[57], one-pot syntheses[58], ligating large biomolecules and tagging biomolecules[59–61]. The impact of CuAAC on modern chemistry has been recognized by the Nobel Prize committee. In 2022 Carolyn Bertozzi, Morten Meldal and Barry Sharpless received Nobel Prize "for the development of click chemistry and bioorthogonal chemistry"[62].

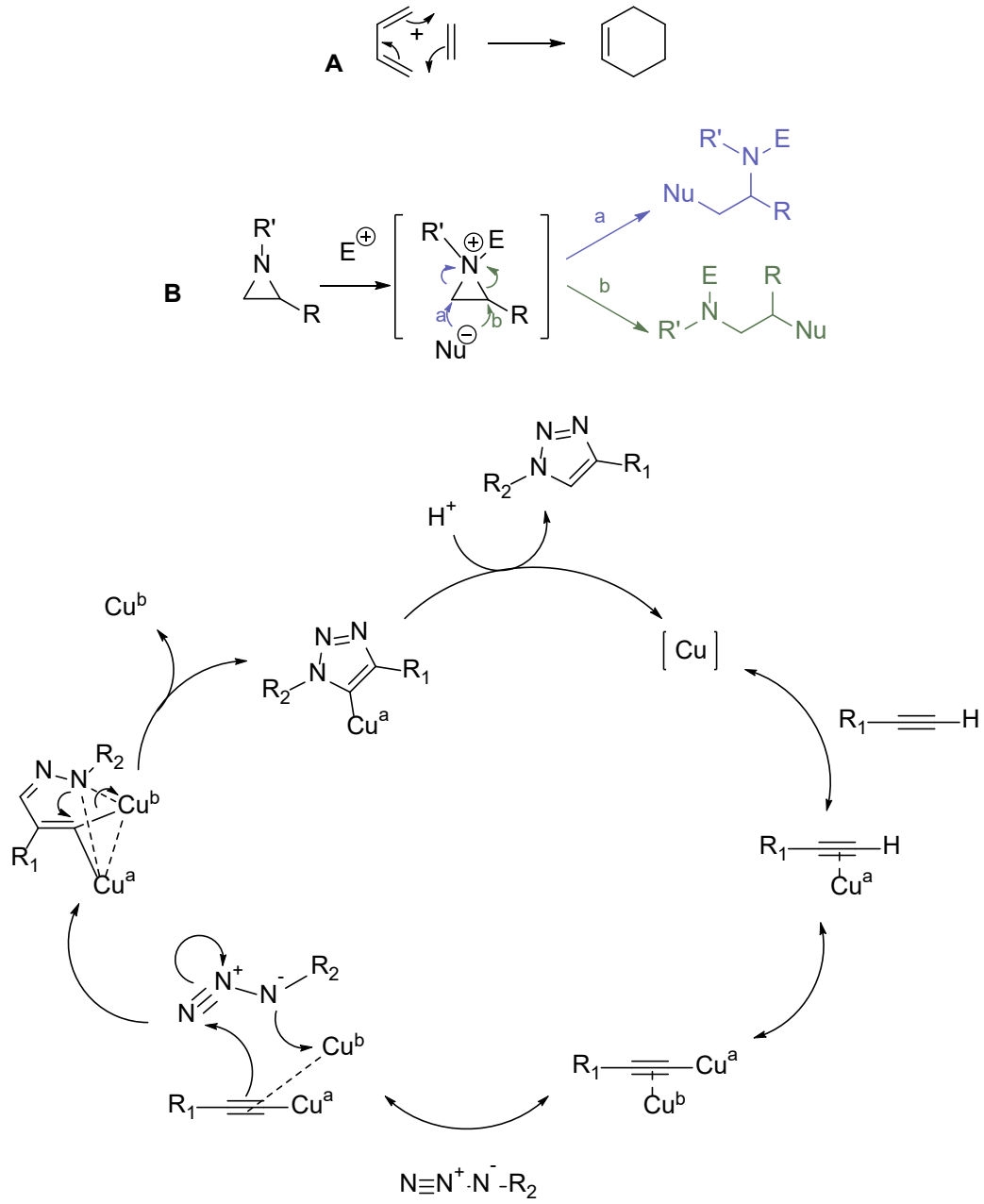


Fig 21 A – The Diels Alder reaction mechanism, **B** – the mechanism of nucleophilic substitution to aziridines, **C** – the copper assisted azide-alkyne cycloaddition mechanism.

2.3.4 Pinner synthesis

Amidines (**Fig 22**) are a subclass of oxo acids[63] where both oxygen atoms are replaced by nitrogen atoms. They are functional bioisosteres of guanidines. While retaining structural similarity to guanidines, amidines are noticeably less basic. Most widely used example of amidine is 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU) – cyclic amidine that acts as non-nucleophilic base in variety of reactions[64].

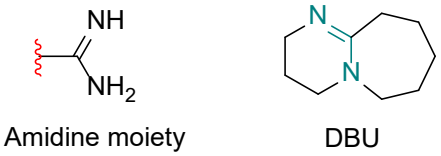


Fig 22 General amidine structure and 1,8-diazabicyclo[5.4.0]undec-7-ene (DBU)

Synthesis of amidines remains an under-researched area. Pinner synthesis, described first in 1877, remains to this day the only common route yielding amidine moiety[65,66]. In the Pinner reaction a nitrile reacts with an alcohol in acidic conditions, forming imino ester salt. This salt can undergo amonolysis to produce unsubstituted amidines (**Fig 23**, in blue) or aminolysis with a primary amine yielding monosubstituted amidines (**Fig 23**, in green).

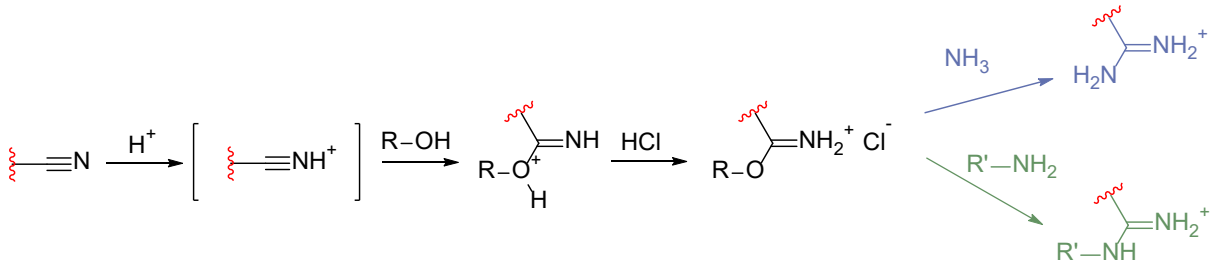


Fig 23 Pinner reaction pathway and its synthetic application

In my research the Pinner synthesis proved to be a fairly reliable route to amidines. Unfortunately, imino ester intermediate reacts readily with water producing carboxylic acid. For that reason, the reaction needs to be conducted in an inert gas atmosphere using dry solvents. This renders the Pinner synthesis labor-intensive and sometimes unpredictable.

2.4 Biological evaluation

To evaluate the impact of the compounds obtained during this work on the cancer cells, a range of testing methods has been employed. The methods were selected not only to test a possible efficacy in cancer cell lines, but also to prove that their mode of action is similar to Cav. For the sake of comparability to previous research on Cav[37], I have also chosen human glioblastoma cell lines U87MG and U251MG.

2.4.1 Cell lines

The Uppsala 87 Malignant Glioma (**Fig 24**) or U87MG is a cell line derived from human glioblastoma. It was first deposited in 1966 at the University of Uppsala[67]. It has been traced to 44-year-old female patient with glioblastoma. The U87MG has been acquired by the American Type Culture Collection (ATCC) in 1982. The ATCC sample has been widely used in research on brain cancer.

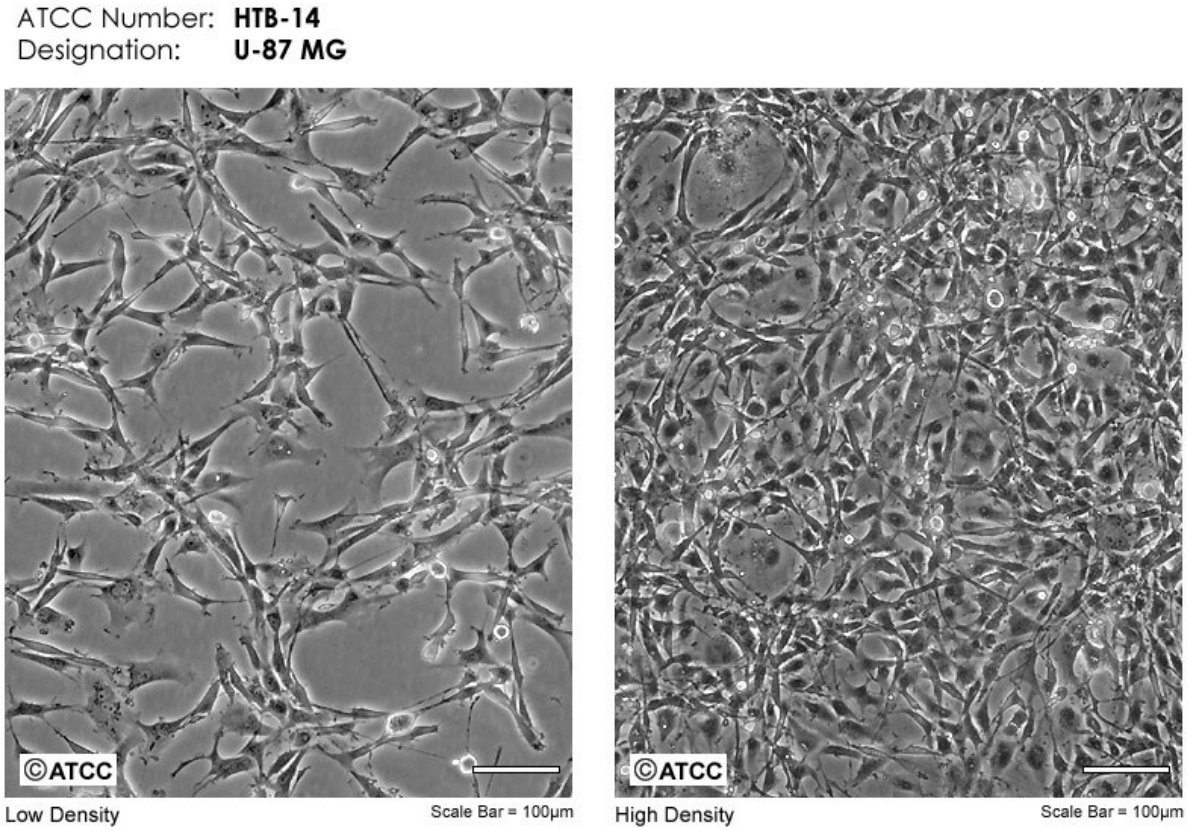


Fig 24 Microscope images of U87MG cell lines (adapted from ATCC)

The Uppsala 251 Malignant Glioma or U251MG (**Fig 25**) is another cell line of human glioblastoma[68]. It has been deposited in 1972 at the University of Uppsala and was derived from a male patient suffering from astrocytoma

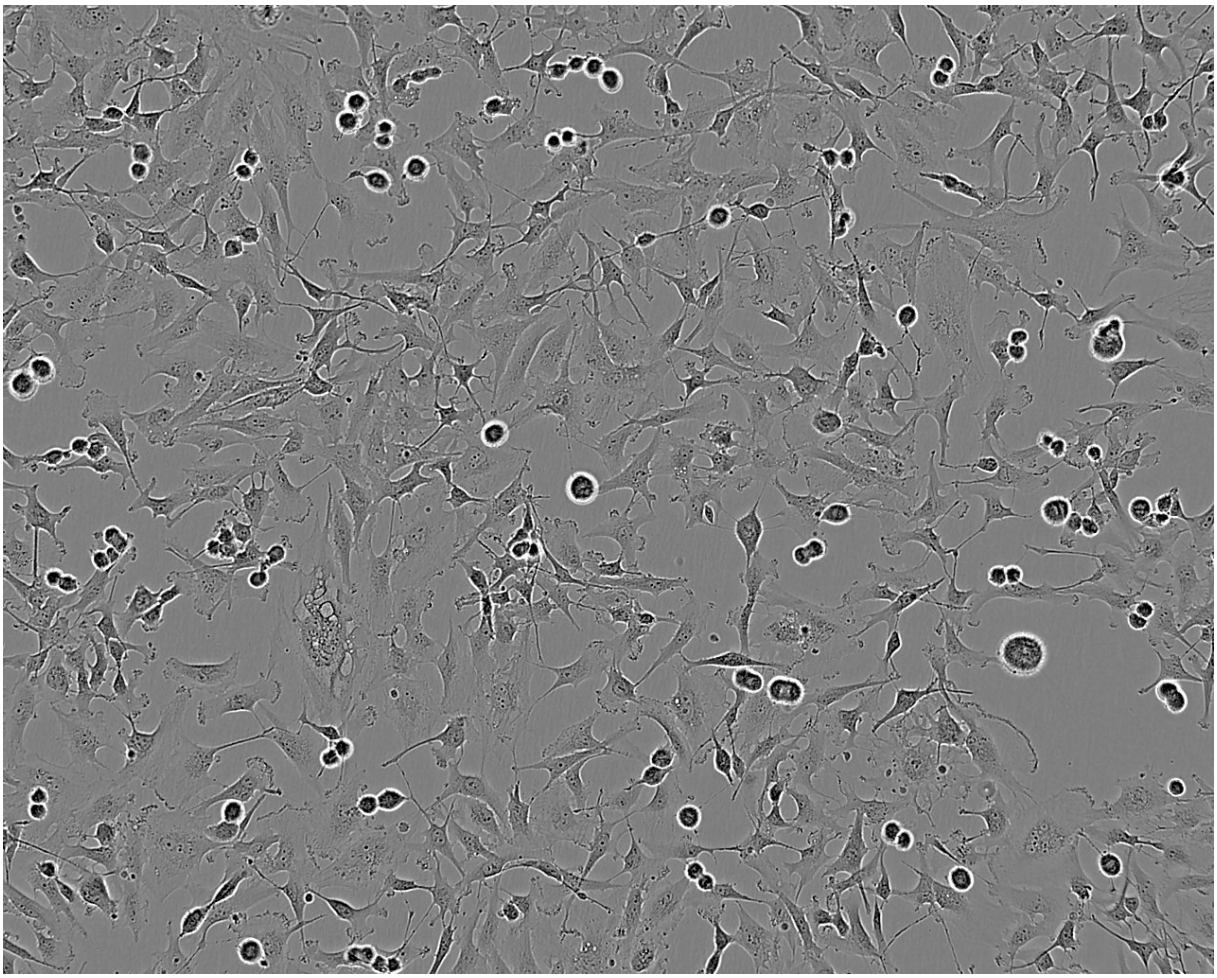


Fig 25 Microscope images of U87MG cell lines (adapted from European Collection of Authenticated Cell Cultures (ECACC))

Both U87MG and U251MG are well-studied cancer cell lines. The Scopus database search on U87MG returns 2803 results and a similar query on U251MG points to 3910 documents. This large background gives room to compare obtained results with agreed standards. Additionally, both cell lines exhibit disruption in Arg synthesis pathways. That makes them a perfect target for testing designed compounds.

2.4.2 MTS Cell viability assay

The MTS assay is the primary test for activity of obtained compounds used in this work. It is a colorimetric assay that quantifies the relative metabolic activity. MTS assay relies on a color shift that happens upon opening of tetrazole ring into a formazan structure (**Fig 26**). The rate of this redox reaction is directly associated with the activity of NADPH dependent dehydrogenase enzymes in the measured sample based on a reductive potential of the tested cell surroundings.

The MTS assay is a quick and reliable method for measuring cell viability. It can be easily conducted in parallel for dozens of samples. For that reason, it is the perfect method for measuring and comparing efficacy of compounds as well as determining their EC50.

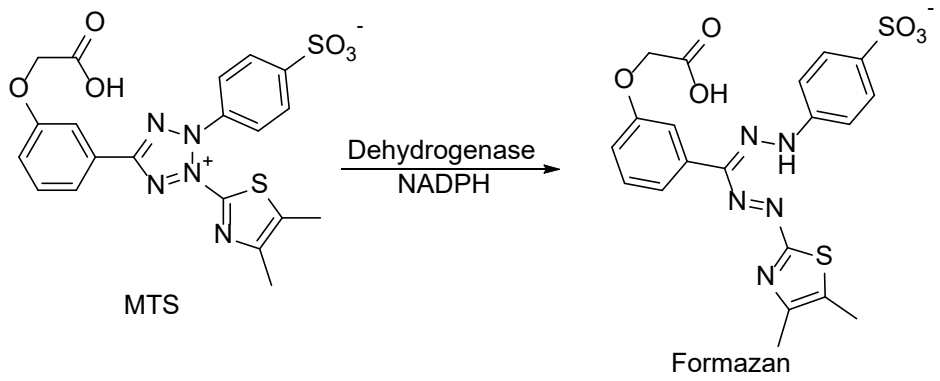


Fig 26 MTS -> formazan reaction

2.4.3 Western blot

One of the most widely used analytical techniques in molecular biology is the western blot utilizing the specific antigen-antibody interaction (WB, **Fig 27**). It allows to detect small quantities of a given protein in samples of tissue extracts and homogenates. Additionally, it allows to assign a protein concentration in relation to other marker proteins and overall protein concentration in the sample[69].

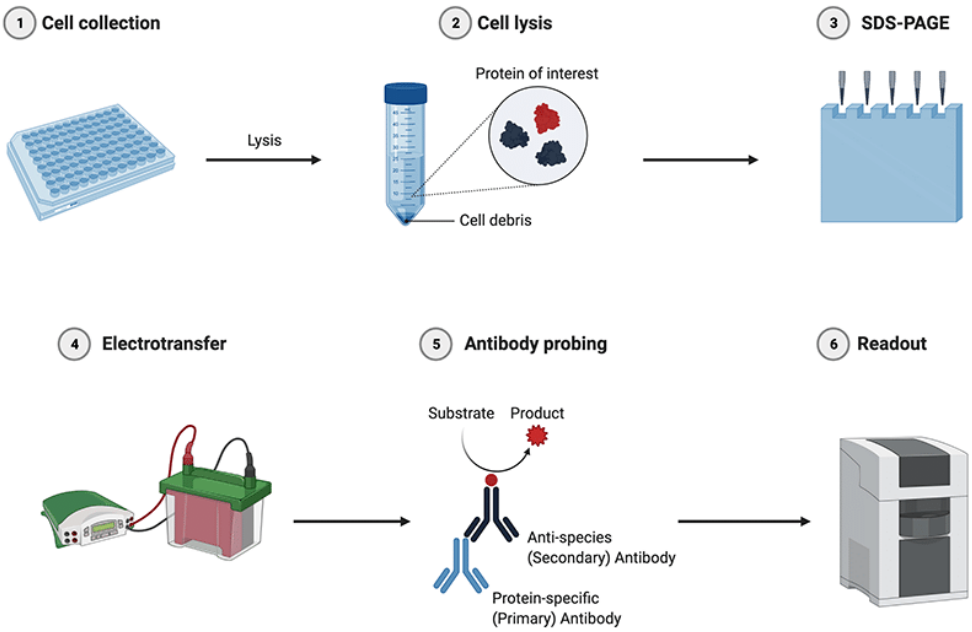


Fig 27 Schematic representation of WB method (adapted from [70])

A high signal-to-noise ratio is achieved by combining separation by size with detection using antibodies. That allows for detection of a single protein in the mass of thousands of different proteins and peptides, with as little difference between them as the presence of phosphoryl group. Ability to indicate multiple samples in one test allows WB to be not only qualitative but also a quantitative method. One of the downsides of WB is its time and labor-exhaustive method as a single test can take a couple days.

2.4.4 Microscopic imaging

Light microscopy (LM, **Fig 28.**) remains the staple method for working with the cells. Visual data brings a lot of information on the condition of a single cell as well as a whole dish. Optical microscopes are widely available in laboratories and LM does not call for any additional equipment besides that. LM is fast and noninvasive making it ideal for day-to-day sample control. Despite the techniques simplicity, choosing the right optical field to represent average conditions in the sample requires a significant experience.

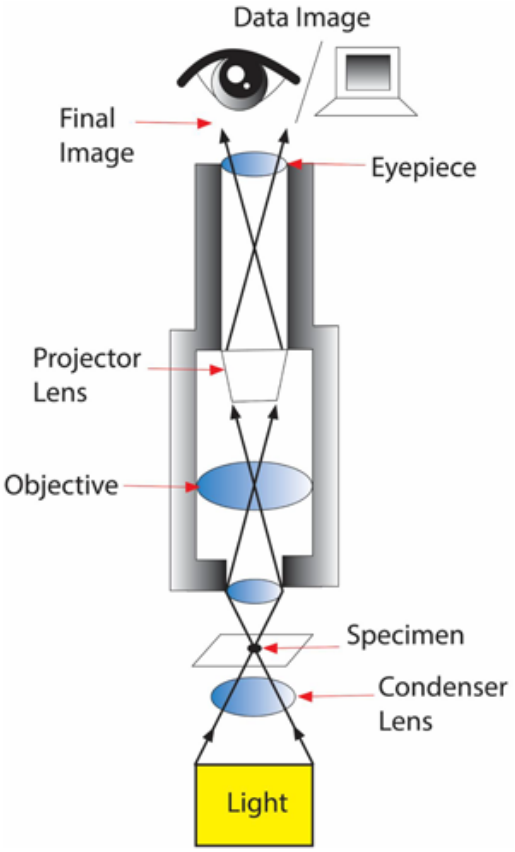


Fig 28 Light microscopy (adapted from [71])

3. Aims

The main objective of this work is to obtain a novel compound that has qualities of an anticancer agent. Using tools of medicinal chemistry, I aimed to design novel analogues of amino acid arginine, to synthesize most promising designs/candidates and to test their efficacy and mode of action *in vitro*.

Arginine, being one of twenty essential amino acids, is crucial for a plethora of metabolic pathways. In a healthy human tissue, arginine is synthesized endogenously. However, certain mutations in cancer cells make them, in contrast to the healthy untransformed cells, to be dependent on exogenous sources of arginine[30,31]. This opens a window for a therapeutic approach based on replacing arginine with a toxic mimetic compound[36,37,40]. Glial cancer carries a significant burden to society. Its location in CNS makes it difficult in diagnose and treatment. For that reason, there is constant and unmet demand for novel therapeutics that could improve the length and quality of life for millions of patients worldwide.

Glial cancer carries significant burden to the society. Its location in the CNS makes its diagnose and treatment difficult. For that reason, there is a constant and unmet demand for a novel specific therapeutic strategy that could improve the length and quality of life for millions of patients worldwide.

Overall research experience proves that successful drug design and testing is laborious and resource exhaustive. In my work, I decided to apply a couple of methods that allowed to minimize time and resources used.

Firstly, in the compound design process I applied the rational drug design approach. This method of designing compounds relies on an *in silico* modelling. Digital modelling and simulations allow to sift through large numbers of designs with the minimal time and resource demand. Candidate structures for synthesis are chosen basing on calculated binding energy (which is a measure of their affinity to ArgRS) and retrosynthetic analysis of their structure.

Secondly, to limit the resources used for synthesis of candidate compounds chosen during *in silico* part of my work, I applied “click chemistry” approach, which allows for obtaining a number of different structures based on the shared building blocks.[56]

Thirdly, to provide reliable and comparable data on biological efficacy of the tested compounds, I decided to follow already established tests and protocols from existing literature on Arg replacement therapy in human glioblastoma cell lines.[37]

The studies I performed allowed for obtaining a novel arginine analogue, which exhibit cytotoxic activity towards the examined glioblastoma cell lines at concentrations in micromolar range.

4. Results

4.1 Arginyl tRNA synthetase

To achieve the main goal of this project, I decided to focus on the compounds that would be disruptive to the cytoskeleton of the cancerous cells. As mentioned in chapter **2.3**, compounds mimicking Arg get incorporated into proteins, including that of the cytoskeleton. This process is represented schematically on **Fig 29**. Firstly, Arg is ligated with its transporter RNA ($tRNA^{ARG}$) by an enzyme arginyl-tRNA synthetase (ArgRS). This happens because enzymes responsible for Arg incorporation into protein lack the ability to recognize Arg**.** Both ribosome and ATE1 lack the ability to recognize Arg itself and require $tRNA^{ARG}$ tag to interact with it. The Arg **-** $tRNA^{ARG}$ conjugate serves as the identifiable substrate for protein synthesis and arginylation.

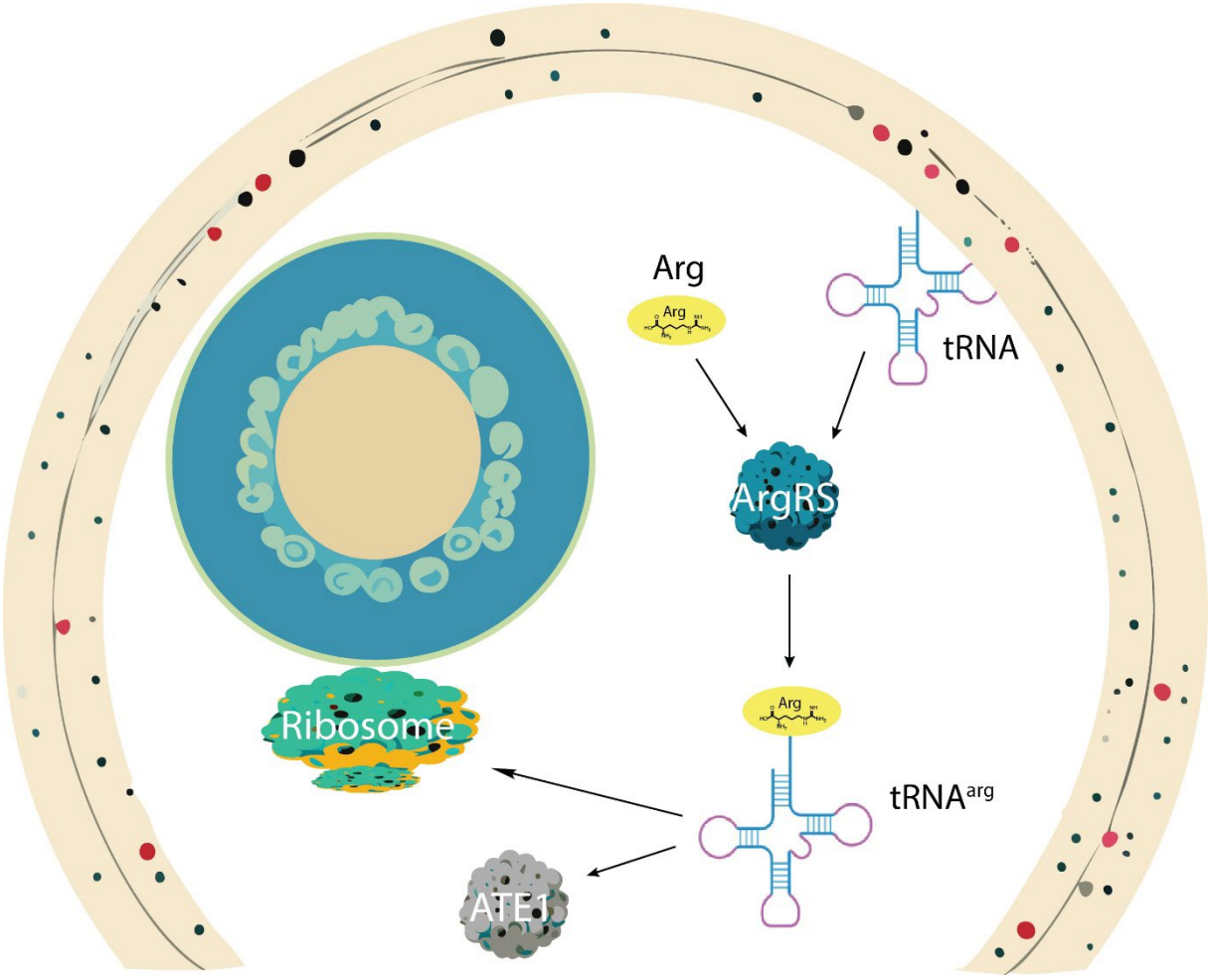


Fig 29 Process of Arg incorporation into protein structure

The only enzyme in this process that recognizes Arg structure is ArgRS. Its role is crucial for the action of all Arg analogues, and I am discussing it here in detail. ArgRS is an ATP dependent ligase that joins $tRNA^{Arg}$ to carboxylic acid group of Arg. The structure of human ArgRS is available with good resolution in protein data bank (entry 4q2t). For the above reasons, ArgRS was an ideal target for docking studies in my project. Surface representation of the Arg binding pocket of ArgRS is displayed on **Fig 30**. Crucial interactions between Arg and aminoacids in the binding pocket are displayed on Table **1**. Worth noting are numerous interactions of the Arg's amino acid moiety. Dual interaction with asparagine 130 (Asn 130) anchors it well in the binding pocket and synergizes with histidine 139 (His 139) in pushing negative charge onto carboxylic acid. This change in charge distribution activates the carboxylic acid, making it easier to ligate with $tRNA^{Arg}$. Finally, there is a symmetric interaction of guanidine moiety with aspartate 316 (Asp 316). This symmetry allows for strong binding regardless of the tautomeric form of guanidinium.

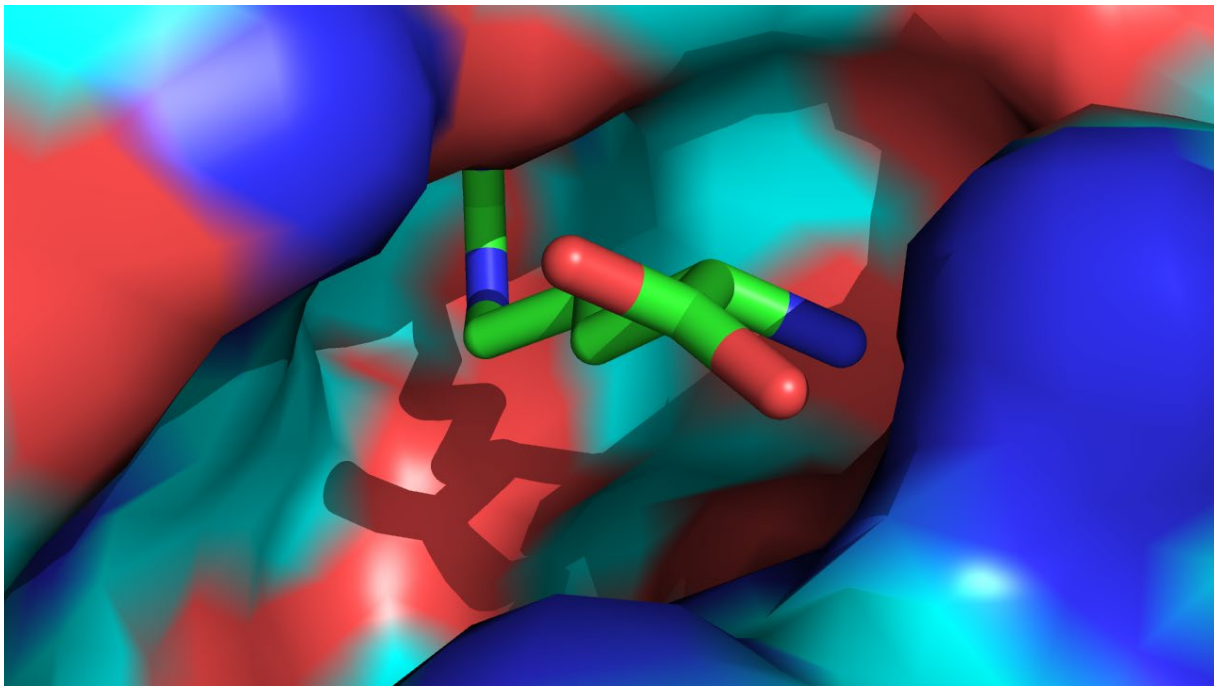
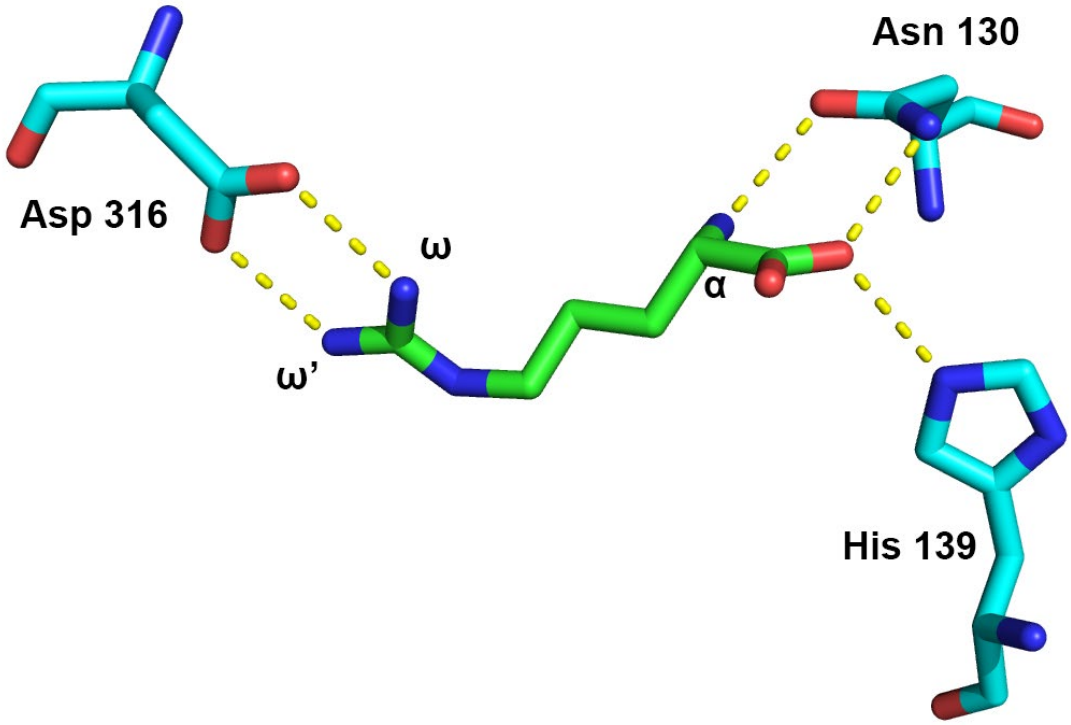


Fig 30 Surface representation of Arg binding pocket of the ArgRS (visualization based on PDB entry 4q2t) Colors: C(ligand)=green, N=dark blue, Inert protein surface = cyan, negatively charged protein surface=dark blue, positively charged protein surface=red

Table 1 Arg inside the binding pocket of ArgRS (visualization based on PDB entry 4q2t)



Residue	Atomic group	Distance (Å)[b]
(Asn130)O^{g}	(Nα)H	2.9
(Asn130)N^{g}	(C1)O^{1}	2.8
(His139)NH	(C1)O^{1}	3.1
(Asp316)O^{g1}H	(Nω)H	3.1
(Asp316)O^{g1}H	(Nω')H	3.0

[a]Colors: C(ligand)=green, C(protein residue=cyan, N=dark blue, O=red, H-bond=yellow.
[b]From the crystal structure.

From the point of view of my research, a number of important conclusions on structural components of Arg molecule can be drawn from the structure of the ArgRS binding pocket:

- amine moiety serves as an important anchor point,
- carboxylic acid is crucial for binding to $tRNA^{ARG,}$
- maintaining the L-configuration on stereogenic carbon is necessary,
- the most crucial interaction is at the end of the side chain,
- possible interactions of the side chain backbone are fairly limited.

4.2 Compound design

Basing on the above conclusions, I have designed a general structure of my target compound. This design is represented on **Fig 31**. I decided that all compounds will retain L-amino acid configuration (**Fig 31** in green) with a guanidine isostere (**Fig 31** in blue) at the end of the side chain. Guanidine isostere should be a planar structure with either hydrogen bond donor or hydrogen bond acceptor. Wide space in the binding pocket allows for various sidechain backbones (**Fig 31** in red) to serve as a linker between two interaction centers.

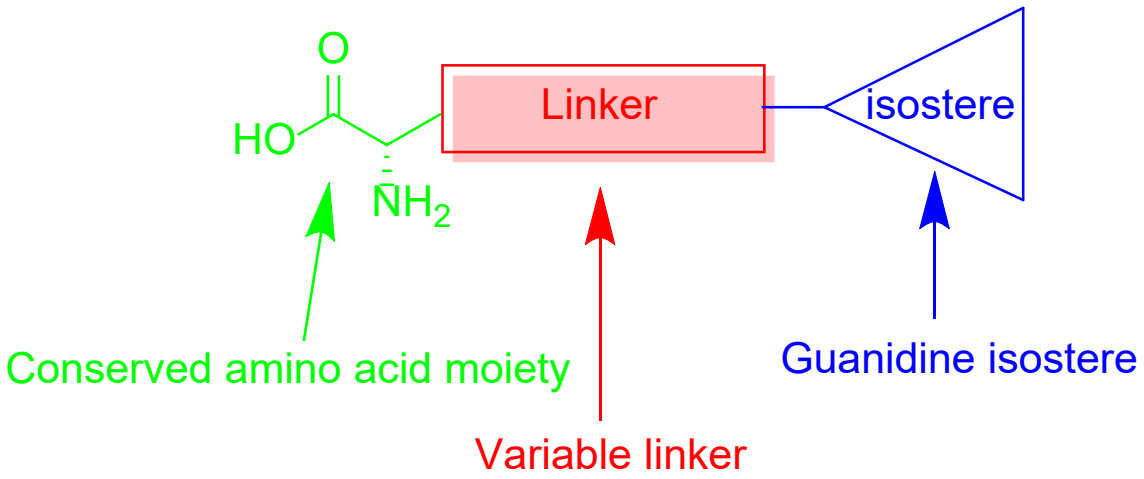


Fig 31 General structure of the designed analogues. Conserved amino acid moiety. Variable linker. Guanidine isostere at the end of the sidechain.

Having results of the above *in silico* investigation in consideration, I have designed three groups of compounds:

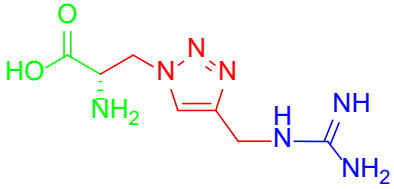
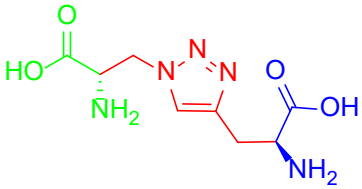
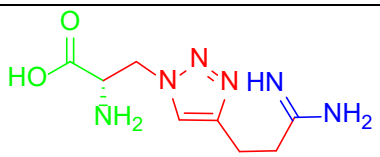
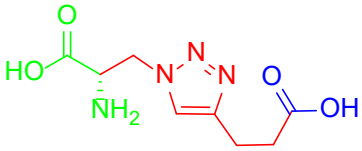
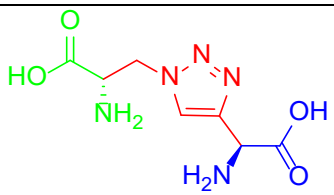
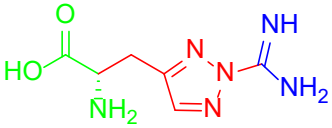
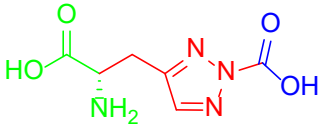
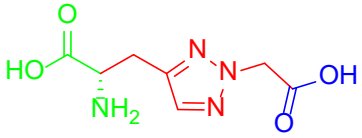
1. Arg analogues with 1*H*-1,2,3-triazole ring in the side chain backbone (**Table 2**),
2. selected natural or synthetic amino acids with aliphatic sidechains and their homo-analogues (**Table 3**),
3. Arg analogues with 1*H*-1,2,3-triazole or tetrazole moiety as guanidine bioisostere (**Table 4**).

The first group is composed of compounds **1** retaining general L-amino acid structure. They are systematically represented in **Table 2.** Major variability in this group is a type of moiety

that mimics the guanidinyl. Their common trait is 1*H*-1,2,3-triazole linker. This linker allows for click chemistry in the synthetic approach. For that reason, all of the compounds in this group can be synthesized from a common substrate. There are two main advantages of this approach. Primarily, it significantly cuts the time requirement to obtain all of the variants. Additionally, it allows to test various guanidine bioisosteres with limited devotion of resources and labor.

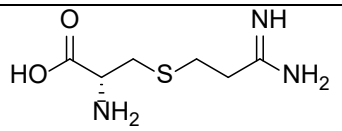
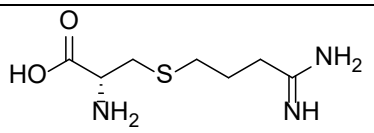
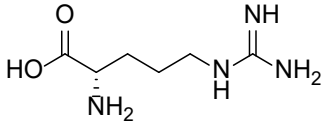
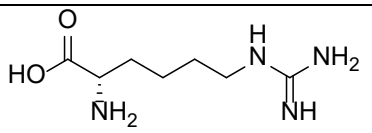
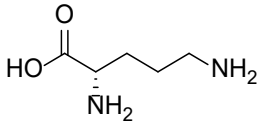
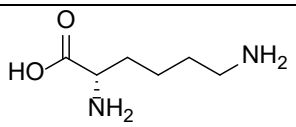
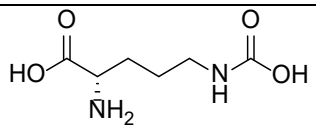
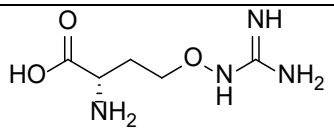
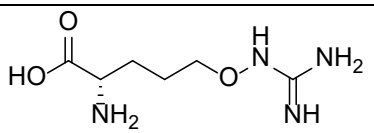
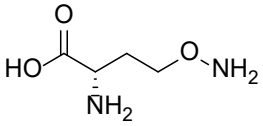
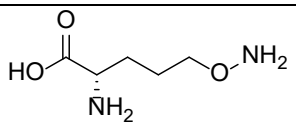
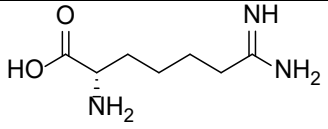
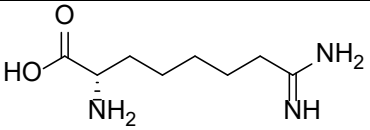
(Table with structures of important compounds discussed in this dissertation is available in appendix **8.1**)

Table 2 Structures of compounds from group **1**.

No	Scheme	Type of guanidinyl mimic
1a		guanidine
1b		α-amino acid
1c		amidine
1d		carboxylic acid
1e		α-amino acid
1f		amidine
1g		carboxylic acid
1h		carboxylic acid

The bulk of the second group of compounds **2** are natural amino acids and their homo-analogues. They are systematically represented in **Table 3.** These aliphatic structures are either Arg analogues or their metabolites. The only exceptions here are compounds **2a** and **2b**. These are novel structures, and their design was inspired by Isp (**2l**). Introduction of sulfur atom into position γ of Isp (**2l**) significantly simplifies synthesis.

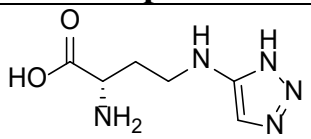
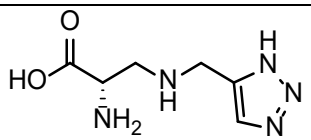
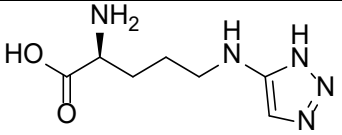
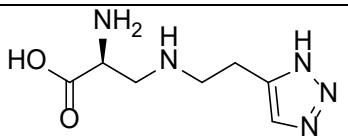
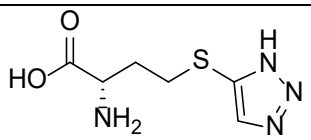
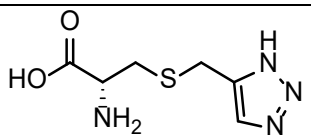
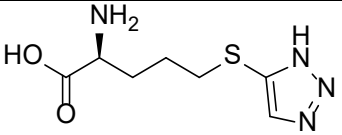
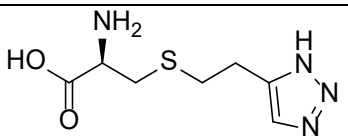
Table 3 Structures of compounds from group **2.**

No	Scheme	name
2a		*S*-(3-amino-3-iminopropyl)-L-cysteine
2b		*S*-(4-amino-4-iminobutyl)-L-cysteine
2c		L-arginine
2d		L-homoarginine
2e		L-ornithine
2f		L-homoornithine
2g		*N*-carboxyl-L-ornithine
2h		L-canavanine
2i		L-homocanavanine
2j		L-canaline
2k		L-homocanaline
2l		L-Indospicine
2m		L-homoindospicine

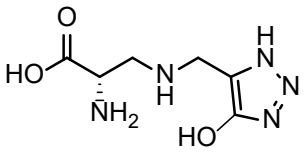
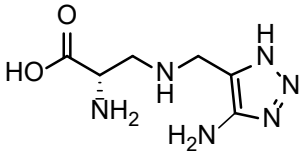
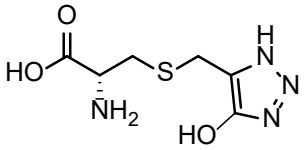
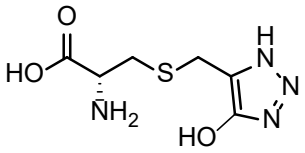
The third group of compounds **3** is represented systematically in Table **4a-d**. This group composes of variants combining the general L-amino acid with several linkers and guanidine isosteres based on 1*H*-1,2,3-triazole or tetrazole ring structure. Ten different linkers of lengths of three or four atoms incorporate nitrogen or sulfur atom in varying positions. Compounds **3** were designed to test the binding pocket *in silico.* This systematic set of compounds might prove to be a precious tool for the docking studies. Results of the molecular docking of these compounds will allow exploring the possible interactions inside the binding pocket.

Table 4 Structures of compounds from group **3**

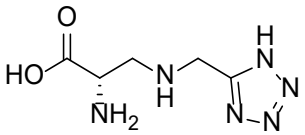
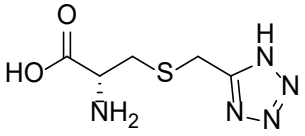
a) 5-nonsubstituted *NH*-1,2,3-triazoles

No	Compound	Linker
3a		1-azaprop-1,3-diyl
3b		2-azaprop-1,3-diyl
3c		1-azabut-1,4-diyl
3d		2-azabut-1,4-diyl
3e		1-tiaprop-1,3-diyl
3f		2-tiaprop-1,3-diyl
3g		1-tiabut-1,4-diyl
3h		2-tiabut-1,4-diyl

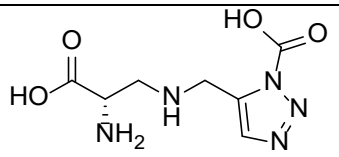
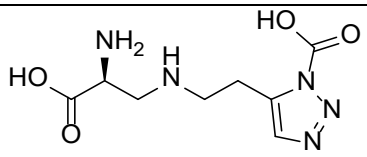
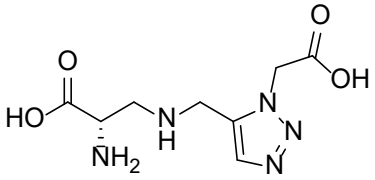
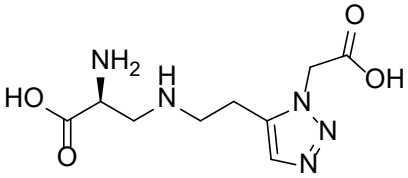
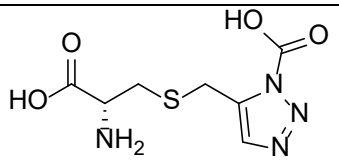
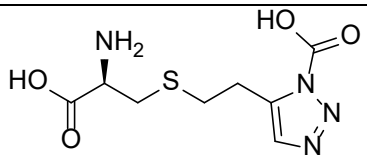
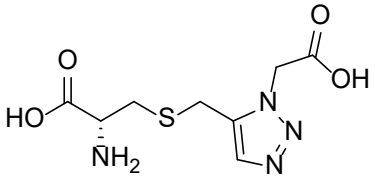
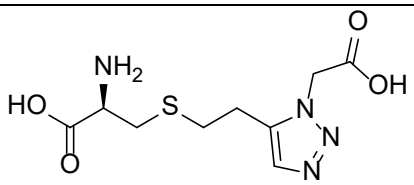
b) 5-substituted *NH*-1,2,3-triazoles

No	Scheme	Linker
3i		2-azaprop-1,3-diyl
3j		2-azaprop-1,3-diyl
3k		2-tiaprop-1,3-diyl
3l		2-tiaprop-1,3-diyl

c) tetrazoles

No	Scheme	Linker
3m		2-azaprop-1,3-diyl
3n		2-tiaprop-1,3-diyl

d) 3-substituted *NH*-1,2,3-triazole

No	Scheme	Linker
3o		2-azaprop-1,3-diyl
3p		2-azabut-1,4-diyl
3q		2-azaprop-1,3-diyl
3r		2-azabut-1,4-diyl
3s		2-tiaprop-1,3-diyl
3t		2-tiabut-1,4-diyl
3u		2-tiaprop-1,3-diyl
3v		2-tiabut-1,4-diyl

4.3 Docking studies

Next, the three groups of compounds **1-3** from the previous chapter (Tables 2-4) were studied *in silico*. Every compound has been docked against the crystal structure of the human ArgRS obtained from the Protein Data Bank (PDB) entry 4q2t. I have used version 1.2.0 of AutoDock Vina as a docking engine. Target protein structure has been prepared for docking using AutoDockTools v 1.5.7. Most of the structures available at PDB do not include hydrogen atoms, thus preparation step for docking always involves introducing polar hydrogen atoms into protein structure. Additionally, Arg and glycerin molecules have been removed from the binding pocket. The search box location has been chosen to include the location of Arg molecule in the original structure as well as all amino acids within 4 Å distance from it. The three-dimensional structure of each docked compound has been generated using AutoDockTools and its rotamer optimized before docking. Structure **2c** is identical with Arg (**Table 5** marked in blue) and serves as control for the docking parameters as well as to obtain benchmark docking score to compare. Each structure has been docked using random seed in two modes:

-with rigid protein structure (**Table 5** "Rigid"),

-with flexible amino acids residues within search box (**Table 5** "Flexible").

Analysis of docking results for each of structures docked in this work are available in the appendix **8.3.**

Out of 43 docked structures, 28 docked in alignment with the position of Arg in available ArgRS crystal structure. The results of the docking are all available in appendices 8.3. Autodock Vina scores output docking modes by assigning predicted binding energy. Because this calculation is conducted on a static structure it is fairly inaccurate and should not be considered for any other purpose than comparing the docking results. Calculated values of the binding energy for all structures are combined in **Table 5.**

Structures from group **1** achieved the most promising results, with binding modes most similar to those of Arg (**Table 1**). Compounds **1e (**-6.7 kcal/mol**)** and **1f** (-7.1 kcal/mol**)** display the lowest value of the predicted binding energy *i.e.*, have highest affinity towards the receptor.

All structures from group **2** docked in alignment with the position of Arg in the crystal structure of the receptor. The values of their binding energy were slightly higher (less affinity towards the receptor) than those of **2c** but on comparable levels.

A significant number of group **3** compounds had difficulties fitting inside the active site. This is not true for all of the compounds. This fact shows that the size of the azole ring is at the limit for the capacity of this binding pocket. Some structures align properly to flexible receptor and do not fit in rigid (**3a**, **3c**, **3e**, **3i**, **3o**, **3r**), some fit inside binding pocket but out of alignment (**3b**, **3d**, **3f**, **3k**, **3l**, **3s**, **3t**, **3m**, **3n**) and some do not enter the binding pocket at all (**3g**, **3h**, **3p**, **3u**, **3v**). This dramatic difference between fairly similar compounds suggests that azole structure is borderline size to fit inside the binding pocket.

Three general binding modes could be distinguished from the results.

- The first mode is the most desired one. It presents compound aligned with the Arg position in the PDB crystal structure. As well as maintaining carboxylic acid moiety location. **Table 5** rigid **1d**, **1e**, **1h**, **2a-2m**, **3j**. **Table 5** flexible **1a-1h**, **2a-2i**, **2k-2m**, **3a**, **3c**, **3e**, **3i**, **3j**, **3o**, **3q**, **3r**.
- The second mode, I defined as “out of position”. Compound structure generally aligns with the Arg position in the PDB crystal structure but position of the carboxylic acid deviates strongly. For that reason, there is no probability for the protein to ligate the structure to $tRNA^{ARG}$. Such results are marked with an asterisk. **Table 5** rigid **1a**, **1b**, **1c**, **1g**, **3b**, **3c**, **3f**, **3l**, **3m**, **3n**, **Table 5** flexible **2j**, **3b**, **3d**, **3f**, **3k**, **3l**, **3s**, **3t**, **3m**, **3n**
- The third mode, I defined as “out of pocket”. There is no alignment between the docked structure and Arg position in the PDB crystal structure. Such results are marked with a double asterisk. **Table 5** rigid **3a**, **3d**, **3e**, **3g**, **3h**, **3k-3t**. **Table 5** flexible **3g**, **3h**, **3p**, **3u**, **3r**.

Table 5 Results of molecular docking. Calculated binding energy for each structure with rigid target receptor (column "rigid protein structure") and with flexible target receptor (column "flexible protein structure"). Results without an asterisk align with the position of Arg in the crystal structure. Arg binding energy was calculated for comparison, highlighted in blue.

	Binding energy ΔG (kcal/mol)	
Compound No	rigid protein structure	flexible protein structure
1a	-6.0*	-6.5
1b	-5.5*	-6.5
1c	-6.2*	-6.4
1d	-5.4	-6.2
1e	-5.4	-6.7
1f	-5.7	-7.1
1g	-5.3*	-5.9
1h	-5.8	-5.3
2a	-5.4	-5.9
2b	-4.5	-5.6
2c (Arg**)**	-6.0	-6.0
2d	-5.6	-6.0
2e	-4.0	-5.1
2f	-4.7	-5.2
2g	-5.3	-5.7
2h	-5.1	-5.9
2i	-5.4	-6.3
2j	-4.4	-4.8*
2k	-4.7	-5.1
2l	-5.4	-5.8
2m	-5.4	-5.1
3a	-5.3**	-6.0
3b	-5.2*	-6.7*
3c	-5.1*	-4.7
3d	-5.3**	-6.5*
3e	-5.3**	-5.6
3f	-5.0*	-5.7*
3g	-5.1**	-5.5**
3h	-5.2**	-5.1**
3i	-5.7*	-4.5
3j	-4.8	-6.0
3k	-4.9**	-6.0*
3l	-5.0**	-6.0*
3m	-5.2*	-5.9*
3n	-5.0*	-5.6*
3o	-5.6**	-5.9
3p	-6.0**	-6.0**
3q	-5.3**	-5.8
3r	-6.1**	-4.8
3s	-5.3**	-5.9*
3t	-5.6**	-5.9*
3u	-5.6**	-5.2**
3v	-6.1**	-5.8**

*Structure docked out of position. **Structure docked out of binding pocket.

As a general rule, docking with flexible amino acid residues within the binding pocket resulted in a more preferred outcome, *i.e.* proper positioning and higher score. This is predictable behavior; in nature protein structure is dynamic and to a certain degree adjusts to the ligand. That results in less strain on ligand conformation and allows interaction centers to adjust position. However, in depth analysis reveals that there are a couple exemptions to this.

- Compound **1h** is the only one that shows a higher affinity to a rigid than flexible receptor. This might be caused by short rigid side chain being forced into contact with residues in proximity.

- Compound **2j** docks out of position to a flexible receptor while simulation binds properly to a rigid protein. This is probably caused by the shorter sidechain of the canaline (**2j**) that does not allow it to interact with residues at the end of the pocket. Rigid receptor structure forces **2j** into proper position while flexible allows it to find more energetically favorable position.

- Compounds **3g, 3h, 3p, 3u, 3v** do not align inside the binding pocket in neither rigid nor flexible mode. This might be a result of the side chain being too long, either due to the long linker (**3g**, **3h**, **3p**) or significantly big substitutes in the azole ring (**3u**, **3v**). The size of the sidechain does not allow molecules to conform to the binding pocket.

To sum up, the performed docking simulations show the results of my initial *in silico* investigation (chapter **4.1**) correct.

- Maintaining L-amino acid structure seems to force carboxylic acid moiety into proper position.
- The most important of the obtained data regards the group replacing the guanidine. As the binding pocket narrows down at its end, the guanidinyl isostere needs to be a planar three-atom structure resembling a carbon atom in the sp^2 hybridization state. There also needs to be at least one hydrogen bond receptor, but a combination of one donor and one receptor is preferred. The 1*H*-1,2,3-triazole ring seems to be too bulky to serve in this role well – my observations do not point to any interactions between 1*H*-1,2,3-triazole structure and binding pocket. Among the tested isosteres, amidinyl group seems to be the best replacement for guanidinyl.
- Extending the side chain by one atom does not impact binding. Introducing hetero atoms (nitrogen or sulfur) into the linker does not have significant influence on binding. There is also limited difference between positions of the heteroatom

in the linker. Additionally, introducing stiff 1*H*-1,2,3-triazole ring into the linker seems to have limited effect on the binding position as well.

4.4 Impact of docking results on synthetic considerations

The primary criteria for the selection of structures that were taken to synthesis were as follows.

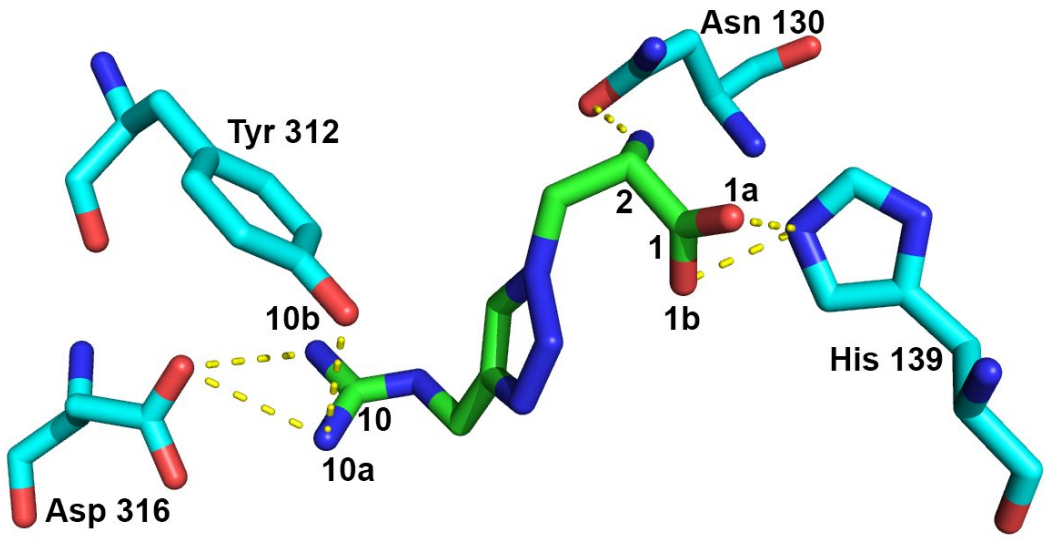
1. Alignment within the binding pocket. Promising candidate needs to align in the binding pocket similarly to Arg. This excludes all structures marked with asterisk(s) in **Table 5**.
2. Novelty of the structure. The aim of the project was to produce novel compounds. This excludes structures **2c**-**2m** as they are natural amino acids.
3. Calculated binding energy **ΔG**. The lower the value of calculated binding energy the higher affinity of the structure towards the receptor. Structures with higher affinity promise better efficacy.

Structures from the group **1** (structures in **Table 2**) are all novel and achieved most promising results (lowest values of the binding energy (**ΔG**, **Table 5**), with binding modes most similar to those of Arg (**Table 1**). However, taking into consideration synthetic steps, compounds **1a-1c** are the most promising. Their predicted binding energy (**ΔG**) is still above the score for Arg and all four can be synthesized from a common substrate – azidoalanine. Compound **2a** has structure similar to Isp and binding mode similar to that of Arg.

For those reasons I decided to carry structures **1a**, **1b**, **1c**, **2a**, to the synthetic step and I would like to discuss their docking modes further.

Compound **1a** (**Table 6**) maintains crucial interaction with the His 139. However, the orientation of the carboxyl prevents it from binding with Asn 130. Comparing with Arg, I observed 90° rotation of the carbonyl plane that also results in neither of its oxygen atoms being favored for interaction with His 139. The guanidinyl group at the end of the side chain is out of plane with Asp 316 but there is a new possible interaction with the tyrosine 312 (Tyr 312).

Table 6 Calculated binding mode of compound **1a**

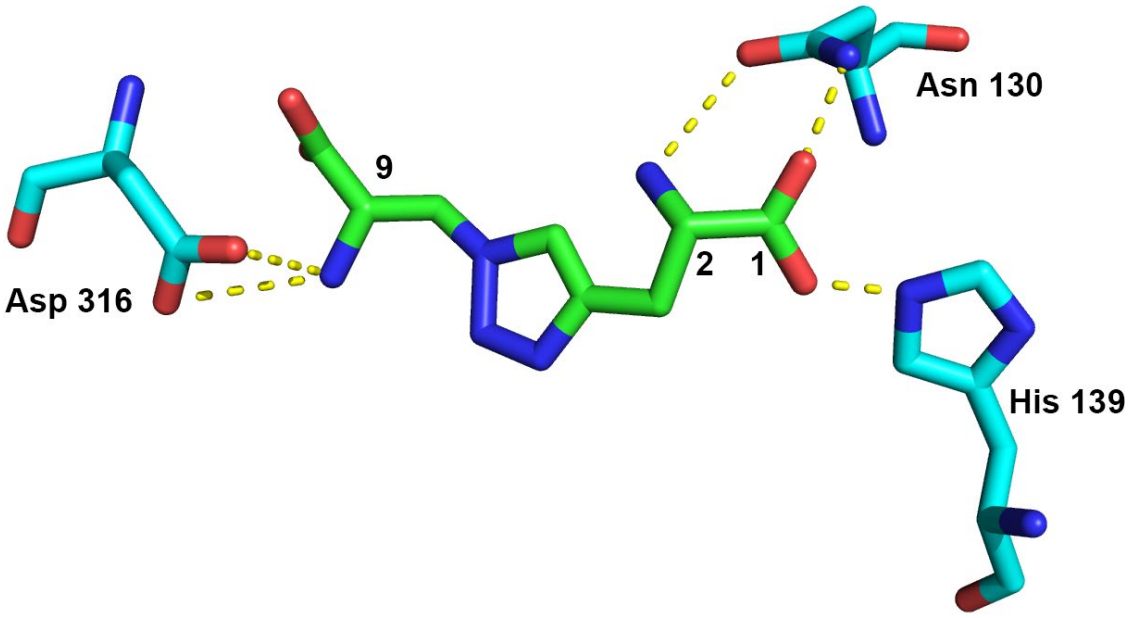


Residue	Atomic group	Distance (Å)[b]
(Asn130)O^{g}	(N2a)H	2.4
(His139)NH	(C1)O1a	2.3
(His139)NH	(C1)O1b	2.3
(Tyr312)O^{4}H	(N11a)H	3.4
(Asp316)O^{g1}H	(N11a)H	3.1
(Asp316)O^{g1}H	(N11b)H	3.2

[a]Colors: C(ligand)=green, C(protein residue=cyan, N=dark blue, O=red, H-bond=yellow.
[b]From the docking experiment.

Compound **1b** (**Table 7**) anchors its amino acid moiety well with both His 139 and Asn 130. The bond lengths to both residues are also shorter compared with Arg (**Table 1**). However, it has a single interaction at the end of the sidechain at the limit of the typical hydrogen bond length (3.3 Å[72])

Table 7 Calculated binding mode of compound **1b**

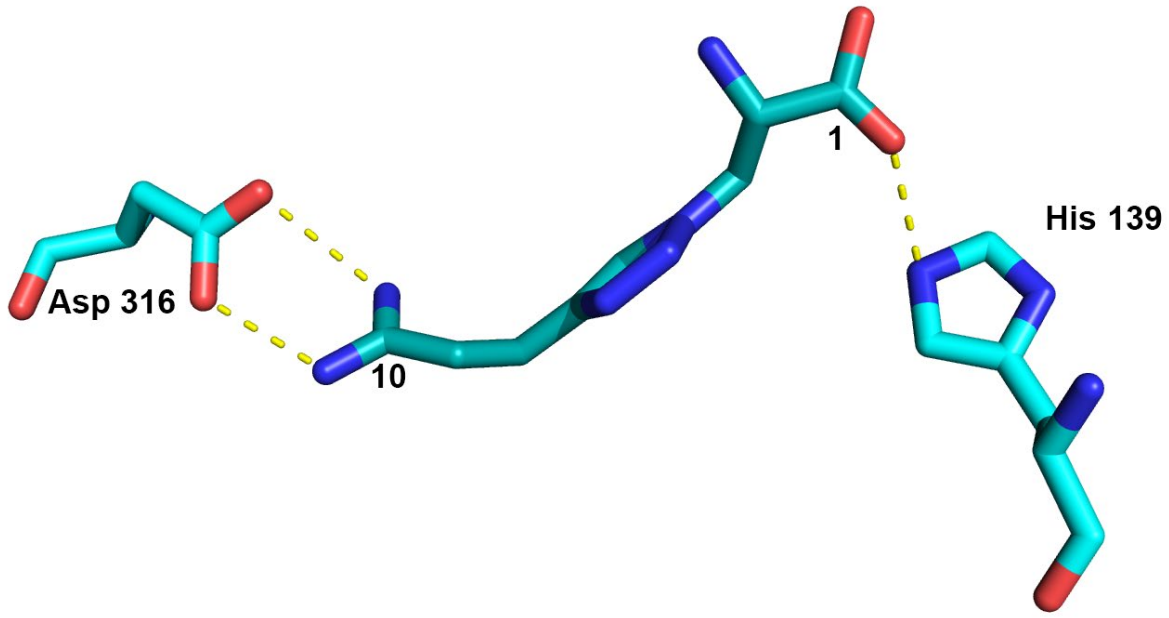


Residue	Atomic group	Distance (Å)[b]
(Asn130)O^{g}	(N2)H	3.2
(Asn130)N^{g}	(C1)O^{1}	2.1
(His139)NH	(C1)O^{2}	2.0
(Asp316)O^{g1}H	(N9)H	3.0
(Asp316)O^{g2}H	(N9)H	3.2

[a]Colors: C(ligand)=green, C(protein residue=cyan, N=dark blue, O=red, H-bond=yellow.
[b]From the docking experiment.

Compound **1c** (**Table 8**) displays similar behavior. It maintains a single interaction between carboxylic acid and His 139 with strong almost planar interaction with Asp 316. Despite limited polar interactions with the binding pocket, the position of the carboxylic acid is similar to that of Arg (**Table 1**).

Table 8 Calculated binding mode of compound **1c**

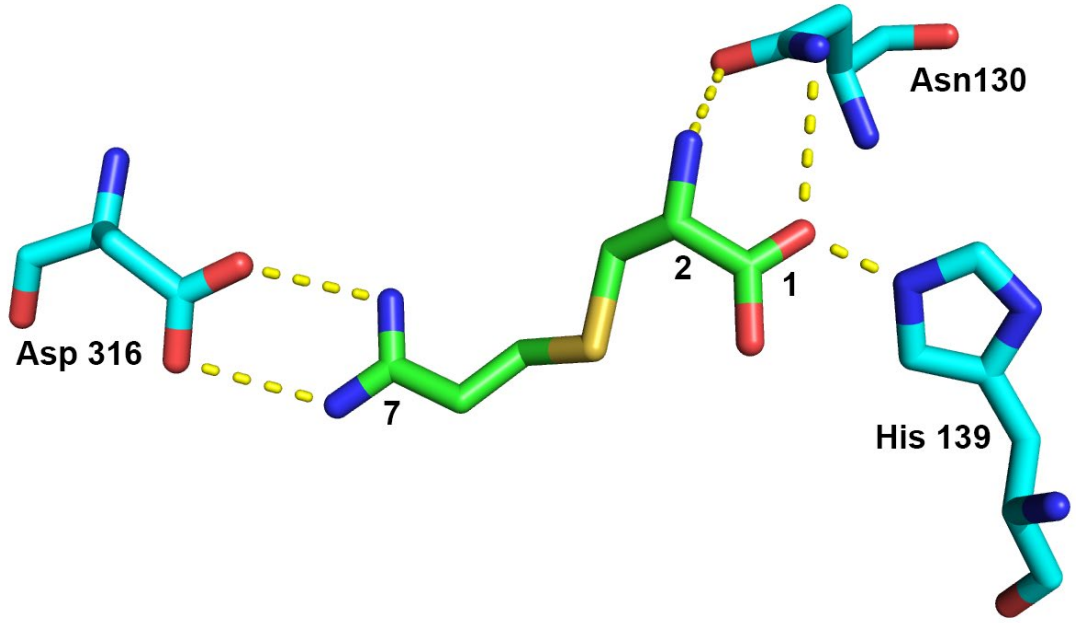


Residue	Atomic group	Distance (Å)[b]
(His139)NH	(C1)O1a	2.6
(Asp316)O^{g1}H	(N10a)H	3.1
(Asp316)O^{g2}H	(N10b)H	2.7

[a]Colors: C(ligand)=green, C(protein residue=cyan, N=dark blue, O=red, H-bond=yellow.
[b]From the docking experiment.

Compound **2a** (**Table 9**) has almost identical binding mode to Arg (**Table 1**). It interacts with the receptor protein in all crucial spots while its calculated affinity is only slightly lower than that of the Arg.

Table 9 Calculated binding mode of compound **2a**



Residue	Atomic group	Distance (Å)[b]
(Asn130)O^{g}	(N2)H	2.7
(Asn130)N^{g}	(C1)O^{1}	3.3
(His139)NH	(C1)O^{1}	2.1
(Asp316)O^{g1}H	(N7a)H	3.1
(Asp316)O^{g2}H	(N7b)H	3.0

[a]Colors: C(ligand)=green, C(protein residue=cyan, N=dark blue, O=red, H-bond=yellow.
[b]From the docking experiment.

4.5 Synthesis

Analyzing possibilities, I decided to firstly synthesize the three most promising compounds from group **1** (*i.e.* compounds **1a-1c**). Those compounds have shown calculated binding energy **ΔG** similar to that of Arg (**Table 5 2c**). A schematic diagram of my synthetic strategy is represented on **Fig 32**. Compounds **1a-1c** might be obtained by employing CuAAC approach (general information on CuAAC is described in chapter **3.3.3**). Azidoalanine (**4**) would be a common substrate and would provide azide moiety in this reaction. Compound **4** can be synthesized from commercially available L-serine (Compound **4** synthesis is explained in detail in chapter **4.5.1**). Alkyne component of the CuAAC reaction (**5-7**) depends on the final compound (**1a-1c** structure). In synthesis of compound **1a** this alkyne component is compound **5**. Compound **5** would be synthesized from 1-aminoprop-2-yn **8** (Compound **5** synthesis is described in chapter **4.5.2)**. In the case of compound **1b**, alkyne component comes from (*S*)-2-(*tert*-butylcarbooxy)aminohex-5-ynoic acid (**6**). Compound **6** is commercially available and was not synthesized in this work. Lastly, the necessary substrate to obtain compound **1c** is pent-4-ynimidamide (**7**). Compound **7** would be obtained from the corresponding cyanide - pent-4-ynenitrile (**9**) on route of Pinner reaction. General information on Pinner reaction is described in chapter **3.3.4**.

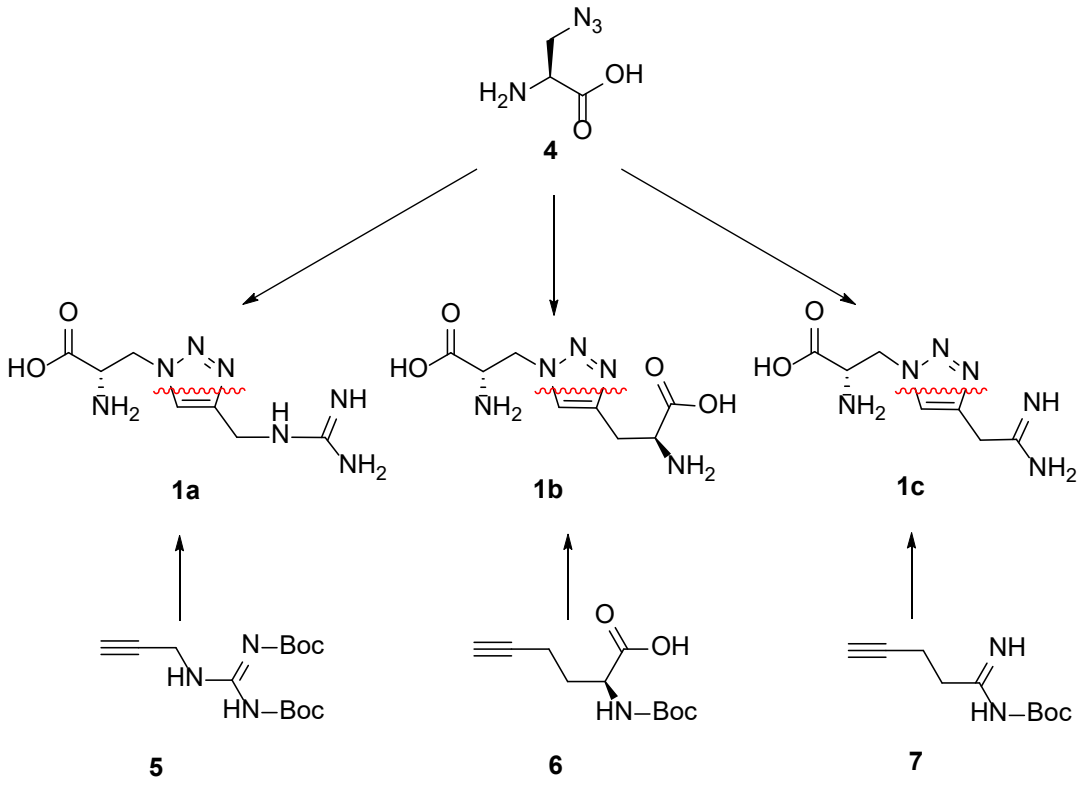


Fig 32 Schematic diagram of planned synthesis of compounds **1a-1c**

Compound **2a** is another organic particle that has been obtained in the scope of this work. It belongs to the group of compounds with aliphatic sidechain **(Fig 33)**. Compound **2a** is a structural analogue of Isp **(Fig 33)** with the sulfur atom replacing carbon atom in the γ- position. This subtle change in the structure might significantly simplify the synthesis from readily available substrates, such as L-cysteine **10** and acrylonitrile **11**. To transform compound **12** into compound **2a,** I will again employ Pinner reaction. Docking results of this compound (**Table 10**) suggests that it aligns itself with Arg in the binding pocket of ArgRS. Additionally, it shows a high calculated affinity towards the receptor – higher than that of Arg (6.2 vs 6.0). Detailed synthesis of **2a** is described in chapter **4.5.4**.

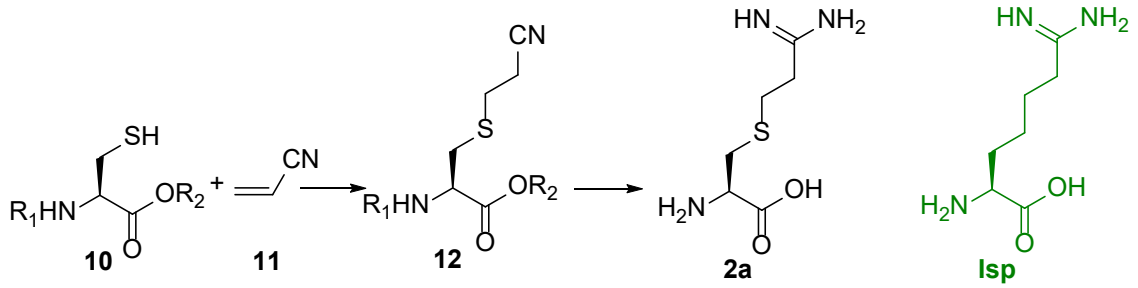


Fig 33 Schematic diagram of planned synthesis of compound **2a**

4.5.1 Synthesis of azidoalanine 17

Literature data point towards the fact that CuAAC tolerates various functional groups[57]. In the early stages of synthesis, I decided to use a protected derivative of azidoalanine **4** to avoid working with highly polar compounds. Application of the Boc group for protection of the NH function, and the OBn group to protect the carboxylic acid function was advantageous due to the mild deprotection conditions (*i.e.* TFA treatment for Boc removal, and catalytic hydrogenation for removal of the benzyl group). In the SciFinder® database I have found two literature methods. First, through replacement of hydroxyl group in compound **13** with the azide moiety in the Mitsunobu reaction conditions (**Fig 34a**)[73]. Second method involves mesylation of a L-serine followed by the substitution with sodium azide (**Fig 34b**)[74].

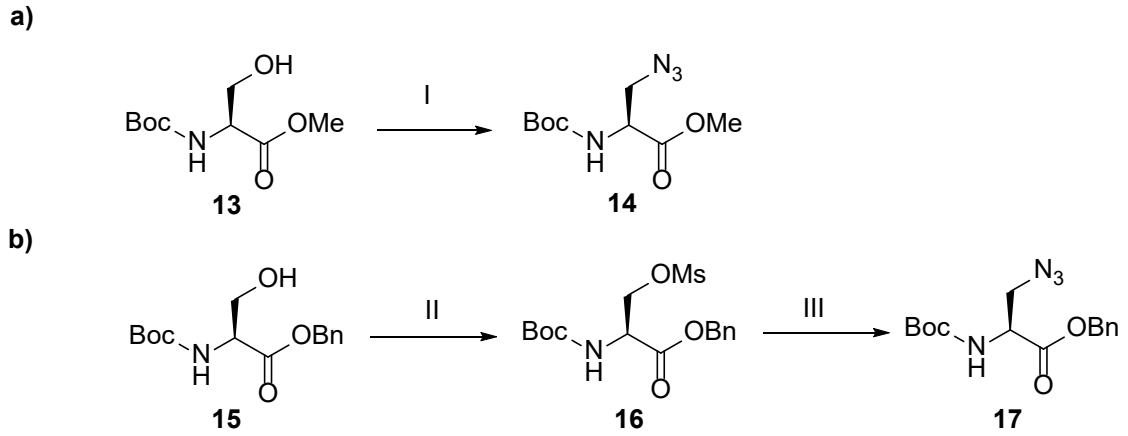


Fig 34 Literature methods of synthesis of protected azidoalanine **4; I –** DEAD, PPh_3, HN_3, THF; **II** – MsCl, Et_3N, DCM; **III** – NaN_3, DMF

I decided to employ the mesylate method for practical reasons, *i.e.* simpler procedure and reactant handling. Two papers report this reaction to undergo readily with satisfactory yields[74,75]. According to those reports this process has three steps: protection of L-serine, reaction of **15** with mesyl chloride, reaction of **16** with sodium azide (**Fig 35**). In my work however, I have met a couple of impediments when trying this approach. First problem was with protecting of L-serine to obtain compound **15**. According to the literature method, the reaction mixture is dissolved in a relatively large volume of DMF. DMF is a solvent with a high boiling point (153°C) and it requires significant heating and time to be evaporated, even at low pressures. Proceeding according to the literature method resulted in decomposition of product **15**. I circumvented this issue by diluting reaction mixture with large quantity of water followed by extraction with ethyl acetate. After evaporating ethyl acetate, I obtained pale yellow oil. That oil was washed multiple times with hexane to yield white fluffy solid – compound **15** in the yield of 74% (procedure 7.2.8). The next step, *i.e.* formation of the mesylate **16** (**Fig 34**) worked without any problems and provided the product **16** in a satisfactory purity and the yield of 82%.

The second obstruction in this pathway was the reaction of compound **16** with sodium azide (**Fig 35**). In my experiments the reaction led to a multicomponent mixture of products. I isolated the major product chromatographically according to the literature method. NMR analysis shows strong signals characteristic for dehydroalanine[76,77] - 5.80 ppm and 6.19 ppm in ^{1}HNMR as well as 105.6 ppm and 131.5 ppm in ^{13}CNMR (**Fig. 36**), leading me to believe that the major product of this reaction was **18** (Appendix **8.2.9**). This analysis allowed me to calculate that dehydroalanine was produced with the yield of 63%. That would suggest that sodium azide

in this reaction plays the role of a base, deprotonating stereogenic carbon and leading to elimination reaction. In both literature works containing azidoalanine synthesis[74,75], the authors do not mention formation of alkene **18**. In the light of my findings, it is obvious why Friscourt et al obtained azidoalanine in a low yield of 18%[75].

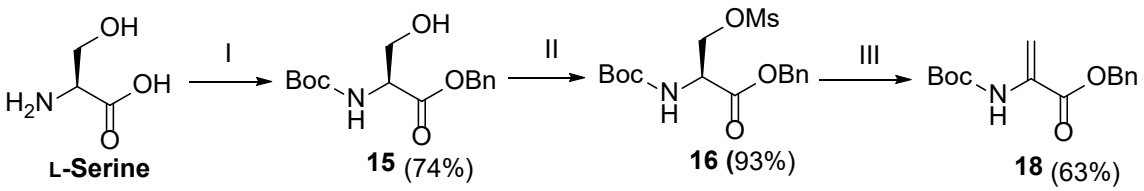


Fig 35 Synthesis of azidoalanine **17** through mesylate intermediate[74] **I** – a) Boc_2O, K_2CO_3, 1,4-dioxane, H_2O b) BnBr, DMF; **II** – MsCl, TEA, DCM; **III** – NaN_3, DMF,

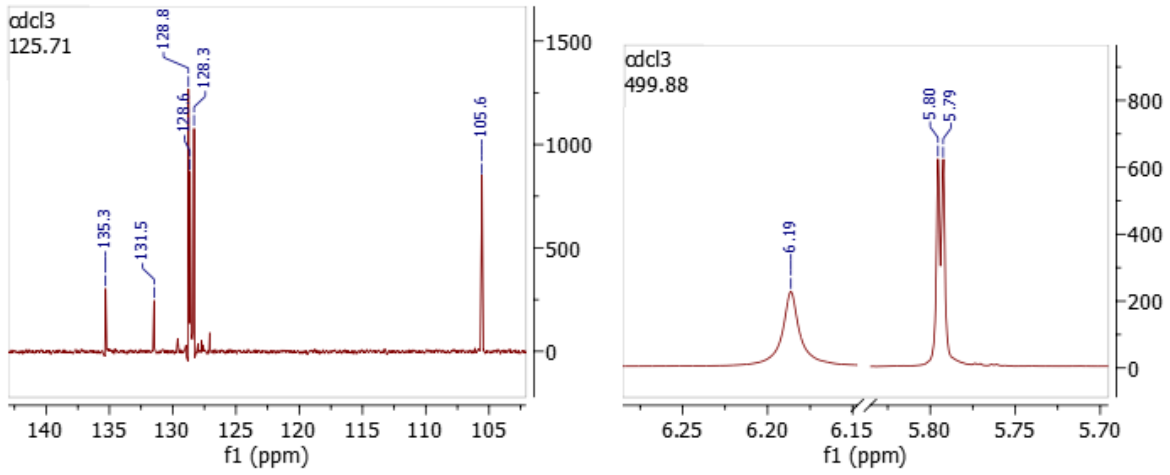


Fig 36 Excerpt from NMR analysis of compound **18** obtained in the reaction of **16** with sodium azide in the literature conditions, showing signals characteristic for **18** ((13C{1H} NMR (125.71 MHz, CDCl3): □C 28.4 (3 x CH3), 67.7 (Ph-CH2-O), 80.8 (C(CH3)3), 105.6 (CH2=C), 128.3 (Ar-C), 128.6 (Ar-C), 128.8 (Ar-C), 131.5 (>CH-NHBoc), 135.3 (Ar-C), 155.9 (t-BuO-CO-NH), 170.9 (-COOBn); 1H NMR (499.87 MHz, CDCl3): □H 1.49 (9H, bs), 5.26 (2H, s), 5.80 (1H, d 2JHH = 1.5 Hz), 6.19 (1H, bs), 7.34 (5H, m))

I have made multiple attempts to modify reaction conditions to obtain desired azidoalanine **17** (**Table 10**). Adjusting the reaction temperature in range from 0 °C to 65°C influenced only substrate conversion rate and had minimal impact on the final mixture composition (**Table 10,** entries **1-4**). Similarly, switching reaction medium from DMF to acetone did not improve the rate of azidoalanine formation (**Table 10,** entry **5**). Reaction also did not occur in acetonitrile showing minimal conversion rate of the substrate **16** after 48 hours (detected by thin layer chromatography, TLC) (**Table 10,** entry **6**). Substrate was recovered with the yield of 82%.

Table 10 Reaction conditions tested to optimize formation of azidoalanine **17** from compound **16**

	Temperature	Solvent	Result
1.	0 °C	DMF	**No product yielded**
2.	RT	DMF	**No product yielded**
3.	45 °C	DMF	**No product yielded**
4.	65°C	DMF	**No product yielded**
5.	RT	Acetone	**No product yielded**
6.	RT	Acetonitrile	**No conversion**

Given the difficulties in obtaining compound **17** with satisfactory yield and purity, I decided to modify my approach and employ cyclic sulfamidate strategy (**Fig 37**). This method has been previously used in the literature to obtain azido and diamine derivatives of amino acids[78–80]. Cyclic sulfamidates (chapter **3.3.2**) are well researched class of compounds. They are known to react readily with various nucleophiles including sodium azide. According to the literature procedure, preparation of the cyclic sulfamidate was a two-step process. First, protected serine **15** reacted with thionyl chloride forming sulfamidite with an oxathiazolidine ring structure **19**. This fast-paced reaction undergoes at low temperature, in dry ice/acetone bath. Compound **19** was stable in short time window and was later oxidized with metaperiodate to form stable cyclic sulfamidate **20**. During my work, I observed that letting reaction mixture reach room temperature slowly before quenching with water leads to 10% better yields. Compound **20** is a handy precursor to various α-mono- or α,α-disubstituted amino acids. It undergoes nucleophilic substitution readily with various nucleophiles[80]. It is worth noting that no step in this process carries a risk of losing chiral purity. For the purpose of this work, the obtained compound **20** was later reacted with sodium azide forming desired azidoalanine **17**. Reaction occurs at room temperature with a mixture of acetone and water as a solvent. Azidoalanine **17** is produced in the yield of 96%. The ^{1}H and ^{13}C NMR spectra were consistent with those available in the literature[75].

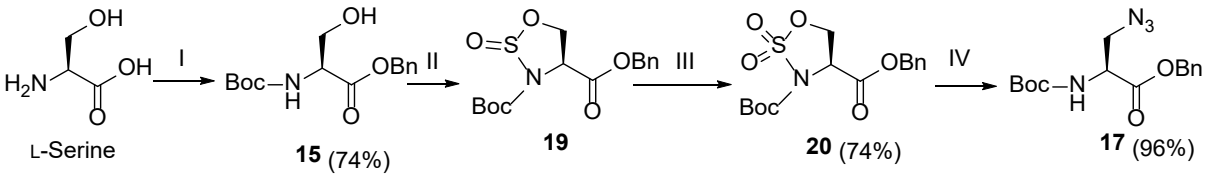


Fig 37 Schematic diagram of sulfamidate strategy for preparation of azidoalanine **17; I –** a) Boc_2O, K_2CO_3, 1,4-dioxane, H_2O b) BnBr, DMF; **II –** SO_2Cl, pyridine, DCM; **III –** $RuCl_{(cat)}$, $NaIO_4$, AcCN, H_2O; **IV –** NaN_3, Acetone, H_2O

4.5.2 Transformation of azidoalanine 17 into compound 1a

Compound **1a** was obtained *via* CuAAC reaction from azidoalanine **17** and compound **5** (**Fig 38**). Compound **5** was obtained according to literature method from *S*-methylisothiourea sulfate **21** [81]. First, compound **21** was -protected in typical conditions (Lewis base, H_2O/1,4-dioxane, overnight). Protected thiourea **22** was then stirred overnight with 1-amineprop-2-yn **8** in THF to form compound **5**.

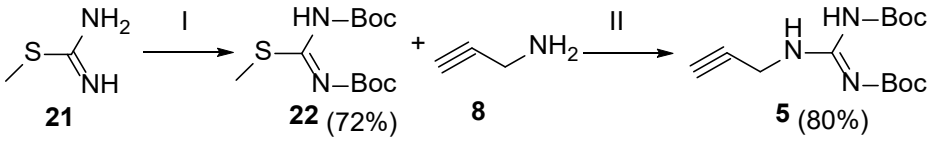


Fig 38 Schematic diagram of synthetic strategy for compound **5**; **I –** Boc_2O, $NaOH_{(aq)}$, 1,4-dioxane; **II –** THF

Azidoalanine **17** and **5** undergo CuAAC in a mixture of THF and water as a reaction medium yielding intermediate compound **24 (Fig 39)**. I decided to start the deprotection process with removal of the benzyl group *via* hydrogenation on palladium catalyst. After removing catalyst, debenzylated intermediate **25** was directly dissolved in 1:1 mixture of TFA/DCM and stirred for one hour to remove Boc protecting groups. Solvents were removed by evaporation under the stream of dry argon gas, and the residue was suspended in cold diethyl ether. The resulting precipitate was filtered out and dried on the air to yield **1a** with high yield (99%) and purity (NMR).

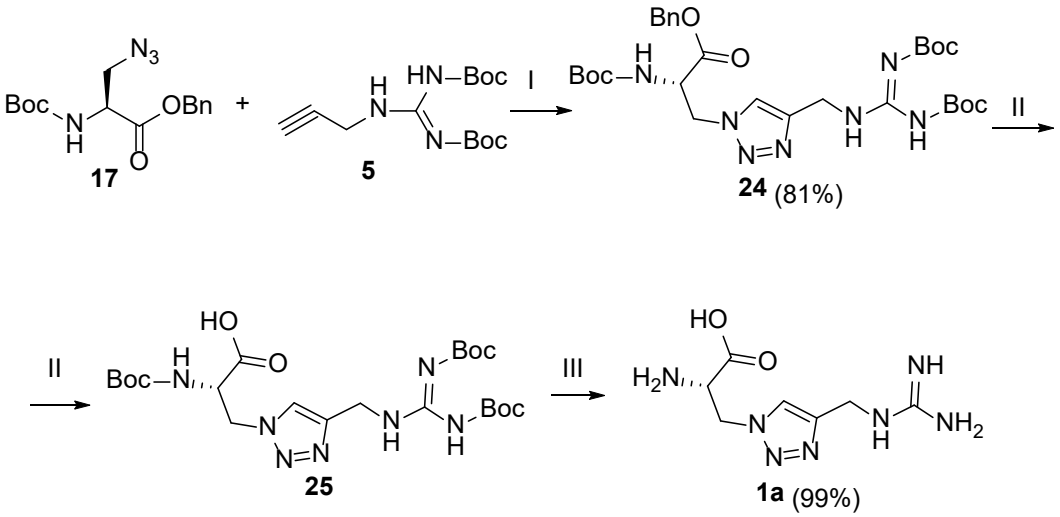


Fig 39 Transformation of azidoalanine **17** into compound **1a**; I – $CuSO_4$, sodium ascorbate, THF, H_2O; II – Pd/C, MeOH; III – TFA, DCM

4.5.3 Transformation of azidoalanine 17 into compound 1b

Compound **1b** was prepared *via* CuAAC reaction from azidoalanine **17** and commercially available 2-((*tert*-butoxycarbonyl)amino)pent-4-ynoic acid **6** (**Fig 40**). In contrast to the previously described synthesis of compound **1a**, this CuAAC reaction did not undergo in THF/water mixture as a reaction medium. For unknown reasons a product of cycloaddition was not detected in the THF/water reaction mixture. A number of different solvent systems have been tested: DMF/water, acetonitrile/water, *tert*-butanol/water and water. Only changing reaction medium to *tert*-butanol/water produced the intermediate **26** in satisfactory yields (71%-76%). Worth noting is the fact that compound **26** chelated copper ions strongly and remained in the **26**-Cu(OAc_2) complex even after flash column chromatography. This behavior is similar to that of similar structure described in patent[82]. The presence of a significant amount of copper impurity in the sample led to a number of problems. Primarily, the molecular mass of the complex was unknown, making calculating reaction efficiency impossible. Additionally, copper impurity in the sample impacts negatively the sensitivity of the NMR probe leading to broad peaks in ^{1}H NMR and extremely weak signals in ^{13}C NMR (**Fig 41**). For that reason, additional purification with ion exchange Chelex® resin was necessary. Purified compound **26** was later deprotected in a similar fashion to compound **23** to yield the target compound **1b**.

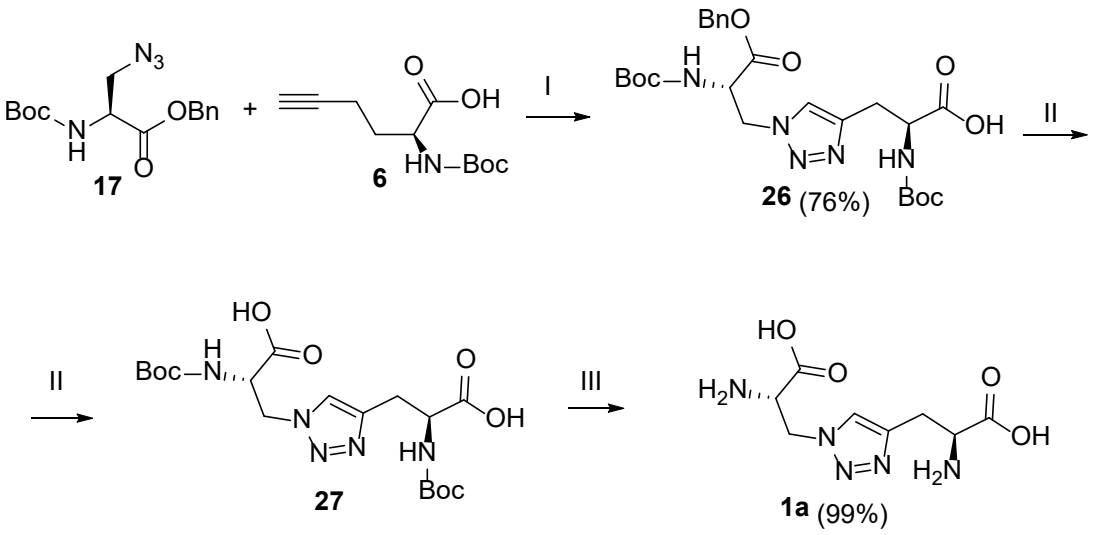


Fig 40 Transformation of azidoalanine **17** into compound **1b**; I – $Cu(AcO)_2$, sodium ascorbate, tBuOH, H_2O; II – Pd/C, MeOH; III – TFA, DCM

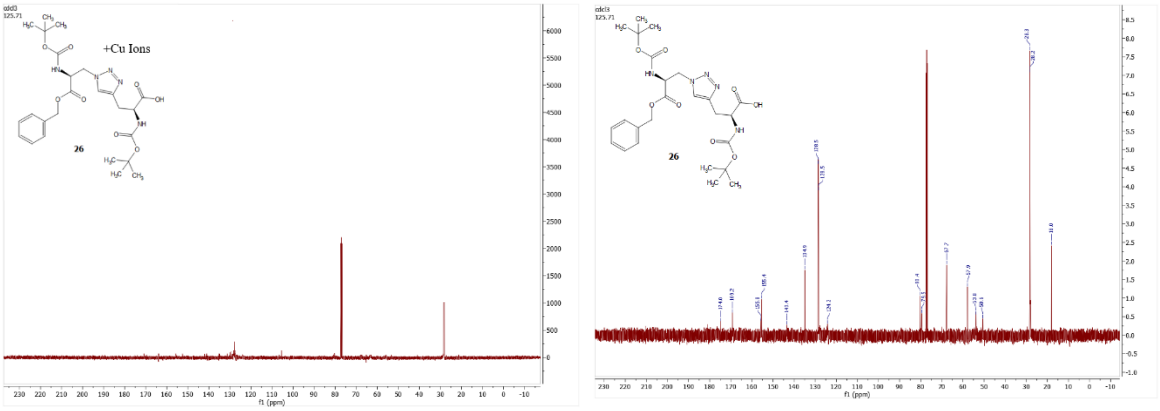


Fig 41. Comparison of ^{13}C NMR spectra of compound **25** in presence of copper ion(left) and after treatment with Chelex® resin(right)(Appendix **8.2.11**)

4.5.4 Transformation of azidoalanine 17 into compound 1c

Compound **1c** was prepared *via* CuAAC reaction from azidoalanine **17** and compound **7**. (**Fig 42a**). Before that, compound **7** was prepared from 4-pentynenitrile **27** employing the Pinner reaction. In the Pinner reaction, nitrile **9** reacts under anhydrous conditions with methanol and hydrochloride to form highly reactive iminoester **28**. Compound **28** reacted readily with water to form carboxylic acid. For that reason, every step in this synthesis was confined in inert gas atmosphere and with dry solvents. After stirring overnight at 0 °C, reaction mixture was blown dry with a stream of dry argon gas. Obtained residue was dissolved in small

amount of dry methanol and dry ammonia saturated methanol is added dropwise. Reaction was left stirring at 0 °C for 1 to 6 days with minimal impact on yields and purity. Produced compound **7** was treated as two previous CuAAC substrates and was reacted with azidoalanine **17** in similar conditions as mentioned in chapter **4.5.3,** *i.e.* in a mixture of THF and water as a reaction medium yielding intermediate compound **24** (**Fig 40b**)

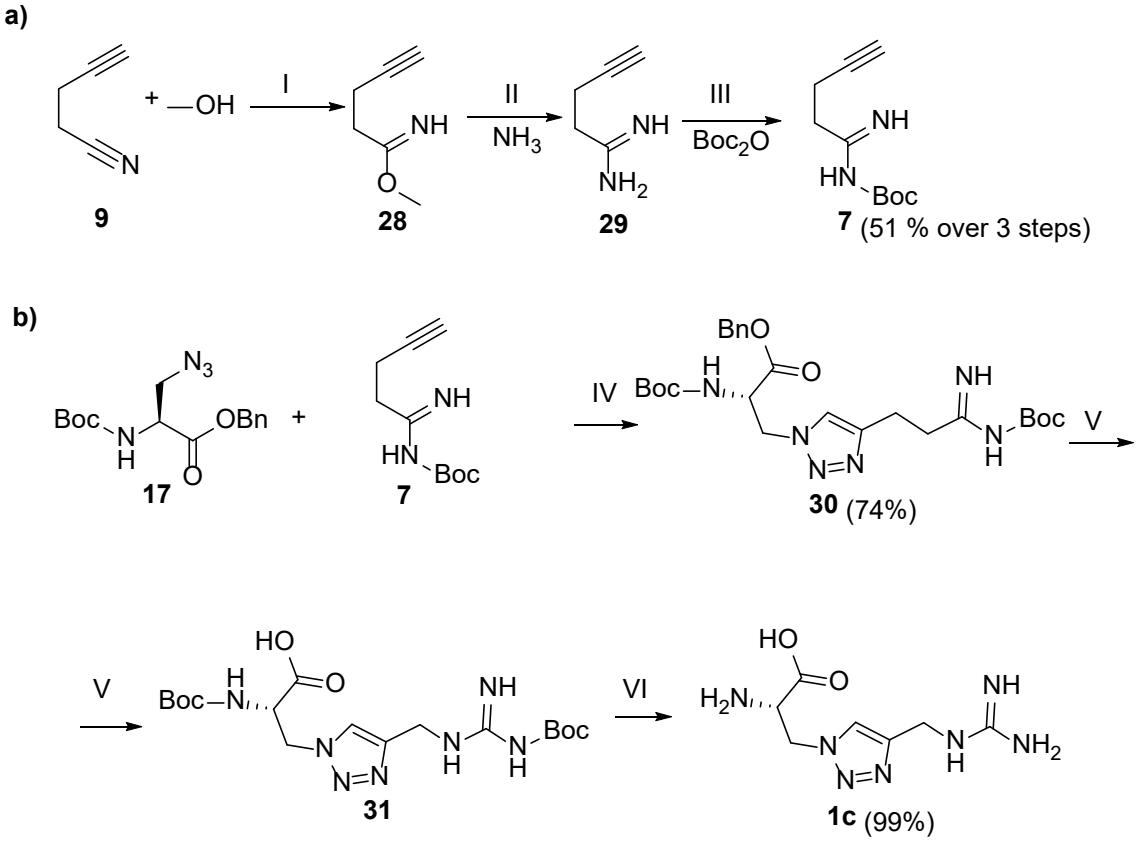


Fig 42 Schematic diagram of synthetic strategy for compound **1c. a)** synthesis of compound **7** ; I – AcCl, $MeOH_{(dry)}$, Ar; II –$MeOH_{(ammonia)}$; III – Boc_2O, 1,4-dioxane, $NaOH_{(aq)}$; **b)** transformation of **17** into **1c;** IV – $CuSO_4$, sodium ascorbate, THF, H_2O; V – Pd/C, MeOH;VI – TFA, DCM

4.5.5 Transformation of cystine 32 into compound 2a

Compound **2a** required quite a different approach (**Fig 43**). To form an amidine group at the sidechain, I also employed the Pinner reaction. However aliphatic sidechain backbone required classic retrosynthetic approach, in contrast to "building block" approach of click chemistry. Commercially available cystine ester **32** was the starting substrate. Reduction of cystine with triphenylphosphine is a reliable way to yield cysteine with satisfying efficacy[83]. Protected cysteine **33** was reacted readily with acrylonitrile in the presence of TEA to produce *S*-(2-cyanoethyl)cysteinate **34.** This substitution was also a straightforward reaction and did not

require any specific measures. Because compound **34** was soluble in methanol it was a perfect substrate for the Pinner reaction. The next step was significantly slower than for compound **29**. Formation of iminoester in this case required at least 6 days to achieve satisfactory efficacy. The reaction mixture was again evaporated to dryness under flow of dry argon gas to remove residual hydrochloride. The residue was dissolved in methanol and then ammonia saturated-methanol was added dropwise. Reaction achieved complete conversion to product **35** after stirring overnight at 0 °C. Deprotection of intermediate **36** was simple in this process. Stirring overnight in 1:1 solution of DCM/TFA removed both Boc and methyl ester to afford the target compound **2a**.

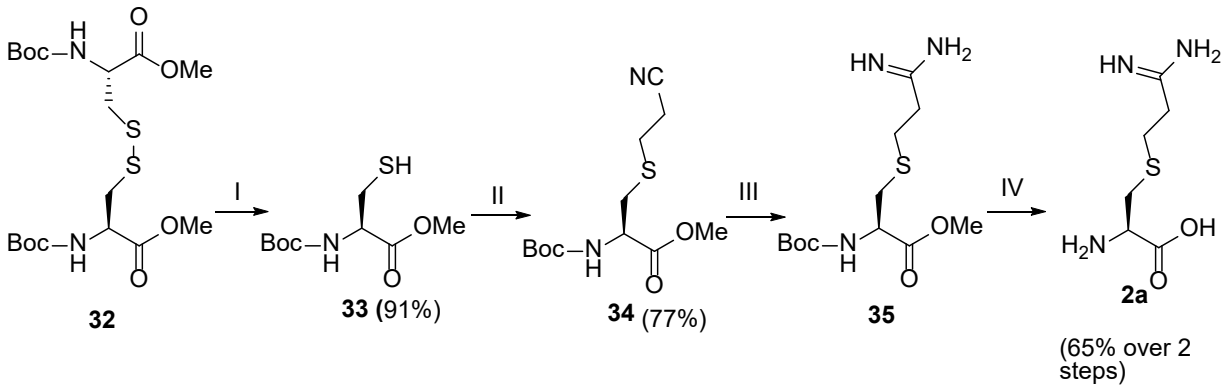


Fig 43 Schematic diagram of synthetic strategy for compound **2a** ; I – PPh_3, MeOH, NaOAc,; II –11, TEA, MeOH, H_2O; III – a) AcCl, $MeOH_{(dry)}$, Ar b) $MeOH_{(ammonia)}$; IV – a) Pd/C, MeOH b) TFA, DCM

4.6 Biological efficacy

The scope of this work is not only the design and synthesis of novel Arg analogues, but also to assess the efficacy of these compounds *in vitro*. For that reason, I designed a three-step procedure based on the work of Karatsai *et al*.[37]. The procedure has been designed to limit employed resources to a minimum while providing comparable results for the most promising Arg analogue.

The first test in this procedure was the MTS assay(**Fig 44**). I started the examination from U251MG cell line. MTS assay allows to measure comparative efficacy of the compound on cancer cell viability. The results of this assay enable the selection of the most promising candidate compound. This most efficacious compound is then selected for further tests to assess its influence on cancer cell viability, morphology and proteome.

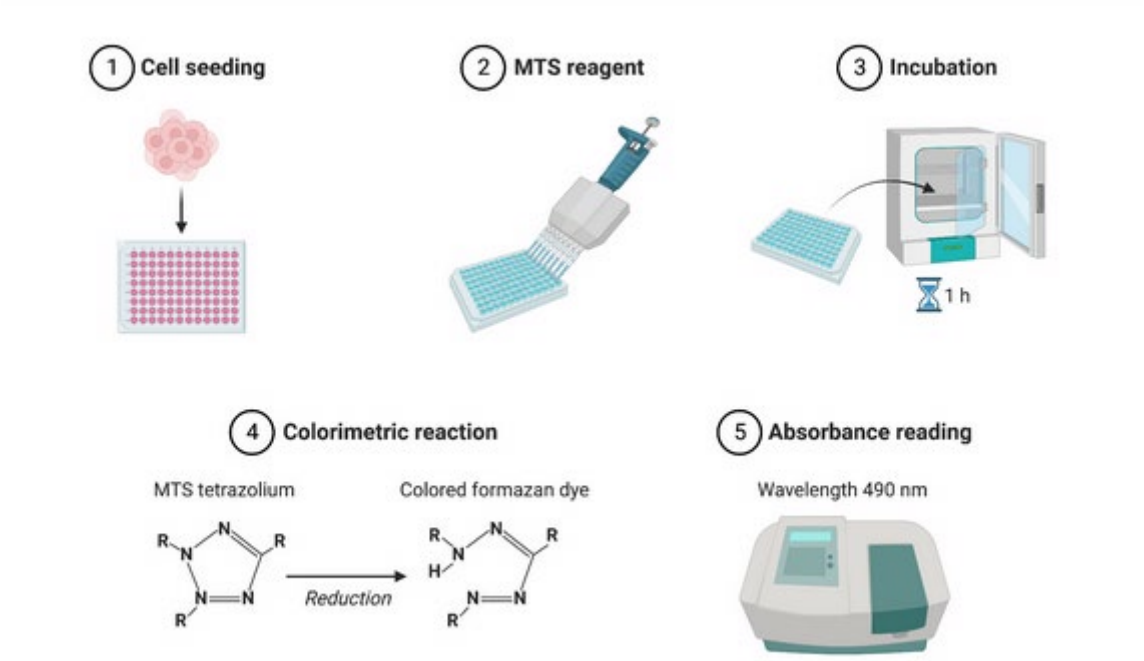


Fig 44 Schematic diagram of MTS assay process[84]

Next, I performed the MTS assay on U87MG cancer cell line. Conducting these tests on two cell lines recognized in the literature as *in vitro* model of glial cancer allows me to assess the results of this work in the light of efficacy of other putative therapies.

Third part was the assessment of cell morphology under the light microscope. The most visible effect of Arg analogues on the microscopic level is on the cytoskeleton formation and

maintenance. The impact of a given compound will be visible as a change of cell shape/morphology.

Final assessment was a semi-quantitative analysis of proteins involved in pro-deaf pathways. This was achieved with a Western blot assay by probing the levels of the following marker proteins: cleaved poly-ADP-ribose polymerase (c-PARP), procaspase 3 (P3), caspase 3 (C3), procaspase 9 (P9), and caspase 9 (C9)

4.6.1 Cell viability assessment

MTS assays are the cornerstone of the efficacy testing in this work. This assessment became the main evaluation tool to compare the efficacy of the tested compounds. If the tested compounds have an impact on Arg dependent metabolic pathways, there should be no statistically significant influence of the compound concentration on cell viability in cultures grown with Arg source (Complete medium, CM). Analogously, in cultures incubated without Arg (Arg free medium, AFM) there should be a significant decrease in cell viability in relation to the exponential increase of the compound concentration.

Figure 45 presents graphs created from combined results of MTS assays conducted on U251MG cell line seeded in complete medium and then transferred to the CM and AFM with concentrations of the compounds from 0 μM to 500 μM, respectively. The results for all four compounds (**1a**, **1b**, **1c**, **2a**) are the average of three biological repetitions run in technical triplicates.

For the four compounds I synthesized, there is no noticeable impact of the compound concentration on cell viability of the cells incubated in the CM up to 72 hours (**Fig 45** a, c, e, g). This means that neither of the compounds is inherently cytotoxic towards U251MG cells when there is external source of Arg available. Interestingly, compounds **1a**-**1c** also exhibit no significant influence on cell viability when no Arg is present in the incubation medium (**Fig 45** b, d and f) suggesting that they are not biologically active in this cell line in the AFM conditions.

However, in the case of compound **2a**, already after 24 hours of incubation there is an observable decrease of cell viability related to the concentration of the compound (**Fig 45** h). The relation between concentration and efficacy of the compound is the handbook inhibitory dose-response curve for the results after 24, 48 and 72 hours. A curve modelled through nonlinear regression model for the data after 72 hours of incubation points towards IC_{50} of 182.2 μM.

Thus, analysis of the results indicates that compounds **1a-1c** are inert towards the U251MG cell line and there is high probability that they are non-cytotoxic to other cell lines in the relevant concentration range. The remarkable difference between CM results and AFM results of compound **2a** point not only to it being cytotoxic towards tested cell line but also to the fact that its efficacy is the result of its interaction with Arg dependent metabolic pathways. For that reason, I decided to further investigate only compound **2a** and its efficacy on the other glial cancer cell line.

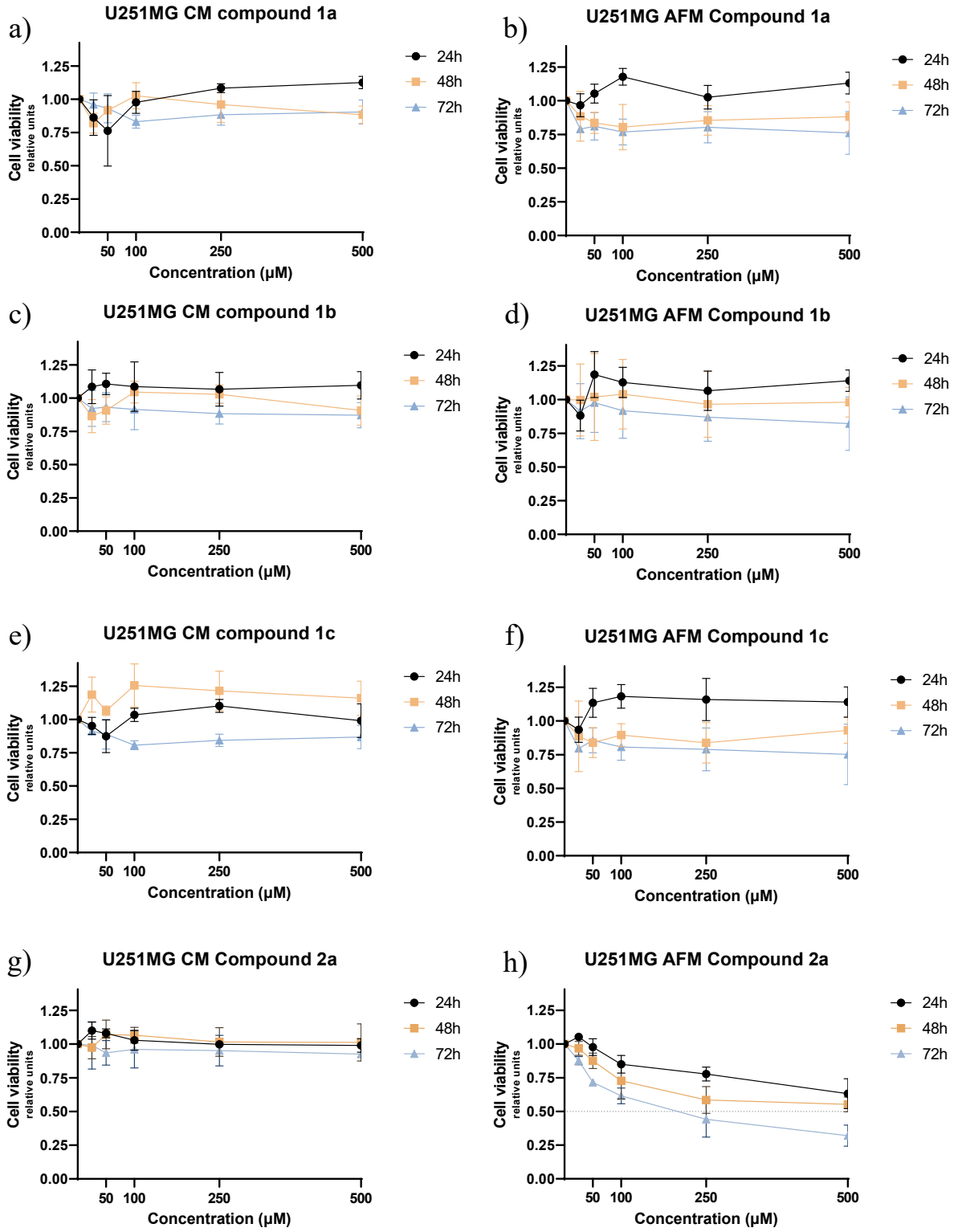


Fig 45 Results of MTS assessment of compounds 1a (a, b), 1b (c, d), 1c (e, f), 2a (g, h) on U251MG cell line.

To confirm the above findings and improve the comparability of my results to the literature, I decided to perform the MTS assessment of compound **2a** also on U87MG cell line. **Fig 46** contains graphs created from combined results of MTS assays conducted on U87MG cell line seeded in complete medium and transferred to the CM and AFM with concentrations of the compounds from 0 µM to 500 µM, respectively. The results are the average of three biological repetitions run in technical triplicates.

Analogously to the U251MG cell line, there is no observable impact on the cell viability up to 72 hours of incubation in CM with compound concentrations up to 500 µM. Additionally, there is a strong observable dose-dependent response already after 24 hours of incubation in AFM. The effect of the compound **2a** is highest after 72 hours of incubation in AFM conditions. A curve modelled through nonlinear regression model for the data after 72 hours of incubation points towards IC_{50} of 11.5 µM.

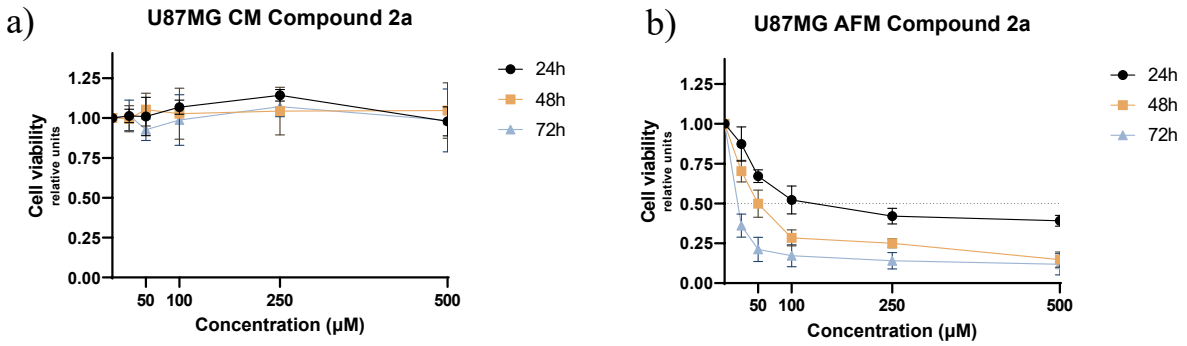


Fig 46 Results of MTS assessment of compound 2a on cell line U87MG

Findings in MTS assessment of both U251MG and U87MG cell line point towards micromolar potency of compound **2a** against glial cancer. Compounds **1a-1c** proved to be inert and were not further evaluated. Additionally, compound **2a** exhibited efficacy only in experiments where Arg free medium was applied. That suggests that the mechanism behind the observed compounds **2a** cytotoxicity is its ability to mimic Arg **in** the metabolic pathways as it was earlier shown for Cav[37].

4.6.2 Morphology assessment

Next, I decided to evaluate the impact of compound **2a** on cell morphology to further investigate whether the cause of the cytotoxicity is the fact that compound **2a** mimics Arg and disrupts Arg-dependent metabolic pathways. One of those pathways is the formation and upkeep of the

cytoskeleton. It has been shown previously that Arg starvation itself affects cell shape and motility, and addition of Cav exacerbates this effect [36,37]. In layman terms, abnormally formed cytoskeleton results in cell losing their tonality and becoming more concave. Those changes to morphology should be visible under normal phase-contrast microscopy.

Effects of Arg deprivation on the U87MG cell line are visible already after 24 hours of incubation already. Naturally spindle-shaped cells (**Fig 47a**) lose their volume when Arg is deprived (**Fig 47b**) and become elongated with visible bulge of nucleus in majority of visible cells. 24-hour treatment in combination with 25 μM concentration of compound **2a** (**Fig 47c**) makes all the cells in the field of view lose their proper morphology. Additionally, bleb-like structures are also visible suggesting that apoptotic processes undergo already. Interestingly, a slight loss of volume can be noticed also when cells are treated with compound **2a** in the presence of Arg. That might suggest that a tested compound competes with Arg with some success.

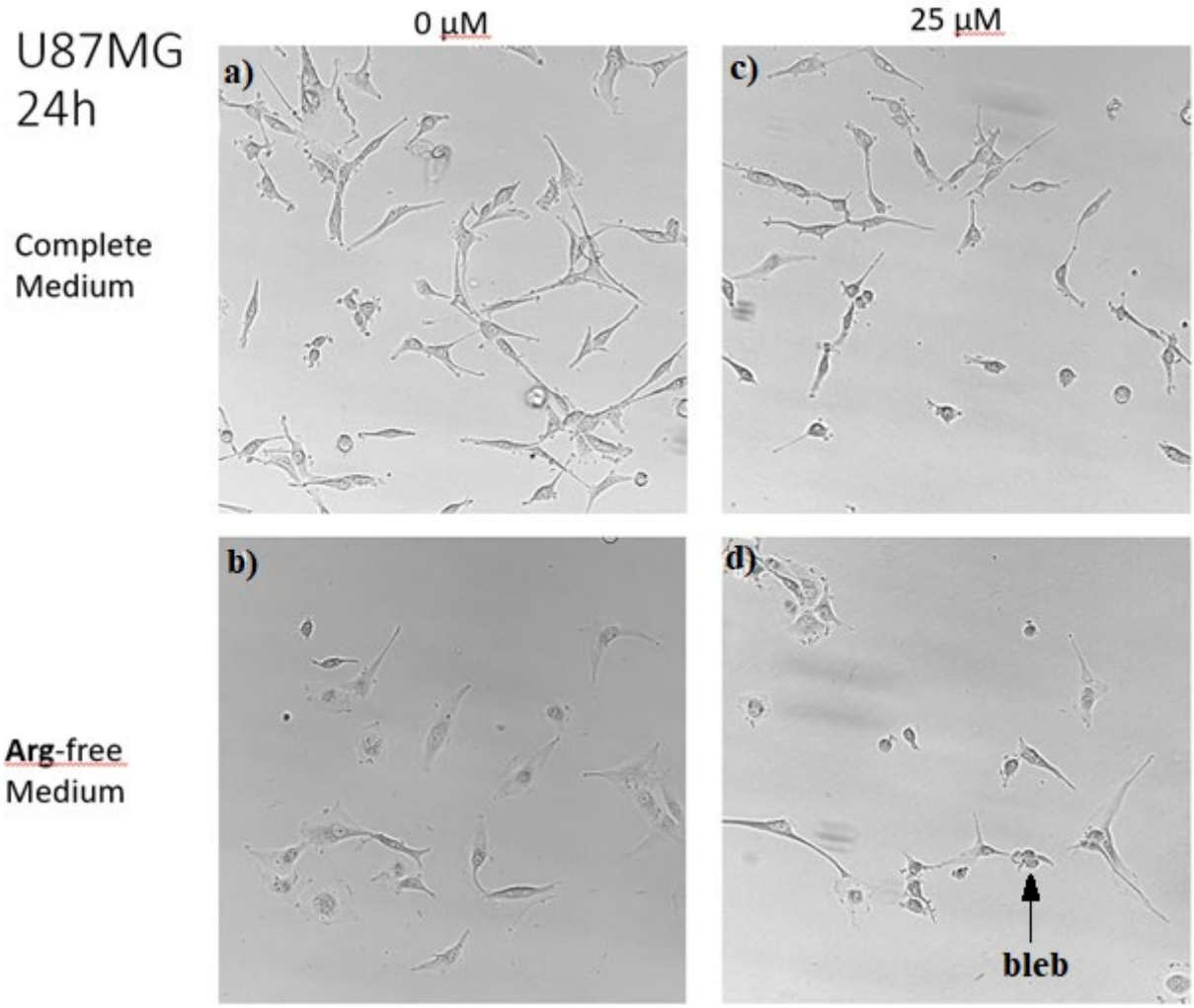


Fig 47 Microscopic assessment of U87MG cell line morphology after 24-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d), concentration of compound **2a,** respectively. Magnification x100

After a 48-hour incubation period in AFM, the above described changes in morphology are exacerbated. Almost all Arg starved cells lose their shape (**Fig 48b**). The number of cells is distinctly lower when incubated in the AFM/compound **2a** medium (**Fig 48d**).

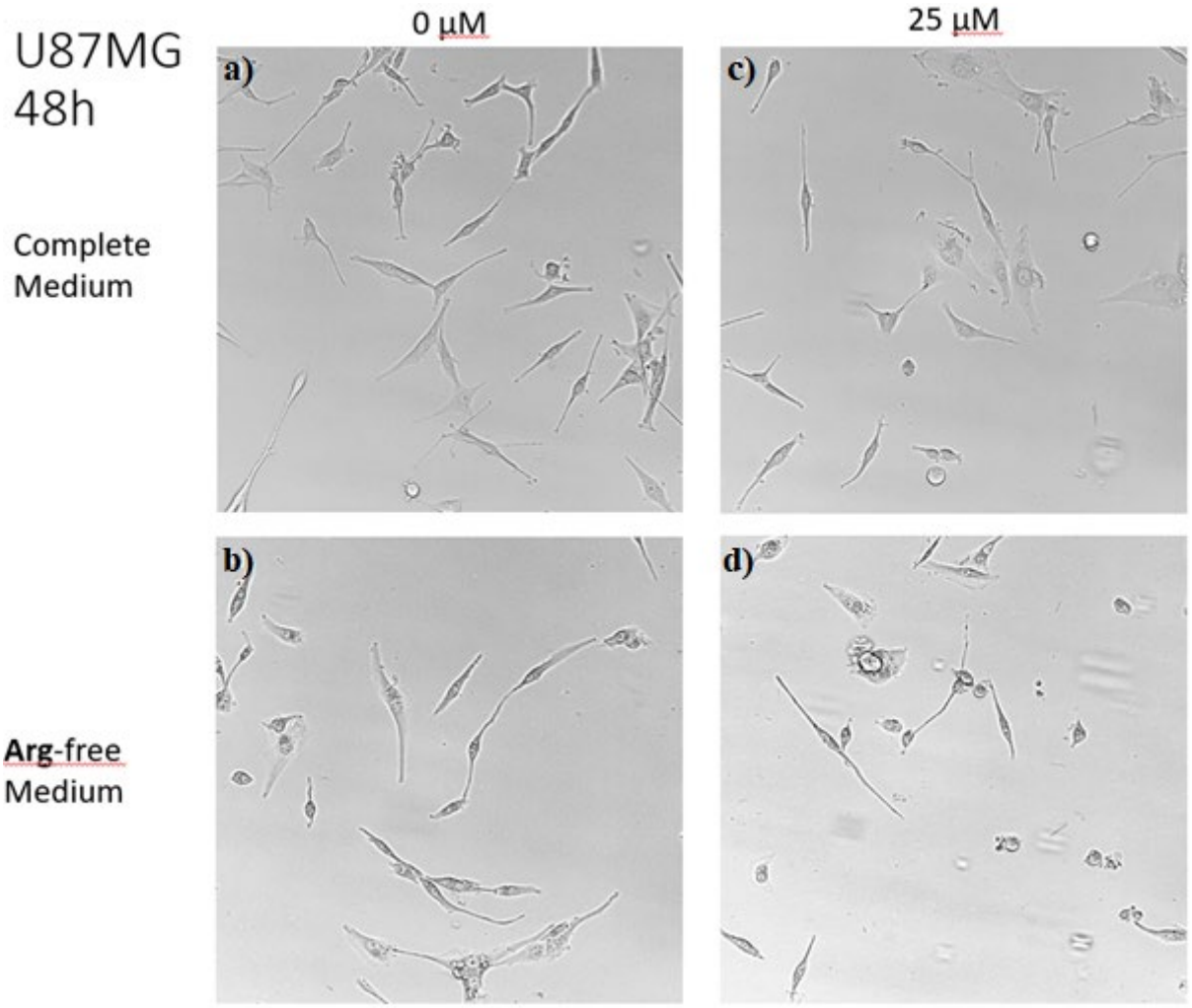


Fig 48 Microscopic assessment of U87MG cell line morphology after 48-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d), concentration of compound **2a,** respectively. Magnification x100

The 72-hour treatment revealed that cells treated with **2a** in the complete medium seem to divide slower (**Fig 49c**) than control group (**Fig 49a**) as their number was substantially lower. Analysis of the images of the cells incubated in AFM for 24 and 72 hours indicates that the cell number seems to be retained despite the changes in their morphology (**Fig 49b**). No "healthy" cells are visible in the field of view in the case of combination therapy (**Fig 49d**).

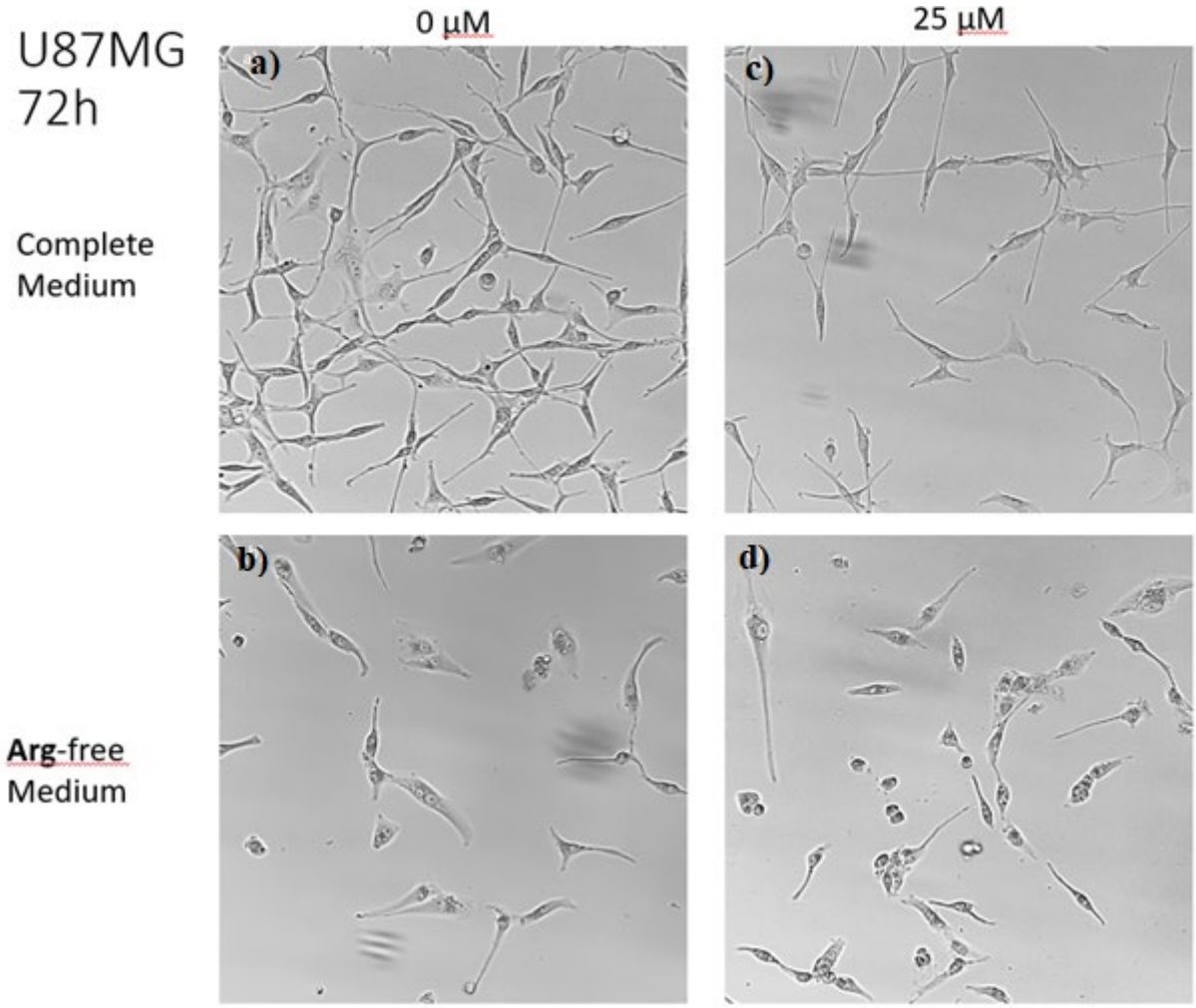


Fig 49 Microscopic assessment of U87MG cell line morphology after 72-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d) concentration of compound **2a** respectively. Magnification x100

U251MG cell line seems to be more resistant to effects of Arg deprivation. After 24 hours, virtually no change in morphology can be distinguished, both images display proper astrocyte-like morphology with characteristic lamellipodia visible (**Fig 50a** and **b**). Also compound **2a** seems to show no impact in CM (**Fig 50c**) with a barely noticeable loss of shape in the AFM/compound **2a** medium (**Fig 50d**).

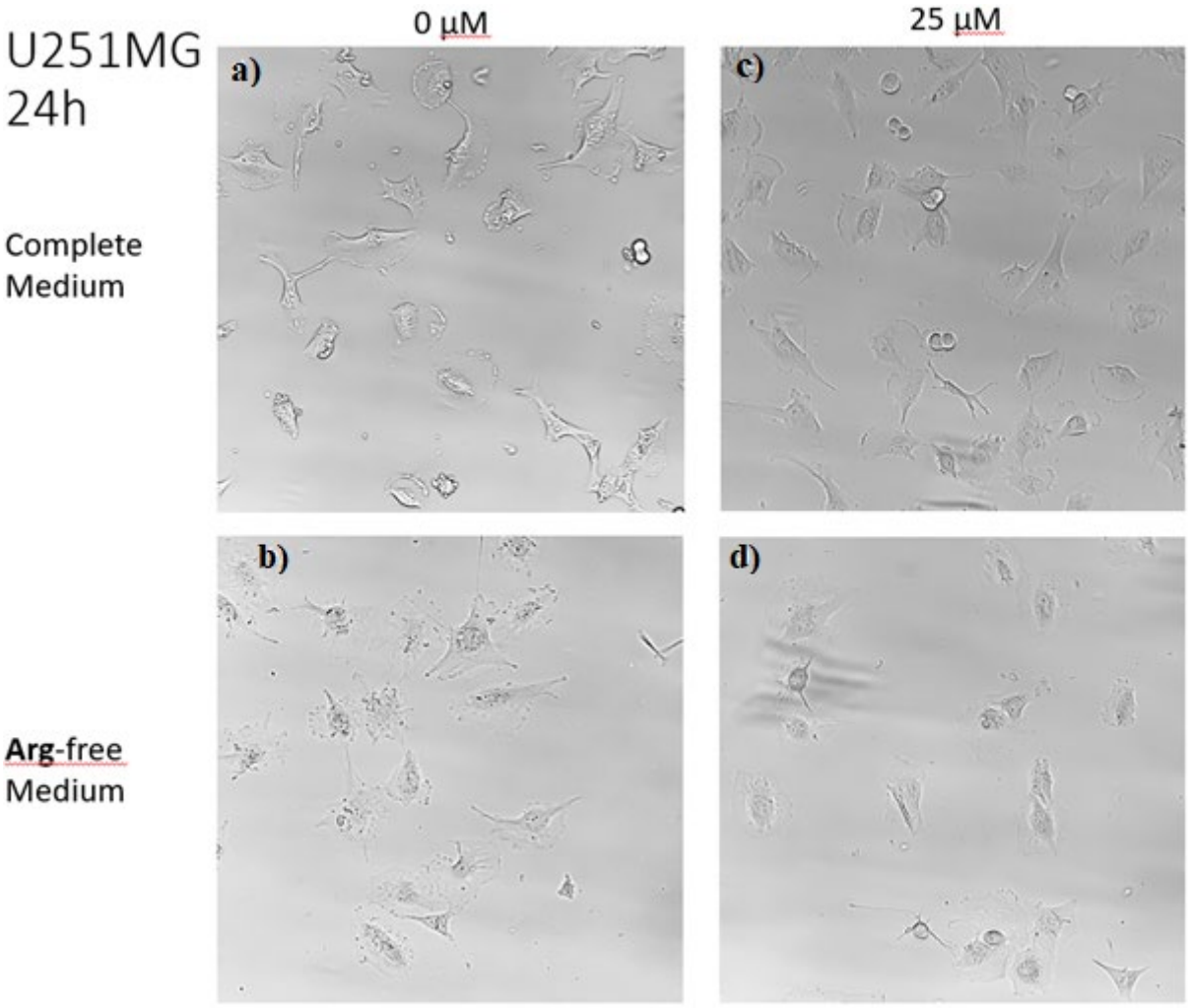


Fig 50 Microscopic assessment of U251MG cell line morphology after 24-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d) concentration of compound **2a** respectively. Magnification x100

After a 48-hour incubation the impact of Arg starvation becomes noticeable. A change in morphology is visible through the lack of distinguishable lammelipodia, additionally cells become more compact as they lose significant part of their volume (**Fig 51b**). Loss of original shape is definite in the combination treatment (**Fig 51d**). Cells become mostly spherical, visibly standing out of the image plane. Again, no distinguishable change is visible in combination of CM/**2a**.

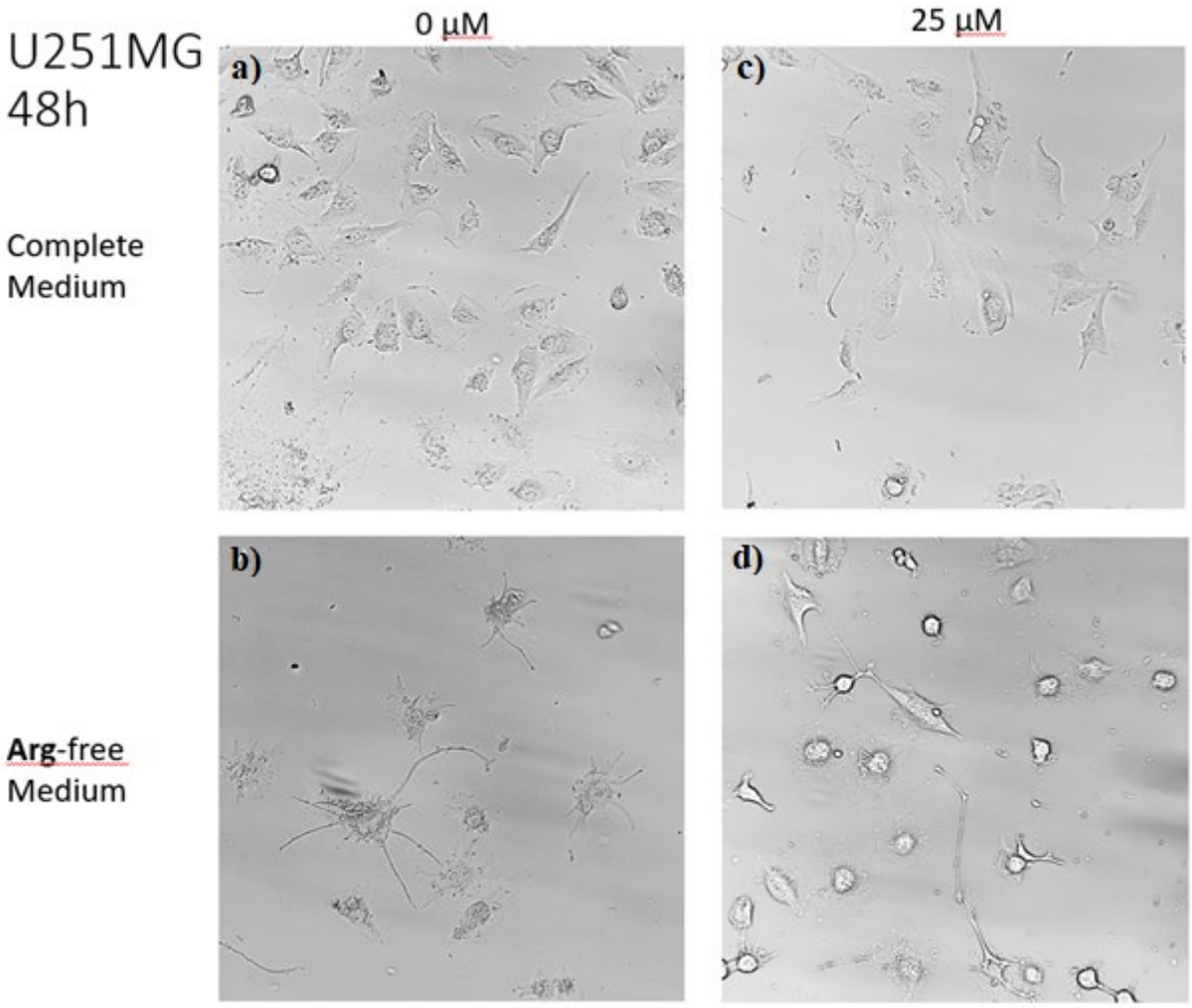


Fig 51 Microscopic assessment of U251MG cell line morphology after 48-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d), concentration of compound **2a** respectively. Magnification x100

Further 24 hours of incubation and changes in the cell morphology are even more evident. Loss of volume is easily distinguishable in AFM (**Fig 52b**). All of the visible cells are spherical when incubated in AFM with compound **2a**, additionally bleb-like structures become visible (**Fig 52d**). Worth noting is a visible but minimal loss of volume in CM with compound **2a** (**Fig 52c**). That again suggest that **2a** competes with Arg in metabolic pathways

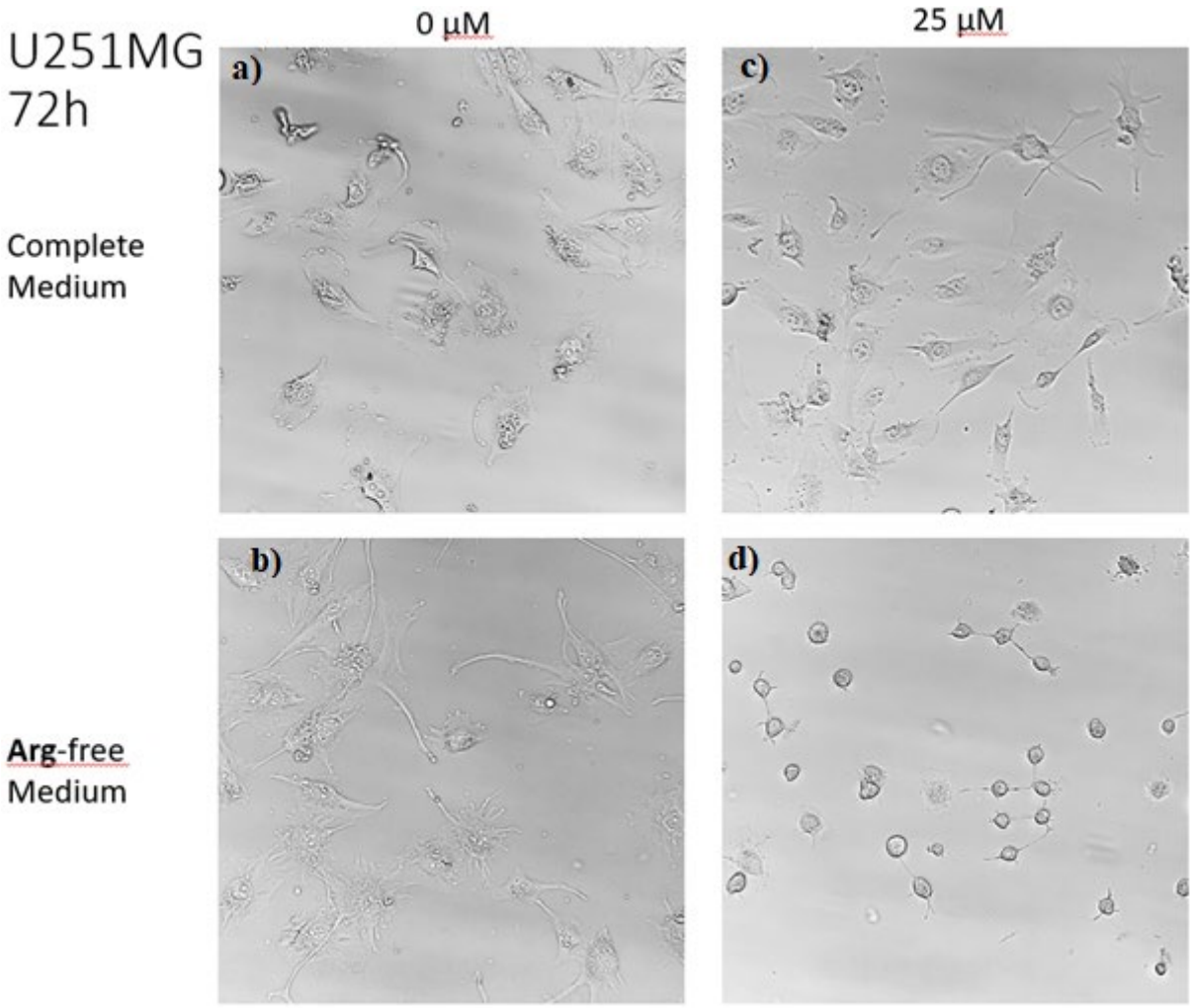


Fig 52 Microscopic assessment of U251MG cell line morphology after 72-hour incubation. Incubation was performed in the complete medium (a, c) and Arg free medium (b, d) and 0 µM (a, b), 25 µM (c, d) concentration of compound **2a** respectively. Magnification x100

Microscopic imaging studies indicate that compound **2a** mode of action is through replacing Arg in cell metabolism. Surprisingly **2a** seems to mimic Arg well enough to compete with it in the metabolic pathways. Slight changes in cell morphology in cells incubated in CM with addition of compound **2a** suggest its incorporation into proteome despite the presence of arginine.

4.6.3 Semi-quantitative protein assessment

An important factor to consider when evaluating novel chemotherapeutics is to assess whether they induce cell death through apoptosis or necrosis. The aftermath of necrosis is cell rupture that puts additional inflammatory stress on human immune system already strained with cancer. For that reason, chemotherapeutics that induce apoptosis are preferred.

Common markers for apoptosis are c-PARP, C3 and C9. c-PARP is cleaved, inactive form of poly (ADP-ribose) polymerase. c-PARP is a product of C3 activity. C3 or caspase 3 and C9 or caspase 9 are members of the cysteine-aspartic acid protease(caspase) family of enzymes. The presence of active caspases is the known indicator of apoptosis.

Figure 53 presents images of WB assays conducted on U87MG and U251MG cell lines. Protein extracts were obtained at 0 hours and after 24, 48 and 72 hours of incubation with 25 μM concentration of compound in **2a** both in the CM and AFM conditions. Protein extract of Jurkat cells after 2h from UV irradiation were analyzed as a positive control for apoptotic factors.

In both cell lines, increasing incubation time in AFM conditions correlates with gradually rising concentration of c-PARP1 protein thus clearly suggesting increasing numbers of cells undergoing apoptosis.What is intriguing in these results is the fact that no active form of C3 (c-caspase 3) was detected in both cell lines. Since c-PARP1 is a product of C3 activity and thus increasing presence of c-PARP1 demands the presence of at least detectable amount of c-caspase 3[49]. Additionally, c-caspase 3 is a product of c-caspase 9 activity[48], and this caspase 9 form is detectable only in U87MG cells treated for 72 hours and is barely visible in the lysates of 24- and 48-hour treated cells as well as in U251MG cells. Interestingly, c-caspase 9, though in lower amounts, is also seen in UMG87 cells cultured in the presence of arginine (see **Fig 53**, upper panels). These data seems to suggest that other pathway(s) could be involved in the observed c-PARP1 activity and the observed apoptotic-like changes in the treated cells..

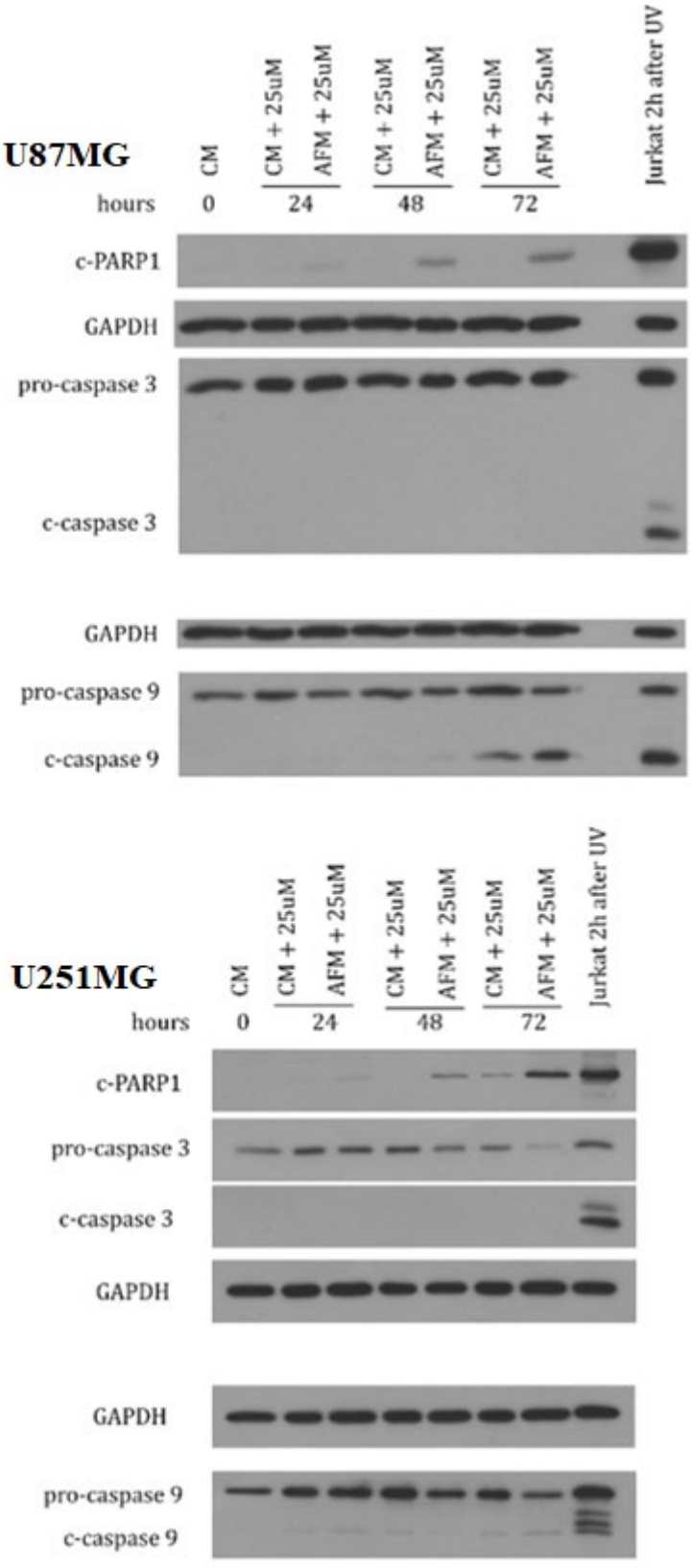


Fig 53 Western blot assessment of apoptotic markers. Cells were treated with 25 µM concentration of compound **2a** under CM and AFM conditions. The lysates were probed with the respective antibodies as described in the Methods section. GAPDH was used as a protein loading control

5. Discussion

In my work I aimed at designing, synthesis and biological evaluation of novel arginine derivatives that could be used as potential drugs in a modification of the arginine deprivation-based chemotherapy.

I started my research from exploration of *in silico* capabilities of the Arg binding pocket of ArgRS to bind to proposed by me 43 different structures derived from Arg. This research allowed me to develop general restrictions to the structure of Arg analogues, described below.

Firstly, to allow for proper ligation with $tRNA^{Arg}$ carboxylic moiety must be conserved. Additionally, to maintain its proper location within binding pocket, amine group (or its isostere) should be located at the adjacent carbon (position α). For this reason, *S*- configuration of the chiral carbon is necessary as well. That narrows down examined structures to the family of L-amino acids. Computational modeling pointed out to the fact that binding pocket should allow significant changes to the sidechain while retaining the guanidinyl isostere at the end. A wide space of the binding pocket seemed to accept even bulky 1*H*-1,2,3-triazole ring as well as sidechains longer by one carbon atom. Good docking scores of the compounds from group **1** (especially compounds **1a-d)** proved viability of that theory. Docked structures bound to the ArgRS with simulated binding energy similar to that of Arg, retaining spatial location of important groups. Despite positive results of molecular docking, the obtained compounds **1a**-**1c** have been proved to be inactive on tested cell lines. It is unknown whether this is the result of sterical hindrances of 1*H*-1,2,3-triazole ring or the fact that the length of proposed sidechains extended one-molecule-length over the length of Arg. Further research both *in silico* and on real molecules with analogues of different length would be needed to solve this problem.

Regarding the end of the sidechain, my research pointed towards a planar three-atom structure resembling a carbon atom in the sp^2 hybridization state. There is also a need for at least one hydrogen bond receptor, but a combination of one donor and one receptor is preferred. Among the tested isosteres, both *in silico* and during laboratory work (compound **1c**, **2a**) amidinyl group proved to be the best replacement for guanidinyl.

Further research on molecular dynamics in the interaction of the Arg analogues with ArgRS would be certainly beneficial. ArgRS is a highly dynamic structure that binds 3 substrates in a close proximity ($tRNA^{Arg}$, Arg and ATP). Binding one of the substrates improves the

complexes affinity towards other substrates, proving that there are conformational changes upon the binding[41].

Designing synthesis around click chemistry approach seemed viable in the conceptual work. Joining together common substrates to produce a wide library of compounds, convinced me that it is a way to obtain significant amount of various structures with the limited labor and time required. In practice, however, joining together parts with a large number of functional groups confined to the limited length of carbon chain backbone, proved to be a demanding task. Additionally, significant density of the reactive groups required application of bulky protecting groups that might have contributed to finnicky, temperamental reactions. Due to the combination of the facts that Arg analogues are amino acids and amounts of compounds needed for *in vitro* assays are relatively small, it is worth exploring the possibility of the solid-state synthesis. A synthesis on solid resins hastens the synthesis thanks to simplified workups and purification, as well as allows for parallel synthesis of dozens of compounds.

Azidoalanine is important synthetic amino acid allowing for straightforward conjugation of peptides with other macromolecules. Surprisingly, I met unexpected problems with the synthesis of commonly used azidoalanine. The synthetic approach presented in the literature proved in multiple attempts to lead to product of elimination (compound **12**). In my work, I tested different, less explored approach through sulfamidate intermediate. Despite more complicated synthesis and requiring specific catalyst, it proved to be a reliable method of producing azidoalanine with satisfactory yields.

Another aspect of my synthesis work that is worth exploring further is Pinner reaction. Amidine moiety proved to be a noteworthy replacement for guanidine both *in silico* and *in vitro*. It retains the affinity to the enzyme comparable to the guanidine while not being substrate for the enzymes degrading Arg's guanidyl group. During this work, I obtained two amidine containing structures through Pinner method. Despite demanding anhydrous condition, it proved to be a reproducible synthesis pathway to obtain amidines from cheap and commonly available substrates. Pinner method despite being discovered in 1877 [65] remains understudied and its applicability for compounds with various reactive moieties remains unknown.

From 43 designed and docked structures I managed to synthesize and evaluate five structures - **1a**, **1b**, **1c**, **1d** and, **2a**. The first four of them were then proved to be inert towards U251MG glial cancer cell line. U251MG cells have been proven to be more resilient towards Arg replacement therapies in the previous research[36,37]. For that reason, it might be feasible to test

those structures on cell lines that are more prone to the effects of the combination therapy (*i.e.* U87MG cell line). This might bring additional data on the structure-activity relationship of ArgRS binding pocket, however, due to the technical limitations it could not be done during my studies.

Out of the obtained structures, only compound **2a** exhibited expected impact on both U251MG and U87MG cell lines. Compound **2a** has been generally inert towards cells grown in the complete medium while being significantly cytotoxic towards cells grown in the Arg free medium. Its cytotoxic effect was shown as the time- and concentration-dependent decrease in viability of both cell lines. Additionally, the results on cell viability can be approximated by a typical dose-response curve. The IC_{50} values of compound **2a** after 72h of incubation are significantly higher than those of Cav for U251MG cell line (182.2 μM(**2a**) vs ~25 μM(Cav)) and comparable for U87MG cell line (11,5 μM(**2a**) vs ~15 μM(Cav)). Values of IC_{50} are arguably one of the most important characteristics of perspective therapeutic. IC_{50} in the low micromolar range is sufficient to investigate compounds **2a** as a drug candidate. Interestingly, a concentration of 25 μM of compound **2a**, *i.e.* over 6 times lower than its IC_{50} in U251MG cells was sufficient to exert the desired biological effect additionally supporting the idea of that compound being a potential drug

Furthermore, images of cells grown in the presence of compound **2a** in the absence of arginine**,** revealed changes in the morphology typical to disruption of Arg-dependent pathways, mostly affecting the cytoskeleton organization. This is in contrast when compound **2a** was applied in the complete medium condition. These observations strongly suggest that compound **2a** acts through replacing Arg in Arg-dependent pathways. It would be interesting to test whether this compound could incorporate into the polypeptide chain as was shown for Cav or to be used as a substrate for post-translational modifications as it was shown for Arg[11,37].

I also conducted a semiquantitative analysis of the level of proteins engaged in pro-death signaling pathways synthesized by the cells treated with compound **2a**. The observed presence of cleaved PARP1 (cPARP1), the hallmark indicator of apoptosis, indicates that compound **2a** may induce pro-apoptotic signaling. However, lack of active, cleaved, caspase 3 is confusing as one should expect its elevated levels in the cells with substantial amounts of cPARP1 which were treated with compound **2a** the examined conditions. Certainly, further analyses are required. It cannot be excluded that longer exposure times of the cells to the compound could be needed for clarification of this dilemma. However, being time constrained and having limited amounts of compound **2a,** I was unable to perform these experiments. Overall, the assessment

of the results presented herein with respect to other research on Arg replacement in the examined glioblastoma cell lines[37] point towards apoptosis as a mechanism of cell death induced by compound **2a** in the absence of arginine.

Additional assessment on healthy glial cells is needed to ensure that compound **2a** is selective towards cancer cells and not to normal, untransformed cells as it was shown for canavanine[37]. Also, as it was abovementioned, proteomic assessment with mass spectroscopy would be additional proof that compound **2a** replaces and maybe even competes with Arg in its metabolism.

Importantly, a direct comparison of efficacy of Isp, Cav and compound **2a** would be beneficial. It would be also important to measure a stability of these three compounds in physiological conditions. It has been shown that Cav has a short half-life that creates a concern for its perspective use in a possible combination therapy[85]. The opposite concern is for Isp as it is toxic since it accumulates in the organisms as it was shown for Australian cattle grazing on *Indigofera* plants[42].

Another issue that should be considered in the usage of compound **2a** in a possible anti-glioblastoma treatment is its ability to cross the brain-blood barrier. Normally, Arg is transported to the brain with the use of specific arginine transporter[85]. It was shown that Cav, differing from Arg only by substitution of the carbon atom in the position δ with oxygen atom, seem to use the same mechanism[86]. Thus, one can assume that compound 2a, being a structural analog of both Arg and Cav could utilize the same mechanism, making it accessible for and effective towards transformed cells in the tumor.

Altogether, the above results of the assessment of the compounds **2a** activity on human glioblastoma U87MG and U251MG cell lines are in line with the previous research performed on a potential Arg replacement therapy *in vitro*. This work is yet another proof that a therapeutical strategy combining the Arg starvation with substitution of Arg analogue is effective *in vitro* and is worth to examine its rationale and effectiveness in other glioblastoma cell lines and/or other cell lines/animal models, and potentially in human glioma patients.

6. Summary and conclusions

1. Molecular docking part of the work presented in this dissertation provided new insights onto the shape and structure of the binding pocket of ArgRS. It proved that amidine containing analogues are viable replacement for Arg. Also, any other alterations to the sidechain structure than at its end have no impact on compound affinity towards receptor as well as their mode of binding.

2. Synthesis conducted during this research proved that cyclic sulfamidates are reliable intermediates for azidoalanine synthesis, providing a novel pathway to obtain this amino acid important for synthesis of conjugates of macromolecules. I also provided a new synthesis approach applying under researched Pinner reaction, proving that it is a viable route to obtain amidines with satisfactory yields.

3. I provided the novel Arg analogue (compound **2a**) with the efficacy and mode of action similar to that of Cav. While this not adds upon previously conducted research on Arg replacement therapy but also provides a candidate for a new potential chemotherapeutic.

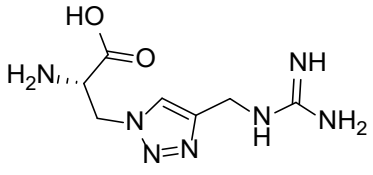
7. Methods

7.1 Molecular docking

Docking calculations were performed using AutoDock Vina software. The experimental crystal structure of hArgRS (PDB entry 4q2t[88]) has been prepared according to protein preparation procedure recommended by software creators[53]. All tested ligands were processed according to the ligand preparation procedure recommended[53]. Default input parameters were used in all computations. Scoring was based on a binding energy score given by AutoDock Vina. All computations were performed on an Intel® Core™ i7-4702MQ 3.2 GHz processor running Fedora 28 Workstation Linux distribution.

7.2 Organic synthesis

Thin-layer chromatography was performed on pre-coated Merck silica gel 60 F254 (TLC, 0.2 mm) with detection under UV light (254 nm) or visualized with phosphomolybdic acid stain. Column chromatography was performed with silica gel (200-400 mesh, Merck). High Resolution Mass Spectra (Electrospray Ionisation, ESI) were performed on a Mariner® spectrometer in positive ionization mode. The NMR spectra were measured on a Varian VNMRS spectrometer (^{1}H NMR at 500 MHz and ^{13}C NMR at 125 MHz) or Varian Mercury spectrometer (^{1}H NMR at 400 MHz and ^{13}C NMR at 100 MHz). ^{1}H and ^{13}C chemical shifts (δ) are reported in parts per million (ppm) relative to the solvent signals: $CDCl_3$, δ_H (residual $CHCl_3$) 7.26 ppm, δ_C 77.16 ppm; or DMSO-d_6, δ_H (residual DMSO-d_5) 2.50 ppm, δ_C 39.52 ppm; signals are quoted as ‘s’ (singlet), ‘d’ (doublet), ‘t’ (triplet), ‘dd’ (dublet of dublets), ‘m’ (multiplet), and ‘bs’ (broad singlet). Solvents were distilled off under reduced pressure on a rotating evaporator.

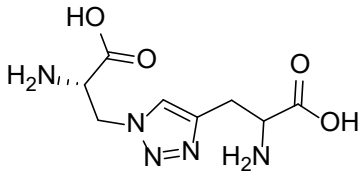


1a

7.2.1 (*S*)-2-Amino-3-(4-(guanidinomethyl)-1*H*-1,2,3-triazol-1-yl)propanoic acid 1a

Compound **24** (500 mg, 0.8 mmol) was dissolved in anhydrous MeOH (5 mL), followed by addition of palladium on activated charcoal (10% wt, 50 mg). The reaction mixture

was stirred for 16 h under hydrogen atmosphere (balloon pressure). Afterwards it was filtered through a Celite® pad and the pad was washed thoroughly with MeOH (15 mL). From the filtrate, volatiles were distilled off under reduced pressure. The residue was dissolved in a mixture of DCM:TFA (5 mL, 1:1, v/v). The solution was stirred for 1 h. Volatiles were removed under the flow of argon. The residue was suspended in cold diethyl ether (5 mL). Formed precipitate was filtered off and washed with a minimal amount of diethyl ether. Compound **1a** was obtained as a white fluffy solid (220mg, 97%). $^{13}C\{^{1}H\}$ NMR (125.71 MHz, D_2O): δ_C 36.8 (C-*C*H_2-NH), 50.3 (>*C*H-NH_2), 54.8 (CH-*C*H_2-N), 125.4 (N-*C*H=C), 143.9 (=*C*-CH_2), 157.6 (NH-*C*NH-NH_2), 170.4 (-COOH). ^{1}H NMR (499.87 MHz D_2O): δ_H 4.98 (2H, s) 5.41-5.42 (2H, m), 8.43 (1H, s). [Appendix 8.2.1] HRMS (ESI) $[M-H]^+$ calc for $C_7H_{14}N_7O_2$ 227.1203, found m/z 227.1207.

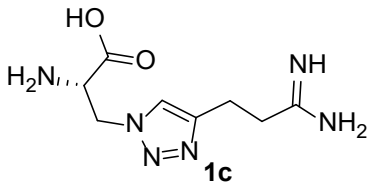


1b

7.2.2 (2*S*)-2-Amino-3-(4-(2-amino-2-carboxyethyl)-1*H*-1,2,3-triazol-1-yl)propanoic acid 1b

Compound **26** (500 mg, 0.94 mmol) was dissolved in anhydrous MeOH (5 mL), followed by addition of palladium on activated charcoal (10% wt, 50 mg). The reaction mixture was stirred for 16 h under hydrogen atmosphere (balloon pressure) and filtered through Celite® pad. The pad was washed thoroughly with MeOH. From the filtrate, volatiles were distilled off under reduced pressure. The residue was dissolved in the DCM:TFA mixture (1:1, v/v, 5 mL) and stirred for 1h at room temperature. Volatiles were removed under the flow of argon gas. The residue was suspended in cold diethyl ether (5 mL). Formed precipitate was filtered off. Compound **1b** was obtained as a pale

yellow fluffy solid (238mg, 98%). $^{13}C\{^{1}H\}$ NMR (125.71 MHz, TFA/$CDCl_3$): 24.8 (C-*C*H_2-CH), 27.3 (C-*C*H_2-N), 50.1 (CH_2-*C*HNH_2-COOH), 53.4 (COOH-*C*HNH_2-CH_2), 125.3 (N-*C*H=C), 132.4 (=*C*N-CH_2), 168.6 (COOH), 171.0 (COOH); ^{1}H NMR (499.87 MHz, TFA/$CDCl_3$): 3.67-3.80 (2H, m), 4.70 (1H, m), 5.01 (1H, m), 5.29 (2H, m), 8.31 (1H, s) [Appendix 8.2.2] HRMS (ESI) $[M-H]^+$ calc for $C_8H_{14}N_5O_4$ 243.0968, found m/z 243.0973.

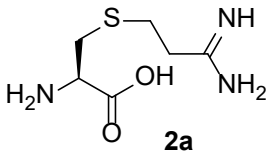


1c

7.2.3 (*S*)-2-Amino-3-(4-(3-amino-3-iminopropyl)-1*H*-1,2,3-triazol-1-yl)propanoic acid 1c

Compound **30** (500 mg, 0.97 mmol) was dissolved in anhydrous MeOH (5 mL), followed by addition of palladium on activated charcoal (10% wt, 50 mg). Reaction mixture was kept for 16 h under hydrogen atmosphere (balloon pressure). The mixture was filtered through a Celite® pad. The pad was washed thoroughly with MeOH. From the filtrate, volatiles were distilled off under reduced pressure. The residue was dissolved in a mixture of DCM:TFA (1:1, v/v, 5 mL) and stirred for 1 h at room temperature. Volatiles were removed under the flow of argon gas. The residue was suspended in cold diethyl ether (5 mL). Formed precipitate was filtered off. Compound **1c** was obtained as a pale yellow fluffy solid (223 mg, 99%). $^{13}C\{^{1}H\}$ NMR (125.71 MHz, TFA-*d*): δ_C 19.15 (C-*C*H_2-CH_2), 32.66 (CH_2-*C*H_2-C), 52.28 (CH-*C*H_2-N), 53.55 (HOOC-*C*HNH_2-CH_2), 130.82 (N-*C*=C), 145.10 (>*C*-CH_2), 168.69 (-COOH), 178.74 (-*C*NH-NH_2); ^{1}H NMR (499.87 MHz, TFA-*d*): δ_H 2.99 (2H, s), 3.35 (2H, m), 3.86 (1H, m) 4.04 (1H, d, $^{3}J_{HH}$ = 17.8), 5.27 (1H, bs), 5.57 (2H, bs), 8.60 (1H, s)

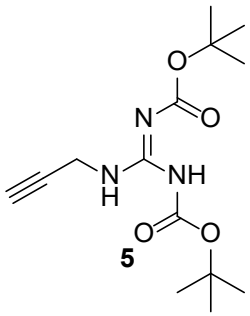
[Appendix 8.2.3]. HRMS (ESI) $[M-H]^+$ calc for $C_8H_{15}N_6O_2$ 227.1251, found m/z 227.1251.



2a

7.2.4 *S*-(3-Amino-3-iminopropyl)-L-cysteine 2a

Compound **35** (0.15 g, 0.49 mmol) was dissolved in a DCM:TFA mixture (1:1, v/v, 5 mL) and stirred overnight at room temperature. Volatiles were under the flow of argon gas. The residue was suspended in cold diethyl ether (5 mL). Formed precipitate was filtered off. Compound **2a** was obtained as a pale yellow fluffy solid (94 mg, 99%). $^{13}C\{^1H\}$ NMR(125.71 MHz, DMSO-d_6): δ_C 26.5 (S-$\underline{C}H_2$-CH_2), 34.1 (CH-$\underline{C}$H2-SH), 45.3 (CH_2-$\underline{C}H_2$-C), 52.6 (>$\underline{C}$H-CH_2), 171.9 (-$\underline{C}$OOH), 172.8 (-$\underline{C}NH_2$=NH). ^1H NMR δ_H (499.87 MHz, DMSO-d_6): 2.73 (2H, m), 2.81 (2H, m), 2.95-3.06 (2H, m), 3.77 (1H, m) 7.62 (1H, bs), 8.60 (1H, s) [Appendix 8.2.4]. HRMS (ESI) $[M-H]^+$ calc for $C_6H_{14}N_3O_2S$ 1920.801, found m/z 191.0806.

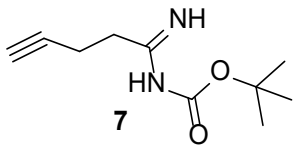


5

7.2.5 *N*,*N*'-Bis(*tert*-butoxycarbonyl)-1-(prop-2-yn-1-yl)guanidine 5 [81]

Propargylamine (2.0 mL, 37 mmol) was added to a solution of compound **22** (3.63 g, 12.5 mmol) in THF (120 mL). The reaction mixture was stirred at 40 °C for 24 h, cooled to room temperature and volatiles were distilled off under reduced pressure. The residue

was subjected to flash column chromatography (SiO_2, AcOEt:hexane, 1:9, v/v). Compound 5 was obtained as a white solid (6.1g, 80%). $^{13}C\{^1H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 28.1 (3xCH_3), 28.4 (3xCH_3), 30.8 (CH_2NH) 72.4 (C≡*C*H), 78.97 (*C*(CH_3)$_3$), 79.72 (CH_2*C*C≡H), 83.53 (*C*(CH_3)$_3$), 153.16 (guanidine-*C*), 155.74 (*t*-BuO-*C*O-NH), 163.38 (*t*-BuO-*C*O-NH). ^1H NMR (499.87 MHz, $CDCl_3$): δ_H 1.52 (6 x CH_3, s), 2.25 (1H, t, $^2J_{HH}$ = 2.5 Hz), 4.22 (2H, dublet of dublets, $^3J_{HH}$ = 4.9, $^2J_{HH}$ = 2.6 Hz), 8.45 (1H, s), 11.43 (1H, s). The NMR spectra are consisted with those published in the chemical literature[81] [Appendix 8.2.5]

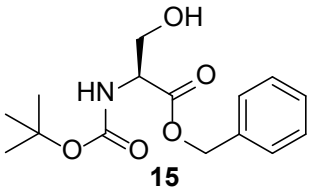


7

7.2.6 *tert*-Butyl (1-iminopent-4-yn-1-yl)carbamate 7

Reaction was conducted under the argon atmosphere. 4-Pentynoatenitrile (500 mg, 6.32 mmol) was dissolved in anhydrous diethyl ether (30 mL). The solution was cooled to 0 °C and acetyl chloride (11.3 mL, 158 mmol) was added dropwise, followed by dropwise addition of anhydrous methanol (10 mL). Reaction was stirred at 0 °C overnight. Volatiles were removed under the flow of anhydrous argon gas. The residue was dissolved in anhydrous methanol (5 mL) and cooled to 0 °C. Afterwards, ammonia saturated anhydrous methanol (23 mL) was added dropwise and the mixture was stirred at 0 °C overnight. The excess of ammonia was blown out with a flow of argon gas. From the residue, volatiles were distilled off under reduced pressure. The residue (670 mg) was obtained as a white fluffy solid and without further purification was dissolved in a mixture of 1,4-dioxane (9 mL) and 0,4 M aqueous NaOH (33 mL), followed by addition of a solution of di-*tert*-butyl dicarbonate (1.7 g, 7.8 mmol) dissolved in 1,4-dioxane (9 mL). The reaction mixture was stirred for 16 h at room temperature and volatiles were

distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM). Compound **11** was obtained as a white solid (0.63 g, 51% over 2 steps). $^{13}C\{^1H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 13.5 (C*C*H_2CH_2) 28.1 (3 x CH_3), 35.3 (CH_2*C*H_2C) 69.0 (CH), 79.9 (C(CH_3)$_3$), 82.8 (C(CH_3)$_3$), 82.9 (CH_2*C*≡CH), 150.6 (*t*-BuO-*C*O-NH), 173.1 (CH_2*C*NH). ^1H NMR (400.11 MHz, $CDCl_3$): δ_H 1.48 (9H, s), 1.96 (1H, m), 2.52 (2H, d, $^3J_{HH}$ = 2.5 Hz), 3.00 (2H, s) 7.62 (1H, bs) [Appendix 8.2.6]. HRMS (ESI) $[M–H]^+$ calc for $C_{10}H_{17}N_2O_2$ 197.1285, found m/z 197.1287.

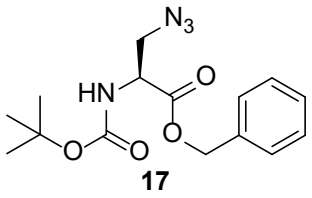


15

7.2.7 *O*-Benzyl-*N*-*tert*-butoxycarbonyl-L-serine 15 [90]

L-Serine (9.5 g, 90 mmol) was added to a mixture of K_2CO_3 (12.5 g, 90 mmol) and water (50 mL), followed by addition of a solution of di-*tert*-butyl dicarbonate (23.6 g, 108 mmol) in 1,4-dioxane (50 mL). The mixture was stirred overnight at room temperature until complete conversion of L-serine (TLC, DCM:MeOH, 99:1, v/v; phosphomolybdic acid stain). Volatiles were distilled off under reduced pressure. Leftover solid was dried for 48 h in a vacuum desiccator over anhydrous P_2O_5. Dried solid was suspended in DMF (50 mL), followed by dropwise addition of benzyl bromide (12 mL, 100 mmol) and stirring overnight at room temperature. Upon complete conversion of the amino acid (TLC, DCM:MeOH, 99:1, v/v; phosphomolybdic acid stain) the reaction mixture was diluted with water (200 mL) and washed with ethyl acetate (3 x 50 mL). Organic extracts were combined and dried over anhydrous $MgSO_4$. The drying agent was filtered off. From the filtrate, volatiles were distilled off under

reduced pressure. The obtained colorless heterogeneous oily mixture was suspended in hexane (200 mL) and sonicated for 1 h. The formed solid was filtered off. Compound **15** was obtained as a white powder (21.8 g, 82% over two steps). $^{13}C\{^{1}H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 28.4 (3 x CH_3), 56.0 (>*C*H-NHBoc), 63.6 (HO-*C*H_2-C), 67.5 (Ph-*C*H_2-O), 80.4 (*C*$(CH_3)_3$), 128.3 (Ar-C), 128.5 (Ar-C), 128.7 (Ar-C), 135.4 (Ar-C), 155.9 (*t*-BuO-*C*O-NH), 170.9 (-COOBn). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.43 (9H, bs), 3.87-3.99 (2H, m), 4.41 (1H, bs), 5.17 and 5.22 (2H, AB-quartet, $^{2}J_{AB}$ 12.3 Hz), 5.56 (1H, d, $^{3}J_{HH}$ 7.8 Hz), 7.34 (5H, m). The NMR spectra are consisted with those published in the chemical literature.[89] [Appendix 8.2.7]

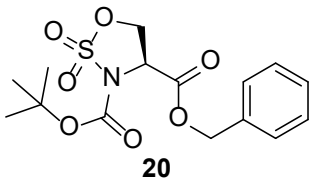


17

7.2.8 Benzyl (*S*)-3-azido-2-((*tert*-butoxycarbonyl)amino)propanoate 17 [74]

A solution of sodium azide (1.6 g, 25.1 mmol) in water (42 mL) was added to a solution of compound **20** (3 g, 8.4 mmol) in acetone (42 mL). The reaction mixture was left stirred overnight at room temperature. Upon total conversion of compound **2** (TLC, DCM, phosphomolybdic acid stain) volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM). Compound **3** was obtained as a pale-yellow oil (2.6g, 99%). $^{13}C\{^{1}H\}$ NMR (125.71, MHz, $CDCl_3$): δ_C 28.4 (3 x CH_3), 52.8 (>*C*H-NHBoc), 53.8 (N_3-*C*H_2-C), 67.9 (Ph-*C*H_2-O), 80.6 (*C*$(CH_3)_3$), 128.8 (Ar-C), 128.8 (Ar-C), 128.9 (Ar-C), 135.0 (Ar-C), 155.2 (t-BuO-*C*O-NH), 169.8 (-COOBn). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.45 (9H, bs), 3.74 (2H, d, $^{3}J_{HH}$ 3.4 Hz), 4.50 (1H, m), 5.19-5.24 (2H, AB-quartet, $^{2}J_{AB}$ 12.2 Hz), 7.36 (5H, m).

The NMR spectra are consisted with those published in the chemical literature.[74] [Appendix 8.2.8]

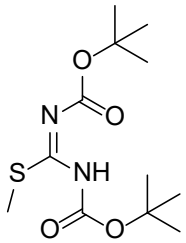


20

7.2.9 4-Benzyl 3-(*tert*-butyl) 1,2,3-oxathiazolidine-3,4-dicarboxylate 2,2-dioxide 20[79]

Compound **15** (2.9 g, 10.2 mmol) was dissolved in anhydrous DCM (50 mL) and added dropwise to a solution of thionyl chloride (1.83 mL, 25.5 mmol) in anhydrous DCM (100 mL) cooled with an acetone/dry ice bath. After stirring for 5 min, pyridine (5 mL, 52 mmol) was added. After further 15 min, the acetone/dry ice bath was removed, and the reaction mixture left stirred for 1 h to reach room temperature and quenched with ice water (100 mL). Organic phase was separated, washed with a saturated aqueous $NaHCO_3$ solution (50 mL), brine (50 mL) and dried over anhydrous $MgSO_4$. The drying agent was filtered off. From the filtrate, volatiles were distilled off under reduced pressure. The residue was suspended in acetonitrile (50 mL) and cooled in an ice/water bath. Ruthenium chloride (42 mg, 0.2 mmol) and sodium periodate (3.42 g, 16 mmol) were dissolved in cold water (50 mL) and added to the reaction mixture. Reaction was kept in an ice/water bath for 1 h and quenched with saturated aqueous $NaHCO_3$ solution. Organic phase was separated, washed with ethyl acetate (3 x 50 mL). Organic extracts were combined, washed with brine (30 mL) and dried over anhydrous $MgSO_4$. The drying agent was filtered off. From the filtrate, volatiles were distilled off under reduced pressure. The residue (dark green powder) was subjected to flash column chromatography (SiO_2, AcOEt:hexane, 1:9, v/v). Compound **20** was obtained as a white powder (2.7g, 74%). $^{13}C\{^1H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 27.8 (3 x CH_3), 57.7

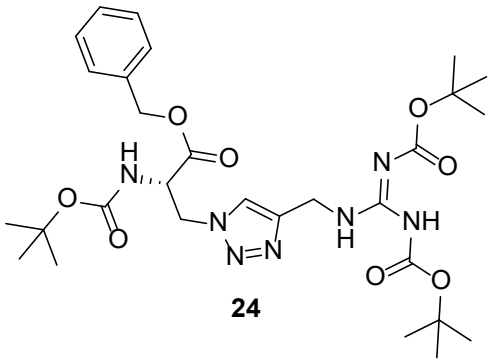
(>*C*H-NHBoc), 67.5 (HO-*C*H_2-C), 68.5 (Ph-*C*H_2-O), 86.3 (*C*$(CH_3)_3$), 128.4 (Ar-C), 128.8 (Ar-C), 128.9 (Ar-C), 134.5 (Ar-C), 148.1 (t-BuO-*C*O-NH), 167.0 (-COOBn). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.48 (9H, bs), 4.65 (1H, ABX system part A, X = H, $^2J_{AB}$ = 9.5 Hz, $^3J_{AX}$ = 2.1 Hz), 4.74 (1H, ABX system part B, X = H, $^2J_{AB}$ = 9.5 Hz, $^3J_{BX}$ = 6.5 Hz), 4.82 (1H, ABX system part X, broad dublet), 5.21-5.31 (2H, AB-quartet, $^2J_{AB}$ = 12.1 Hz), 7.36 (5H, m). The NMR spectra are consisted with those published in the chemical literature.[79] [Appendix 8.2.10]



22

7.2.10 *N*,*N*'-Bis(*tert*-butoxycarbonyl)-*S*-methylisothiourea 22[81]

A solution of di-*tert*-butyl dicarbonate (21.83 g, 100 mmol) in 1,4-dioxane (60 mL) was added to a solution of *S*-methylisothiourea sulfate (5.56 g, 20 mmol) in an aqueous solution of NaOH (0.4 M, 100 mL). The reaction mixture was left stirred for 16 h at room temperature. The reaction mixture was concentrated at rotary evaporator to a half of the starting volume. White precipitate was filtered off and washed with a minimal amount of water. Compound **22** was obtained as a white solid (4.18 g, 72%). $^{13}C\{^1H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 14.5 (CH_3S) 28.1 (6 x CH_3), 81.2 (*C*$(CH_3)_3$), 83.3 (*C*$(CH_3)_3$), 150.87 (*t*-BuO-*C*O-N), 160.87 (*t*-BuO-*C*O-N), 171,55 (urea-*C*). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.52 (6 x CH_3, s), 2.38 (3H, m), 11.43 (1H, bs). The NMR spectra are consistent with those published in the chemical literature[81]. [Appendix 8.2.11]

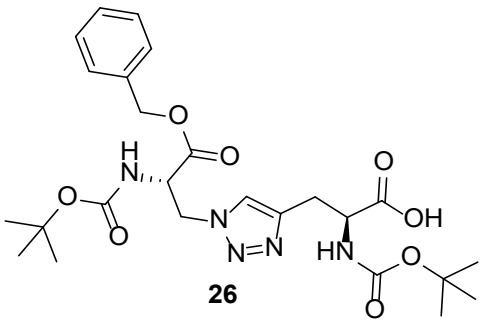


24

7.2.11 Benzyl (*S*,*Z*)-3-(4-((2,3-bis(*tert*-butoxycarbonyl)guanidino)methyl)-1*H*-1,2,3-triazol-1-yl)-2-((*tert*-butoxycarbonyl)amino)propanoate 24

Compound **5** (714 mg, 2.4 mmol) was dissolved in THF (5 mL), followed by addition of $CuSO_4$ (32 mg, 0.2 mmol) and a solution of sodium ascorbate (75 mg, 0.4 mmol) in water (10 mL). After stirring for 5 min at room temperature, a solution of compound **17** **(**640 mg, 2 mmol) in THF (5 mL) was added. The reaction mixture was left stirred for 72 h until complete conversion of compound **17** was detected (TLC, DCM; phosphomolybdic acid stain). After completion, the reaction mixture was diluted with ethyl acetate (30 mL) and washed with a saturated aqueous solution of NH_4Cl (3 x 30 mL) followed by brine (30 mL). Organic phase was dried with anhydrous $MgSO_4$.The drying agent was filtered off. From the filtrate, volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM/MeOH, 99:1, v/v). Compound **6** was obtained as a pale yellow glassy solid (0.99 g, 81%). The chemical shifts in the $^{13}C\{^1H\}$ NMR were assigned from the APT $^{13}C\{^1H\}$ NMR spectrum (125.71 MHz, $CDCl_3$); quaternary and methylene carbon atoms are phased negatively (C-4°↓ or CH_2↓, respectively) while methine and methyl carbon atoms are phased positively (CH_3↑ or CH↑, respectively). $^{13}C\{^1H\}$ NMR (125.71 MHz, $CDCl_3$): δ_C 28.1↑ (3 x CH_3) , 28.3↑ (3 x CH_3), 28.4↑ (3 x CH_3), 36.4↓ (C-*C*H_2-NH), 51.0↓ (CH-*C*H_2-N), 53.8↑ (>*C*H-NHBoc), 68.1↓ (Ph-*C*H_2-O), 79.5↓ (*C*$(CH_3)_3$) ,

80.7↓ (*C*(CH_3)$_3$), 83.4↓ (*C*(CH_3)$_3$), 123.5↑ (N-*C*H=C), 128.8↑ (Ar-C), 128.8↑ (Ar-C), 128.8↑ (Ar-C), 134.8↑ (Ar-C), 144.0↓ (=C-*C*H_2), 153.0↓ (*t*-BuO-*C*O-NH), 155.2↓ (*t*-BuO-*C*O-NH), 156.0↓ (*t*-BuO-*C*O-NH), 163.4↓ (NH-*C*NBoc-NHBoc), 169.0↓ (-COOBn). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.40 (CH_3, s), 1.45 (CH_3, s), 1.49 (CH_3, s), 4.64 (2H, d, $^2J_{HH}$ = 4.2 Hz), 4.71-4.81 (3H, m), 7.32-7.35 (6H, m), 8.07 (NHBoc, bs), 8.76 (NH, broad triplet), 11.43 (NHBoc, s) [Appendix 8.2.12]

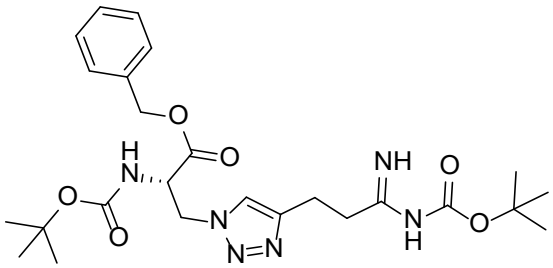


26

7.2.12 3-(1-((*S*)-3-(Benzyloxy)-2-((*tert*-butoxycarbonyl)amino)-3-oxopropyl)-1*H*-1,2,3-triazol-4-yl)-2-((*tert*-butoxycarbonyl)amino)propanoic acid 26

(*S*)-2-((*Tert*-butoxycarbonyl)amino)pent-4-ynoic acid (512 mg, 2.4 mmol) was dissolved in *tert*-butanol (5 mL), followed by addition of $Cu(OAc)_2$ (73 mg, 0.4 mmol) and a solution of sodium ascorbate (150 mg, 0.8 mmol) in water (10 mL). After stirring for 5 min at room temperature, a solution of compound **17 (**640 mg, 2 mmol) in *tert*-butanol (5 mL) was added. The reaction was left stirred for 72 h until complete conversion of compound **17** (TLC, DCM; phosphomolybdic acid stain). After completion, the reaction mixture was diluted with ethyl acetate (30 mL) and washed with saturated aqueous NH_4Cl (3 x 30 mL) and brine (30 mL). The organic phase was dried with anhydrous $MgSO_4$. The drying agent was filtered off. From the filtrate, volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM/MeOH, 99:1, v/v). The obtained bright green

glassy solid was dissolved in a mixture of acetone (1 mL) and of water (5 mL) and shaken with Chelex® resin (1 g) for 30 min. The resin was filtered off. From the filtrate, volatiles were distilled off under reduced pressure. Compound **8** was obtained as a white glassy solid (812 mg, 76%). $^{13}C\{^{1}H\}$ NMR (125.71 MHz, $CDCl_3$): 18.0 (CH-*C*H_2-C), 28.2 (3 x CH_3), 28.3 (3 x CH_3), 50.6 (>*C*H-NHBoc), 53.8 (CH-*C*H_2-N), 57.9 (>*C*H-NHBoc), 67.7 (Ph-*C*H_2-O), 79.5 (*C*$(CH_3)_3$), 80.4 (*C*$(CH_3)_3$), 124.2 (N-*C*H=C), 128.7 5(Ar-C), 134.9 (Ar-C), 143.4 (=C-CH_2), 155.4 (*t*-BuO-*C*O-NH), 155.8 (*t*-BuO-*C*O-NH), 169.2 (-COOBn), 174.0 (-COOH). ^{1}H NMR (499.87 MHz, $CDCl_3$): δ_H 1.37 (2xCH_3, bs), 3.23 (CH_2, bs), 4.69-4.74 (3H, m), 5.12 (2H, bs), 5.67 (1H, m) 7.31 (6H, m), [Appendix 8.2.13]

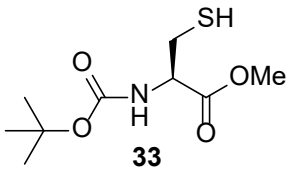


30

7.2.13 Benzyl (*S*)-2-((*tert*-butoxycarbonyl)amino)-3-(4-(3-((*tert*-butoxycarbonyl) amino)-3-iminopropyl)-1*H*-1,2,3-triazol-1-yl)propanoate 30

Compound **7** (470 mg, 2.4 mmol) was dissolved in THF (5 mL), followed by addition of $CuSO_4$ (32 mg, 0,2 mmol and sodium ascorbate (75 mg, 0.4 mmol) in water (10 mL). After stirring for 5 min at room temperature, a solution of compound **17 (**640 mg, 2 mmol) in THF (5 mL) was added. The reaction mixture was left stirred for 72 h until complete conversion of compound **17** (TLC, DCM; phosphomolybdic acid stain). After completion, the reaction mixture was diluted with ethyl acetate (30 mL) and washed with saturated an aqueous solution of NH_4Cl (3 x 30 mL) followed by brine (30 mL).

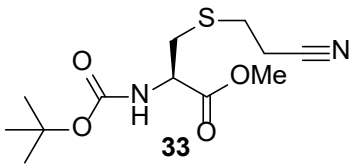
The organic phase was dried with anhydrous $MgSO_4$. From the filtrate, volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM/MeOH, 99:1, v/v). Compound **30** was obtained as a white glassy solid (0.76 g, 74%). $^{13}C\{^1H\}$ NMR (100.62 MHz, $CDCl_3$): δ_C 19.9 (C-*C*H_2-CH_2), 28.1 (3 x CH_3), 28.3 (3 x CH_3), 35.7 (CH_2-*C*H_2-C), 51.0 (CH-*C*H_2-N), 53.8 (BnOOC-*C*HNHBoc-CH_2) 68.2 (Ph-*C*H_2), 80.8 (*C*$(CH_3)_3$), 82.7 (*C*$(CH_3)_3$), 123.3 (N-C=C), 128.8 (Ar-C), 128.9 (Ar-C), 134.8 (Ar-C), 146.3 (C=*C*-*C*H_2), 150.4 (*t*-BuO-*C*O-NH), 155.2 (*t*-BuO-*C*O-NH), 169.1 (-COOBn), 173.4 (-*C*NH-NHBoc); ^1H NMR (400.11 MHz $CDCl_3$): δ_H 1.41 (9H, s), 1.46 (9H, s), 3.00 (2H, m), 3.08 (2H, m), 4.74 (3H, m) 5.16 (2H, m), 7.23 (1H, s),7.35 (5H, s). [Appendix 8.2.14]



33

7.2.14 Methyl ((*tert*-butoxycarbonyl)-L-cysteinate 33[90]

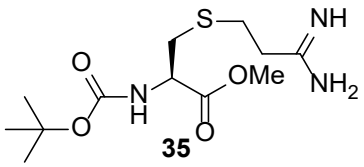
N,*N*'-Bis(*tert*-butoxycarbonyl)-cystine dimethyl ester (2 g, 4.27 mmol) and triphenylphosphine (1.23 g, 4.7 mmol) was dissolved in methanol (17 mL), followed by addition of sodium acetate (0.14g, 1.7 mmol) and a solution of acetic acid (0.12 mL) in water (9 mL). The reaction mixture was left stirred at room temperature for 16 h. After completion (TLC, hexane:ethyl acetate, 4:1, v/v; phosphomolybdic acid stain) reaction mixture was diluted with DCM (200 mL). The organic layer was washed with water (2 x 50 mL) and brine (50 mL). After drying with anhydrous $MgSO_4$, volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM). Compound **33** was obtained as white thick oil (1.8 g, 91%).



33

7.2.15 Methyl *N*-((*tert*-butoxycarbonyl)-*S*-(2-cyanoethyl)-L-cysteinate 34[91]

Compound **33** (1.25 g, 5.4 mmol) was dissolved in mixture of MeOH (14 mL) and water (7 mL), followed by acrylonitrile (0.38 g, 7,2 mmol) of and TEA (1.9 g, 18.9 mmol). The reaction mixture was left stirred overnight at room temperature. Volatiles were distilled off under reduced pressure. The residue was subjected to flash column chromatography (SiO_2, DCM). Compound **34** was obtained as an off-white fluffy solid (1.19 g, 77%).



35

7.2.16 Methyl *S*-(3-amino-3-iminopropyl)-*N*-((*tert*-butoxycarbonyl)-L-cysteinate 35

Reaction was conducted under the argon atmosphere. Compound **34** (0.22 g; 1.26 mmol) and anhydrous MeOH (2 mL) were added to a flame dried Schlenk flask, under inert gas. After cooling in dry ice/acetone bath, acetyl chloride (3.5 g, 31.5 mmol) was added dropwise. The reaction was left stirred at 0 °C for 6 days. Volatiles were removed under the flow of argon. The residue was dissolved in anhydrous MeOH (0.75 mL), cooled in ice bath, followed by addition of ammonia saturated anhydrous MeOH (4 mL). The reaction mixture was left stirred at 0°C for 16 h. Volatiles were distilled off under reduced pressure. The residue was suspended in cold ethyl acetate (5 mL) and filtered. Volatiles were distilled off under reduced pressure from the filtrate. Compound **35** was obtained as a white solid and used in the synthesis of **2a** without purification.

7.3 Biological assays

All biological experiments were performed at the discretion of Prof. Maria Jolanta Rędowicz at the Laboratory of Molecular Basis of Cell Motility the Nencki Institute of Experimental Biology, Polish Academy of Sciences. The author of this work would like to thank Dr. Olena Karatsai from Prof. Rędowicz's group for her invaluable help in obtaining and interpreting that part of the thesis.

7.3.1 Cell lines and culture conditions

Human glioblastoma U251MG and U87MG cell lines were obtained from the Cell Lines Service (Eppelheim, Germany). Cells were cultivated at 37°C with 5% CO2 in Dulbecco′s Modified Eagle Medium (DMEM; Gibco 31966021, Waltham, MA, USA) supplemented with GlutaMAX-1, 10% heat inactivated fetal bovine serum (FBS; Gibco 10500064) and antibiotics, 1% penicillin/streptomycin (Gibco 15140122). Whenever applicable for experiments, DMEM (Sigma-Aldrich D9443) with or without 0.4 mmol/L arginine was formulated (complete medium (CM) and arginine-free medium (AFM), respectively, supplemented with 5% dialyzed serum (Sigma-Aldrich F0392).

7.3.2 MTS Cell Viability Assay

The cell viability MTS assay was performed according to the manufacturer′s instructions[92]. Cells were seeded on 96-well plates (8000 cells per well) in standard DMEM medium and left to adhere for 18 hours. Cell-containing wells were washed twice with PBS followed by addition of CM or AFM containing respective concentration of the tested compound. Tecan SUNRISE XFluor4 plate reader at a wavelength of 490 nm was sued to measure the absorbance. The ratio of the viable cells after treatment was calculated versus the absorbance recorded for the control conditions (*i.e.* 0 μM concentration of the tested compound).

7.3.3 Western Blot Analysis

Cell lysis was performed in RIPA Buffer with addition of 50 mM NaF, 2 mM Na_3VO_4, 1 mM PMSF, and protease and phosphatase inhibitor cocktails (04693116001 and 04906837001, Roche, Mannheim, Germany) at 4°C for 20 min. Cell extracts were obtained after centrifugation at 13,500× g at 4°C for 20 min. Bradford method was applied to determine protein concentrations Lysates were treated with the standard Laemmli buffer for 5 min at 98°C. Protein samples were separated in 12% SDS–polyacrylamide gel (SDS-PAGE) and transferred onto nitrocellulose membranes (Amersham 10600002, Freiburg, Germany). The membranes were

blocked with 5% fat-free milk or 5% BSA (Bioshop ALB001, Burlington, Canada) in TBS/0.2% Triton X-100. GAPDH was used as an internal protein loading control. One-minute-UV-irradiated Jurkat cells lysates were parallelly analyzed as an apoptotic cell death control. c-PARP (#9546), Caspase 3 (#9662), caspase 9 (#9502) were detected using respective antibodies used at 1:1000 dilution (obtained from Cell Signaling Technologies Danvers, MA, USA). GAPDH was detected with antibody from Millipore (MAB374, St. Louis, MO, USA). Secondary HRP conjugated anti-mouse (AP308P) and anti-rabbit (AP307P) IgG were applied. enhanced chemiluminescence reagent method was used for visualizing protein bands corresponding to the proteins of interest.

7.3.4 Microscopy

Cells were seeded on 12-well plates (21000 cells per well) in the standard DMEM medium and left to adhere for 18 h. Cell-containing wells were washed twice with PBS followed by addition of CM or AFM containing a respective concentration of the tested compound. Microphotographies were obtained using Nikon Eclipse Ti-U fluorescent microscope equipped with a 10× objective and DS-Qi2 digital camera, (Tokyo, Japan) after 24, 48 and 72 h of incubation.

7.3.5 Statistical analysis

Experiments were run in technical triplicates. Data is expressed as means ± standard deviation. Statistical analysis was performed using two-sided Student's *t* test in the GraphPad Prism 9.1.1 software (San Diego, CA, USA). Statistical significance was defined as $p < 0.05$.

8. Bibliography

(1) Tapiero, H.; Mathé, G.; Couvreur, P.; Tew, K. D. I. Arginine. Biomed. Pharmacother. 2002, 56 (9), 439–445. https://doi.org/10.1016/S0753-3322(02)00284-6.

(2) Zhou, H.-X.; Pang, X. Electrostatic Interactions in Protein Structure, Folding, Binding, and Condensation. Chem. Rev. 2018, 118 (4), 1691–1741. https://doi.org/10.1021/acs.chemrev.7b00305.

(3) Sokalingam, S.; Raghunathan, G.; Soundrarajan, N.; Lee, S. G. A Study on the Effect of Surface Lysine to Arginine Mutagenesis on Protein Stability and Structure Using Green Fluorescent Protein. PLoS One 2012, 7 (7). https://doi.org/10.1371/journal.pone.0040410.

(4) Saha, S.; Kashina, A. Posttranslational Arginylation as a Global Biological Regulator. 2011. https://doi.org/10.1016/j.ydbio.2011.06.043.

(5) Tanaka, K. The Proteasome: Overview of Structure and Functions. Proc. Japan Acad. Ser. B Phys. Biol. Sci. 2009, 85 (1), 12–36. https://doi.org/10.2183/pjab.85.12.

(6) Wilkinson, K. D. Ubiquitination and Deubiquitination: Targeting of Proteins for Degradation by the Proteasome. Semin. Cell Dev. Biol. 2000, 11 (3), 141–148. https://doi.org/10.1006/scdb.2000.0164.

(7) Tasaki, T.; Sriram, S. M.; Park, K. S.; Kwon, Y. T. The N-End Rule Pathway. https://doi.org/10.1146/annurev-biochem-051710-093308 2012, 81, 261–289.

(8) Kopitz, J.; Rist, B.; Bohley, P. Post-Translational Arginylation of Ornithine Decarboxylase from Rat. Hepatocytes. *Biochem. J* 1990, 267, 343–348.

(9) Eriste, E.; Norberg, Å.; Nepomuceno, D.; Kuei, C.; Kamme, F.; Tran, D. T.; Strupat, K.; Jörnvall, H.; Liu, C.; Lovenberg, T. W.; Sillard, R. A Novel Form of Neurotensin Post-Translationally Modified by Arginylation. J. Biol. Chem. 2005, 280 (42), 35089–35097. https://doi.org/10.1074/JBC.M502567200.

(10) St.-Gelais, F.; Jomphe, C.; Trudeau, L. É. The Role of Neurotensin in Central Nervous System Pathophysiology: What Is the Evidence? J. Psychiatry Neurosci. 2006, 31 (4), 229.

(11) Karakozova, M.; Kozak, M.; Wong, C. C. L.; Bailey, A. O.; Yates, J. R.; Mogilner, A.; Zebroski, H.; Kashina, A. Arginylation of Beta-Actin Regulates Actin Cytoskeleton and Cell Motility. Science 2006, 313 (5784), 192–196. https://doi.org/10.1126/SCIENCE.1129344.

(12) Shambaugh, G. E. Urea Biosynthesis I. The Urea Cycle and Relationships to the Citric Acid Cycle. Am. J. Clin. Nutr. 1977, 30 (12), 2083–2087. https://doi.org/10.1093/AJCN/30.12.2083.

(13) Husson, A.; Brasse-Lagnel, C.; Fairand, A.; Renouf, S.; Lavoinne, A. Argininosuccinate Synthetase from the Urea Cycle to the Citrulline–NO Cycle. Eur. J. Biochem. 2003, 270 (9), 1887–1899. https://doi.org/10.1046/J.1432-1033.2003.03559.X.

(14) Bence, A. K.; Crooks, P. A. The Mechanism of L-Canavanine Cytotoxicity: Arginyl TRNA Synthetase as a Novel Target for Anticancer Drug Discovery. J. Enzyme Inhib. Med. Chem. 2003, 18 (5), 383–394. https://doi.org/10.1080/1475636031000152277.

(15) Sivashanmugam, M.; J., J.; V., U.; K.N., S. Ornithine and Its Role in Metabolic Diseases: An Appraisal. Biomed. Pharmacother. 2017, 86, 185–194. https://doi.org/10.1016/J.BIOPHA.2016.12.024.

(16) Rizk, M.; Witte, M. B.; Barbul, A. Nitric Oxide and Wound Healing. World J. Surg. 2004, 28 (3), 301–306. https://doi.org/10.1007/S00268-003-7396-7/METRICS.

(17) Ginguay, A.; Cynober, L.; Curis, E.; Nicolis, I. Ornithine Aminotransferase, an Important Glutamate-Metabolizing Enzyme at the Crossroads of Multiple Metabolic Pathways. Biology (Basel). 2017, 6 (1), 18. https://doi.org/10.3390/biology6010018.

(18) Pegg, A. E. Mammalian Polyamine Metabolism and Function. IUBMB Life 2009, 61 (9), 880-894. https://doi.org/10.1002/iub.230.

(19) Joncquel-Chevalier Curt, M.; Voicu, P.-M.; Fontaine, M.; Ed Erique Dessein, A.-F.; Porchet, N.; Mention-Mulliez, K.; Dobbelaere, D.; Soto-Ares, G.; Cheillan, D.; Vamecq, J. Creatine Biosynthesis and Transport in Health and Disease. *Biochimie* 2015, 119, 146-165. doi:10.1016/j.biochi.2015.10.022

(20) Brosnan, M. E.; Edison, E. E.; da Silva, R.; Brosnan Ã, J. T. New Insights into Creatine Function and Synthesis. Advan. Enzym. Regul 2007, 47, 252–260. https://doi.org/10.1016/j.advenzreg.2006.12.005.

(21) Sosnowska, A.; Chlebowska-Tuz, J.; Matryba, P.; Pilch, Z.; Greig, A.; Wolny, A.; Grzywa, T. M.; Rydzynska, Z.; Sokolowska, O.; Rygiel, T. P.; Grzybowski, M.; Stanczak, P.; Blaszczyk, R.; Nowis, D.; Golab, J. Inhibition of Arginase Modulates T-Cell Response in the

Tumor Microenvironment of Lung Carcinoma. Oncoimmunology. 2021, 10(1), 1956143. doi:10.1080/2162402X.2021.1956143

(22) Global Cancer Observatory. *Global Cancer Observatory*. https://gco.iarc.fr/today/fact-sheets-cancers (accessed 2023-05-23).

(23) Bradley, C. J.; Yabroff, K. R.; Dahman, B.; Feuer, E. J.; Mariotto, A.; Brown, M. L. Productivity Costs of Cancer Mortality in the United States: 2000-2020. J. Natl. Cancer Inst. 2008, 100 (24), 1763–1770. https://doi.org/10.1093/jnci/djn384.

(24) Ostrom, Q. T.; Francis, S. S.; Barnholtz-Sloan, J. S. Epidemiology of Brain and Other CNS Tumors. Curr. Neurol. Neurosci. Rep. 2021, 21 (12), 68. https://doi.org/10.1007/S11910-021-01152-9.

(25) Fields, R. D.; Araque, A.; Johansen-Berg, H.; Lim, S. S.; Lynch, G.; Nave, K. A.; Nedergaard, M.; Perez, R.; Sejnowski, T.; Wake, H. Glial Biology in Learning and Cognition. Neuroscientist 2014, 20 (5), 426–431. https://doi.org/10.1177/1073858413504465

(26) Bean, B. P. The Action Potential in Mammalian Central Neurons. Nat. Rev. Neurosci. 2007 86 2007, 8 (6), 451–465. https://doi.org/10.1038/nrn2148.

(27) Stummer, W. Surgical Management of Glial Cancers. Emerg. Concepts Neuro-Oncology 2013, 9780857294586, 143–159. https://doi.org/10.1007/978-0-85729-458-6_9

(28) Pilanc, P.; Wojnicki, K.; Roura, A. J.; Cyranowski, S.; Ellert-Miklaszewska, A.; Ochocka, N.; Gielniewski, B.; Grzybowski, M. M.; Błaszczyk, R.; Stańczak, P. S.; Dobrzański, P.; Kaminska, B. A Novel Oral Arginase 1/2 Inhibitor Enhances the Antitumor Effect of PD-1 Inhibition in Murine Experimental Gliomas by Altering the Immunosuppressive Environment. Front. Oncol. 2021, 11, 703465. doi:10.3389/fonc.2021.703465.

(29) R&D: Molecure. https://molecure.com/pipeline/ (accessed 2024-02-01).

(30) Szefel, J.; Danielak, A.; Kruszewski, W. J. Metabolic Pathways of L-Arginine and Therapeutic Consequences in Tumors. Adv. Med. Sci. 2019, 64 (1), 104–110. https://doi.org/10.1016/J.ADVMS.2018.08.018.

(31) Phillips, M. M.; Sheaff, M. T.; Szlosarek, P. W. Targeting Arginine-Dependent Cancers with Arginine-Degrading Enzymes: Opportunities and Challenges. Cancer Res, Treat. 2013;45(4):251-262. doi:10.4143/crt.2013.45.4.251.

(32) Ascierto, P. A.; Scala, S.; Castello, G.; Daponte, A.; Simeone, E.; Ottaiano, A.; Beneduce, G.; De Rosa, V.; Izzo, F.; Melucci, M. T.; Ensor, C. M.; Prestayko, A. W.; Holtsberg, F. W.; Bomalaski, J. S.; Clark, M. A.; Savaraj, N.; Feun, L. G.; Logan, T. F. Pegylated Arginine Deiminase Treatment of Patients with Metastatic Melanoma: Results from Phase I and II Studies. J. Clin. Oncol. 2005, 23 (30), 7660–7668. https://doi.org/10.1200/JCO.2005.02.0933.

(33) Izzo, F.; Marra, P.; Beneduce, G.; Castello, G.; Vallone, P.; De Rosa, V.; Cremona, F.; Ensor, C. M.; Holtsberg, F. W.; Bomalaski, J. S.; Clark, M. A.; Ng, C.; Curley, S. A. Pegylated Arginine Deiminase Treatment of Patients with Unresectable Hepatocellular Carcinoma: Results from Phase I/II Studies. J. Clin. Oncol. 2004, 22 (10), 1815–1822. https://doi.org/10.1200/JCO.2004.11.120.

(34) Ott, P. A.; Carvajal, R. D.; Pandit-Taskar, N.; Jungbluth, A. A.; Hoffman, E. W.; Wu, B. W.; Bomalaski, J. S.; Venhaus, R.; Pan, L.; Old, L. J.; Pavlick, A. C.; Wolchok, J. D. Phase I/II Study of Pegylated Arginine Deiminase (ADI-PEG 20) in Patients with Advanced Melanoma. Invest. New Drugs 2013, 31 (2), 425–434. https://doi.org/10.1007/S10637-012-9862-2.

(35) Yang, T. S.; Lu, S. N.; Chao, Y.; Sheen, I. S.; Lin, C. C.; Wang, T. E.; Chen, S. C.; Wang, J. H.; Liao, L. Y.; Thomson, J. A.; Wang-Peng, J.; Chen, P. J.; Chen, L. T. A Randomised Phase II Study of Pegylated Arginine Deiminase (ADI-PEG 20) in Asian Advanced Hepatocellular Carcinoma Patients. Br. J. Cancer 2010 1037 2010, 103 (7), 954–960. https://doi.org/10.1038/sj.bjc.6605856.

(36) Pavlyk, I.; Rzhepetskyy, Y.; Jagielski, A. K.; Drozak, J.; Wasik, A.; Pereverzieva, G.; Olchowik, M.; Kunz-Schugart, L. A.; Stasyk, O.; Redowicz, M. J. Arginine Deprivation Affects Glioblastoma Cell Adhesion, Invasiveness and Actin Cytoskeleton Organization by Impairment of β-Actin Arginylation. Amino Acids 2015, 47 (1), 199–212. https://doi.org/10.1007/S00726-014-1857-1.

(37) Karatsai, O.; Shliaha, P.; Jensen, O. N.; Stasyk, O.; Rędowicz, M. J. Combinatory Treatment of Canavanine and Arginine Deprivation Efficiently Targets Human Glioblastoma Cells via Pleiotropic Mechanisms. Cells 2020, 9 (10), 2217. doi:10.3390/cells9102217.

(38) Gabarra-Niecko, V.; Schaller, M. D.; Dunty, J. M. FAK Regulates Biological Processes Important for the Pathogenesis of Cancer. Cancer Metastasis Rev. 2003, 22 (4), 359–374. https://doi.org/10.1023/A:1023725029589.

(39) Dawson, J. C.; Serrels, A.; Stupack, D. G.; Schlaepfer, D. D.; Frame, M. C. Targeting FAK in Anticancer Combination Therapies. Nat. Rev. Cancer 2021 215 2021, 21 (5), 313–324. https://doi.org/10.1038/s41568-021-00340-6.

(40) Shuvayeva, G. Y.; Bobak, Y. P.; Vovk, O. I.; Kunz-Schughart, L. A.; Fletcher, M. T.; Stasyk, O. V. Indospicine Combined with Arginine Deprivation Triggers Cancer Cell Death via Caspase-Dependent Apoptosis. Cell Biol. Int. 2021, 45 (3), 518–527. https://doi.org/10.1002/CBIN.11321.

(41) Bence, A. K.; Crooks, P. A. The Mechanism of L-Canavanine Cytotoxicity: Arginyl TRNA Synthetase as a Novel Target for Anticancer Drug Discovery. 2008, 18 (5), 383–394. https://doi.org/10.1080/1475636031000152277.

(42) Fletcher, M. T.; Tan, E. T. T.; Silcock, R. G. Indigofera Linnaei-a Problem Legume across Northern Pastures. In *Joint ISNH/ISRP International Conference 2014: Harnessing the Ecology and Physiology of Herbivores*; Canberra September 8-12, 2014; Vol. 30, pp 285–285.

(43) Chen, Q.; Kang, J.; Fu, C. The Independence of and Associations among Apoptosis, Autophagy, and Necrosis. Signal Transduct. Target. Ther. 2018 31 2018, 3 (1), 1–11. https://doi.org/10.1038/s41392-018-0018-5.

(44) Hitomi, J.; Christofferson, D. E.; Ng, A.; Yao, J.; Degterev, A.; Xavier, R. J.; Yuan, J. Identification of a Molecular Signaling Network That Regulates a Cellular Necrotic Cell Death Pathway. Cell 2008, 135 (7), 1311–1323. https://doi.org/10.1016/J.CELL.2008.10.044.

(45) Kono, H.; Rock, K. L. How Dying Cells Alert the Immune System to Danger. Nat. Rev. Immunol. 2008, 8 (4), 279–289. https://doi.org/10.1038/nri2215.

(46) Ichim, G.; Tait, S. W. G. A Fate Worse than Death: Apoptosis as an Oncogenic Process. Nat. Rev. Cancer 2016 168 2016, 16 (8), 539–548. https://doi.org/10.1038/nrc.2016.58.

(47) Atkin-Smith, G. K.; Poon, I. K. H. Disassembly of the Dying: Mechanisms and Functions. Trends Cell Biol. 2017, 27 (2), 151–162. https://doi.org/10.1016/j.tcb.2016.08.011.

(48) McIlwain, D. R.; Berger, T.; Mak, T. W. Caspase Functions in Cell Death and Disease. Cold Spring Harb. Perspect. Biol. 2013, 5 (4), 1–28. https://doi.org/10.1101/cshperspect.a008656.

(49) Los, M.; Mozoluk, M.; Ferrari, D.; Stepczynska, A.; Stroh, C.; Renz, A.; Herceg, Z.; Wang, Z. Q.; Schulze-Osthoff, K. Activation and Caspase-Mediated Inhibition of PARP: A

Molecular Switch between Fibroblast Necrosis and Apoptosis in Death Receptor Signaling. Mol. Biol. Cell 2002, 13 (3), 978-988. https://doi.org/10.1091/mbc.01-05-0272.

(50) Smith, H. J.; Williams, H. J. Textbook of Drug Design and Discovery, 5th ed.; CRC Press, 2017. https://doi.org/10.1201/b12381.

(51) Allen, W. J.; Balius, T. E.; Mukherjee, S.; Brozell, S. R.; Moustakas, D. T.; Lang, P. T.; Case, D. A.; Kuntz, I. D.; Rizzo, R. C. DOCK 6: Impact of New Features and Current Docking Performance. J. Comput. Chem. 2015, 36 (15), 1132–1156. https://doi.org/10.1002/jcc.23905.

(52) Kuntz, I. D.; Blaney, J. M.; Oatley, S. J.; Langridge, R.; Ferrin, T. E. A Geometric Approach to Macromolecule-Ligand Interactions. J. Mol. Biol. 1982, 161 (2), 269–288. https://doi.org/10.1016/0022-2836(82)90153-x.

(53) Trott, O.; Olson, A. J. AutoDock Vina: Improving the Speed and Accuracy of Docking with a New Scoring Function, Efficient Optimization, and Multithreading. J. Comput. Chem. 2010, 31 (2), 455–461. https://doi.org/10.1002/jcc.21334.

(54) Home | World Community Grid. https://www.worldcommunitygrid.org/ (accessed 2023-03-25).

(55) Isidro-Llobet, A.; Álvarez, M.; Albericio, F. Amino Acid-Protecting Groups. Chem. Rev. 2009, 109 (6), 2455–2504. doi:10.1021/cr800323s.

(56) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. Click Chemistry: Diverse Chemical Function from a Few Good Reactions. Angew. Chemie Int. Ed. 2001, 40 (11), 2004–2021. https://doi.org/10.1002/1521-3773(20010601)40:11<2004::AID-ANIE2004>3.0.CO;2-5.

(57) Haldón, E.; Nicasio, M. C.; Pérez, P. J. Copper-Catalysed Azide–Alkyne Cycloadditions (CuAAC): An Update. Org. Biomol. Chem. 2015, 13 (37), 9528–9550. https://doi.org/10.1039/C5OB01457C.

(58) Kafle, A.; Handy, S. T. A One-Pot, Copper-Catalyzed Azidation/Click Reaction of Aryl and Heteroaryl Bromides in an Environmentally Friendly Deep Eutectic Solvent. Tetrahedron 2017, 73 (50), 7024–7029. https://doi.org/10.1016/J.TET.2017.10.050.

(59) Nwe, K.; Brechbiel, M. W. Growing Applications of “Click Chemistry” for Bioconjugation in Contemporary Biomedical Research. Cancer Biother. Radiopharm. 2009, 24 (3), 289. https://doi.org/10.1089/CBR.2008.0626.

(60) Kaur, J.; Saxena, M.; Rishi, N. An Overview of Recent Advances in Biomedical Applications of Click Chemistry. Bioconjug. Chem. 2021, 32 (8), 1455–1471. doi:10.1021/acs.bioconjchem.1c00247

(61) Fantoni, N. Z.; El-Sagheer, A. H.; Brown, T. A Hitchhiker's Guide to Click-Chemistry with Nucleic Acids. Chem. Rev. 2021, 121 (12), 7122–7154. doi:10.1021/acs.chemrev.0c00928

(62) Press release: The Nobel Prize in Chemistry 2022 - NobelPrize.org. https://www.nobelprize.org/prizes/chemistry/2022/press-release/ (accessed 2023-03-25).

(63) Moss, G. P.; Smith, P. A. S.; Tavernier, D. Glossary of Class Names of Organic Compounds and Reactive Intermediates Based on Structure (IUPAC Recommendations 1995). Pure Appl. Chem. 1995, 67 (8–9), 1307–1375. http://dx.doi.org/10.1351/pac199567081307

(64) Nand, B.; Khanna, G.; Chaudhary, A.; Lumb, A.; M. Khurana, J. 1,8-Diazabicyclo[5.4.0]Undec-7-Ene (DBU): A Versatile Reagent in Organic Synthesis. Curr. Org. Chem. 2015, 19 (9), 790–812. https://doi.org/10.2174/1385272819666150402221133.

(65) Pinner, A.; Klein, F. Umwandlung Der Nitrile in Imide. Berichte der Dtsch. Chem. Gesellschaft 1877, 10 (2), 1889–1897. https://doi.org/10.1002/cber.187701002154.

(66) Li, J. J. Pinner Synthesis. In *Name Reactions*; Springer, Berlin, Heidelberg, 2002; pp 285–285. https://doi.org/10.1007/978-3-662-04835-1_225.

(67) Clark, M. J.; Homer, N.; O'Connor, B. D.; Chen, Z.; Eskin, A.; Lee, H.; Merriman, B.; Nelson, S. F. U87MG Decoded: The Genomic Sequence of a Cytogenetically Aberrant Human Cancer Cell Line. PLoS Genet. 2010, 6 (1), 1000832. https://doi.org/10.1371/journal.pgen.1007392.

(68) Torsvik, A.; Stieber, D.; Enger, P. O.; Golebiewska, A.; Molven, A.; Svendsen, A.; Westermark, B.; Niclou, S. P.; Olsen, T. K.; Chekenya Enger, M.; Bjerkvig, R. U-251 Revisited: Genetic Drift and Phenotypic Consequences of Long-Term Cultures of Glioblastoma Cells. Cancer Med. 2014, 3 (4), 812–824. https://doi.org/10.1002/cam4.219.

(69) Towbin, H.; Staehelin, T.; Gordon, J. Electrophoretic Transfer of Proteins from Polyacrylamide Gels to Nitrocellulose Sheets: Procedure and Some Applications. Proc. Natl. Acad. Sci. U. S. A. 1979, 76 (9), 4350–4354. https://doi.org/10.1073/pnas.76.9.4350.

(70) Western Blotting Technique | Rockland. https://www.rockland.com/resources/western-blotting-technique/ (accessed 2023-03-25).

(71) 2.1. Light microscopy | MOOC: Instrumental analysis of cultural heritage objects. https://sisu.ut.ee/heritage-analysis/21-light-microscopy (accessed 2023-03-25).

(72) Harris, T. K.; Mildvan, A. S. High-Precision Measurement of Hydrogen Bond Lengths in Proteins by Nuclear Magnetic Resonance Methods. https://doi.org/10.1002/(SICI)1097-0134(19990515)35:3.

(73) Johansson, H.; Pedersen, D. S. Azide- and Alkyne-Derivatised α-Amino Acids. European J. Org. Chem. 2012, No. 23, 4267–4281. https://doi.org/10.1002/EJOC.201200496.

(74) Roth, S.; Thomas, N. R. A Concise Route to L-Azidoamino Acids: L-Azidoalanine, L-Azidohomo-Alanine and L-Azidonorvaline. Synlett 2010, No. 4, 607–609. https://doi.org/10.1055/S-0029-1218391.

(75) Friscourt, F.; Fahrni, C. J.; Boons, G. J. A Fluorogenic Probe for the Catalyst-Free Detection of Azide-Tagged Molecules. J. Am. Chem. Soc. 2012, 134 (45), 18809–18815. https://doi.org/10.1021/JA309000S.

(76) Sakai, M.; Hashimoto, K.; Shirahama, H. Synthesis of Optically Pure β-Phenylselenoalanine through Serine-β-Lactone: A Useful Precursor of Dehydroalanine. Heterocycles 1997, 44 (1), 319–324. https://doi.org/10.3987/COM-96-S22.

(77) Ramesh, R.; De, K.; Chandrasekaran, S. An Efficient Synthesis of Dehydroamino Acids and Dehydropeptides from O-Cbz and O-Eoc Derivatives of Serine and Threonine. Tetrahedron 2007, 63 (42), 10534–10542. https://doi.org/10.1016/J.TET.2007.07.094.

(78) Meléndez, R. E.; Lubell, W. D. Synthesis and Reactivity of Cyclic Sulfamidites and Sulfamidates. Tetrahedron 2003, 59 (15), 2581–2616. https://doi.org/10.1016/S0040-4020(03)00284-9.

(79) Malhotra, R.; Dey, T. K.; Dutta, S.; Basu, S.; Hajra, S. Efficient Asymmetric Synthesis of N-Protected-β-Aryloxyamino Acids via Regioselective Ring Opening of Serine Sulfamidate Carboxylic Acid. Org. Biomol. Chem. 2014, 12 (33), 6507–6515. https://doi.org/10.1039/C4OB01047G.

(80) Cohen, S. B.; Halcomb, R. L. Application of Serine- and Threonine-Derived Cyclic Sulfamidates for the Preparation of S-Linked Glycosyl Amino Acids in Solution- and Solid-

Phase Peptide Synthesis. J. Am. Chem. Soc. 2002, 124 (11), 2534–2543. https://doi.org/10.1021/ja011932l.

(81) Okuro, K.; Sasaki, M.; Aida, T. Boronic Acid-Appended Molecular Glues for ATP-Responsive Activity Modulation of Enzymes. J. Am. Chem. Soc. 2016, 138 (17), 5527–5530. https://doi.org/10.1021/jacs.6b02664.

(82) Mindt, T. L.; Schibli, R. Triazole Containing Metal Chelating Agents. WO 2007/147482 A1, 2007.

(83) Humphrey, R. E.; McCrary, A. L.; Webb, R. M. Reduction of Alkyl Disulphides with Triphenylphosphine. Talanta 1965, 12 (8), 727–732. https://doi.org/10.1016/0039-9140(65)80108-4.

(84) MTS Cell Proliferation Assay | BioRender Science Templates. https://www.biorender.com/template/mts-cell-proliferation-assay (accessed 2024-01-15).

(85) Thomas, D. A.; Rosenthal, G. A. Metabolism of L-[Guanidinooxy-14C]Canavanine in the Rat. Toxicol. Appl. Pharmacol. 1987, 91 (3), 406–414. https://doi.org/10.1016/0041-008X(87)90062-7.

(86) Closs, E. I., Simon, A., Vékony, N., & Rotmann, A. Plasma membrane transporters for arginine. The Journal of Nutrition, 2004, 134(10 Suppl), 2752S-2759S; discussion 2765S-2767S. https://doi.org/10.1093/jn/134.10.2752S

(87) D'Mello, J.P.F. (Ed.) Adverse effects In *Amino Acids in Human Nutrition and Health*; CAB International: Wallingford, 2012. pp 340

(88) Kim, H. S.; Cha, S. Y.; Jo, C. H.; Han, A.; Hwang, K. Y. The Crystal Structure of Arginyl-TRNA Synthetase from Homo Sapiens. FEBS Lett. 2014, 588 (14), 2328–2334. https://doi.org/10.1016/j.febslet.2014.05.027.

(89) Denoël, T.; Zervosen, A.; Gerards, T.; Lemaire, C.; Joris, B.; Blanot, D.; Luxen, A. Stereoselective Synthesis of Lanthionine Derivatives in Aqueous Solution and Their Incorporation into the Peptidoglycan of Escherichia Coli. Bioorg. Med. Chem. 2014, 22 (17), 4621–4628. https://doi.org/10.1016/j.bmc.2014.07.023.

(90) Busnel, O.; Carreaux, F.; Carboni, B.; Pethe, S.; Goff, S. V. Le; Mansuy, D.; Boucher, J. L. Synthesis and Evaluation of New ω-Borono-α-Amino Acids as Rat Liver Arginase

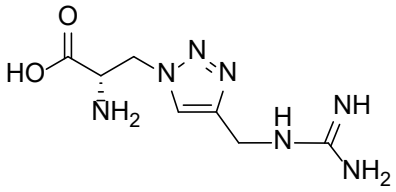
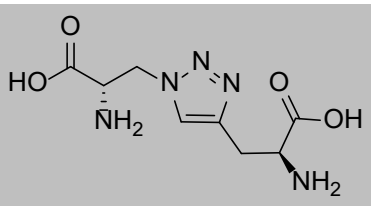
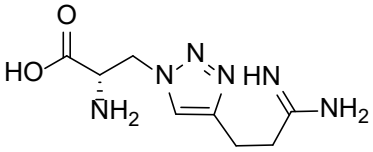
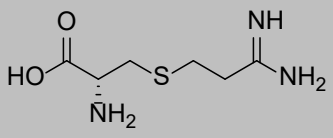
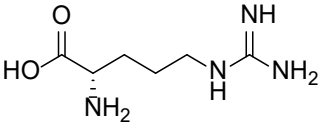
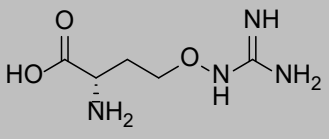
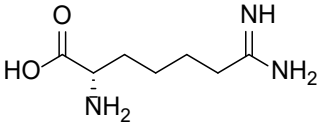
Inhibitors. Bioorg. Med. Chem. 2005, 13 (7), 2373–2379. https://doi.org/10.1016/j.bmc.2005.01.053.

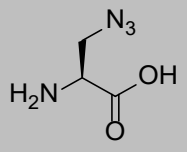
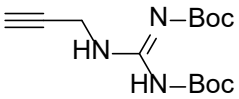
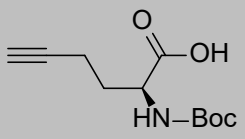
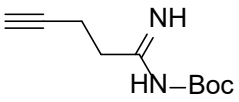
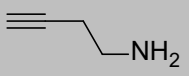
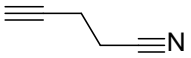
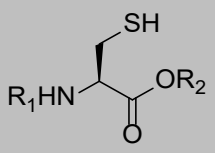
(91) Aversa, M. C.; Barattucci, A.; Bonaccorsi, P.; Giannetto, P. L-Cysteine, a Versatile Source of Sulfenic Acids. Synthesis of Enantiopure Alliin Analogues. J. Org. Chem. 2005, 70 (6), 1986–1992. https://doi.org/10.1021/jO048662k.

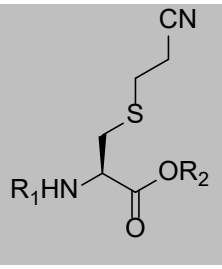
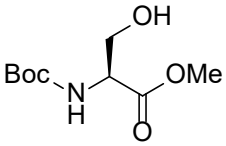
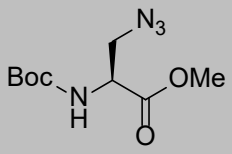
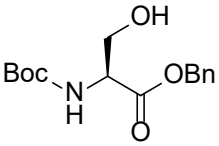
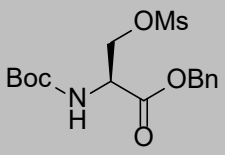
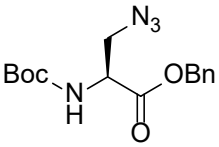
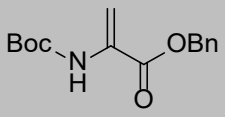
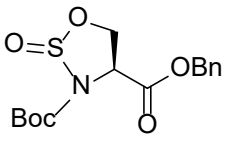
(92) Corporation, P. CellTiter 96 ® AQ Ueous One Solution Cell Proliferation Assay Instructions for Use of Products G3580, G3581 and G3582.

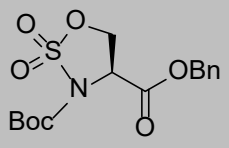
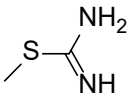
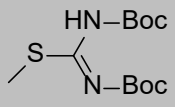
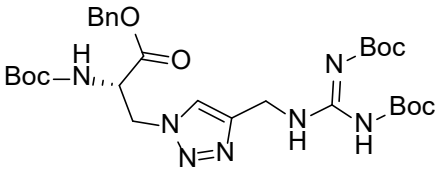
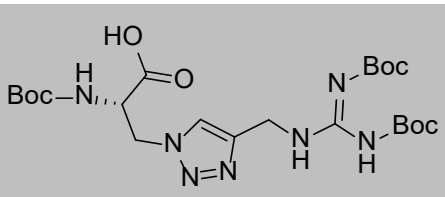
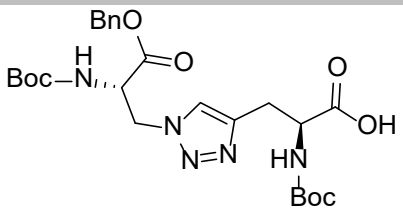
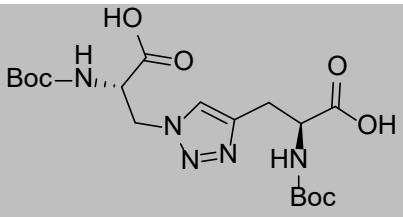
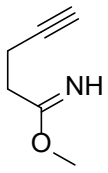
8. Appendices

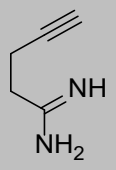
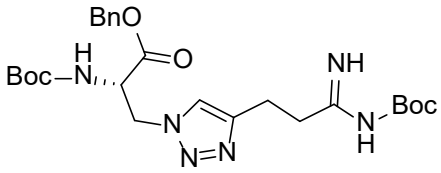
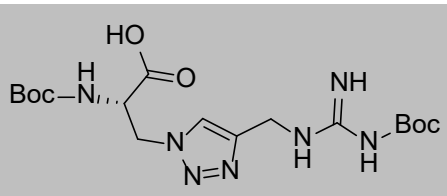
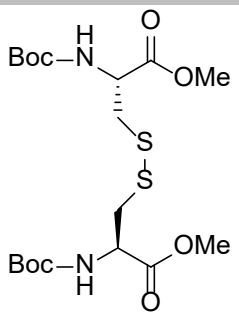
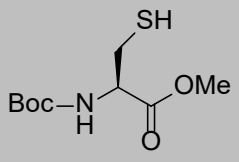
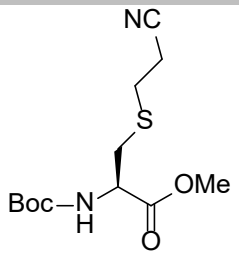
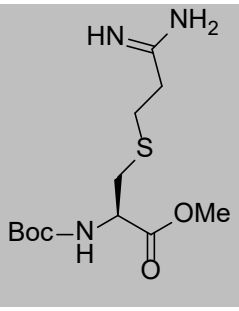
8.1 Important structures

1a	
1b	
1c	
2a	
2c (L-Arginine)	
2h (L-Canavanine)	
2l (L-Indospicine)	

4	
5	
6	
7	
8	
9	
10	
11	

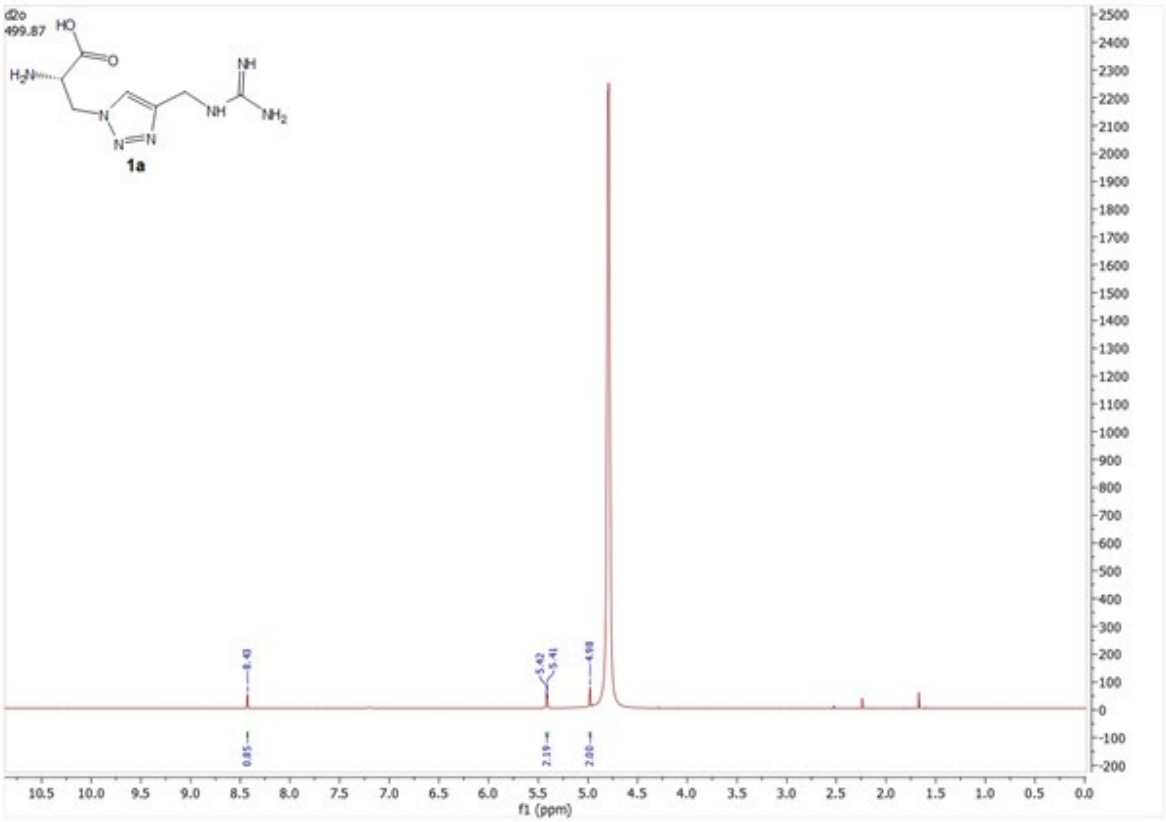
12	
13	
14	
15	
16	
17	
18	
19	

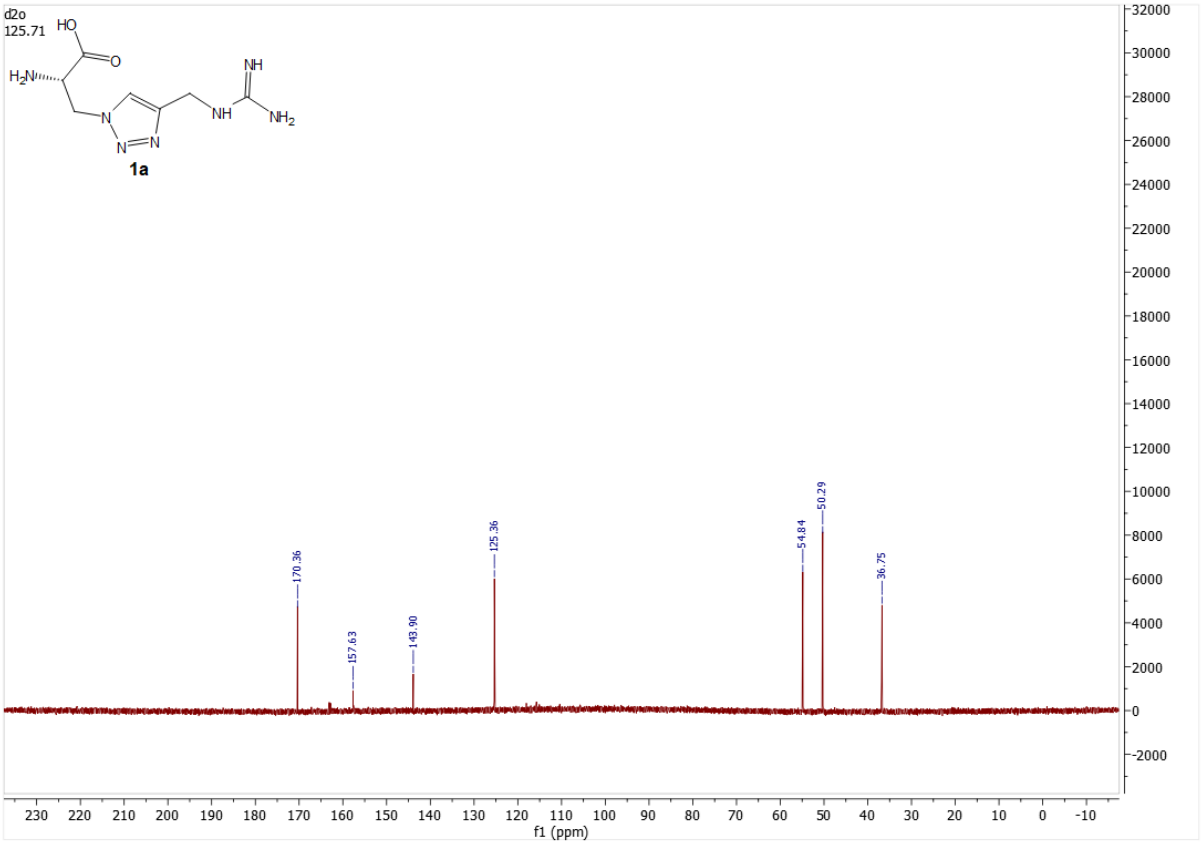
20	
21	
22	
24	
25	
26	
27	
28	

29	
30	
31	
32	
33	
34	
35	

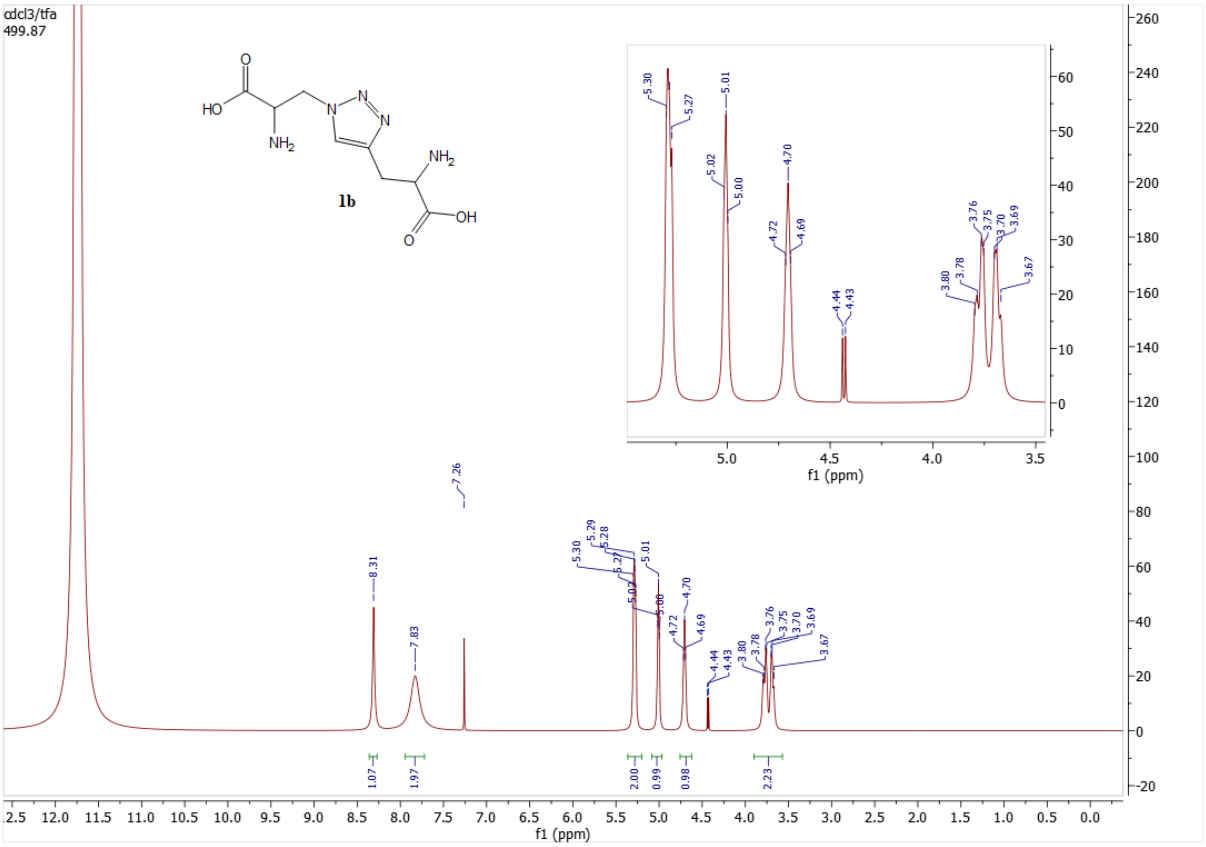
8.2 NMRs

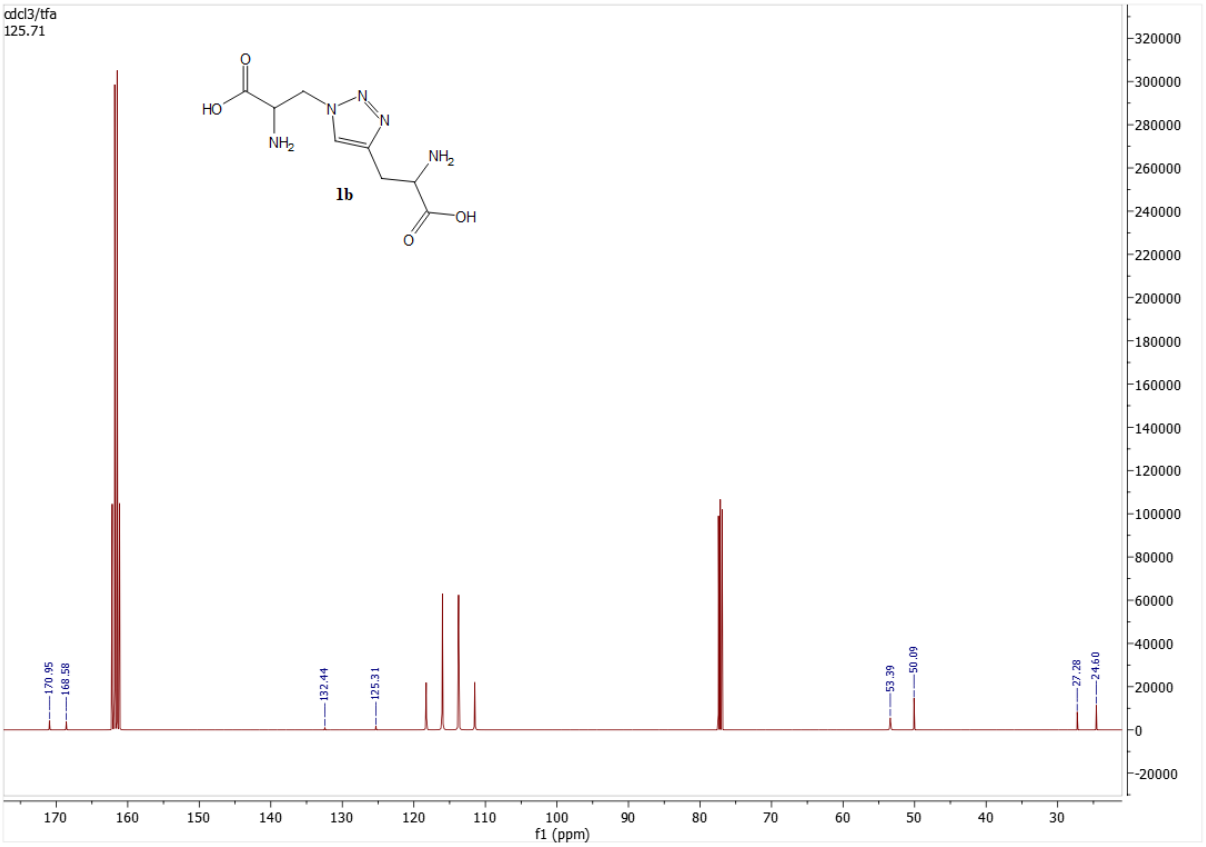
8.2.1 NMR analysis results of compound 1a



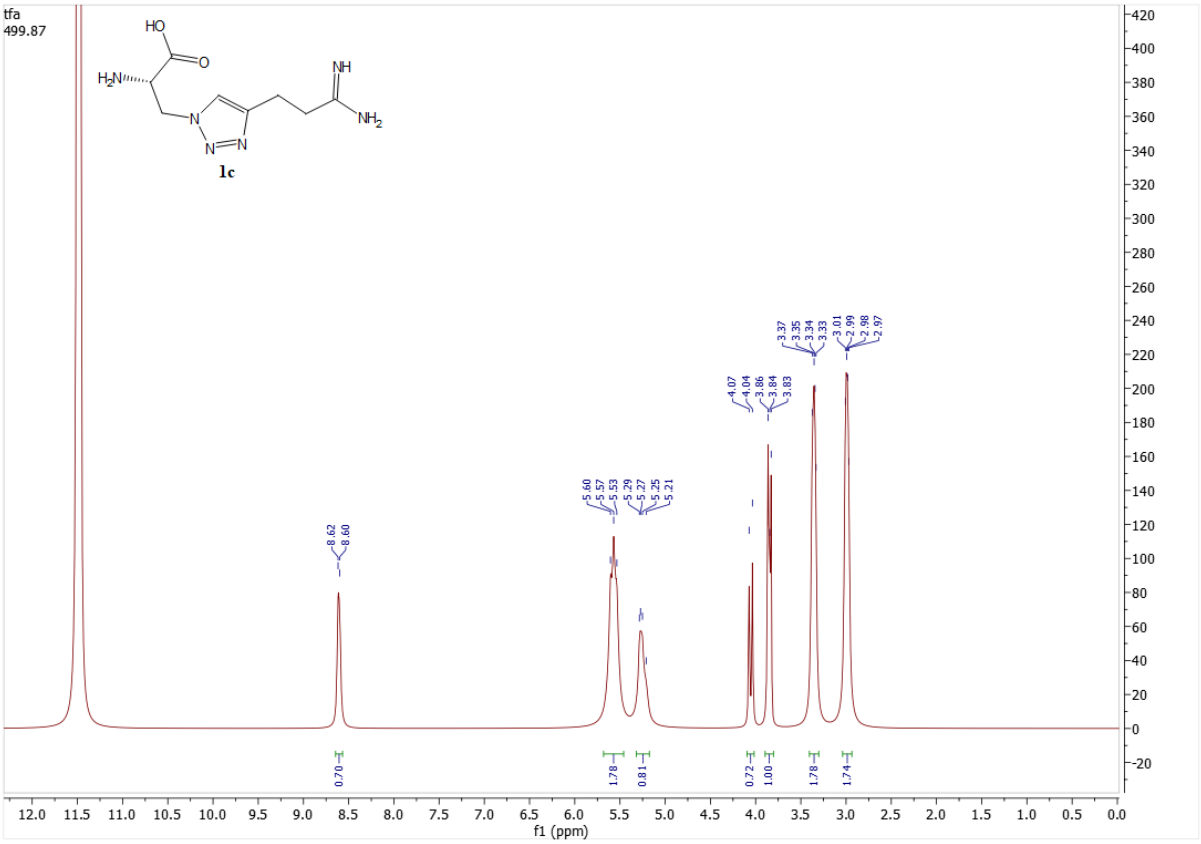


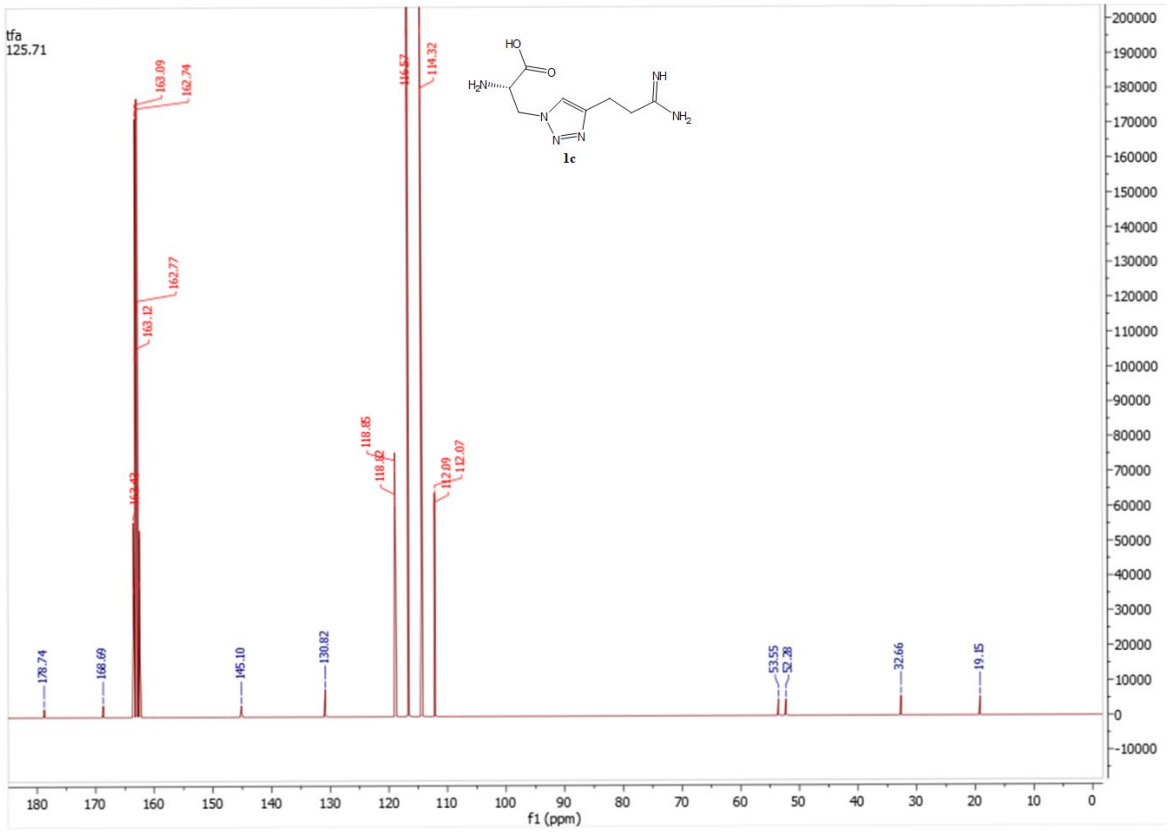
8.2.2 NMR analysis results of compound 1b



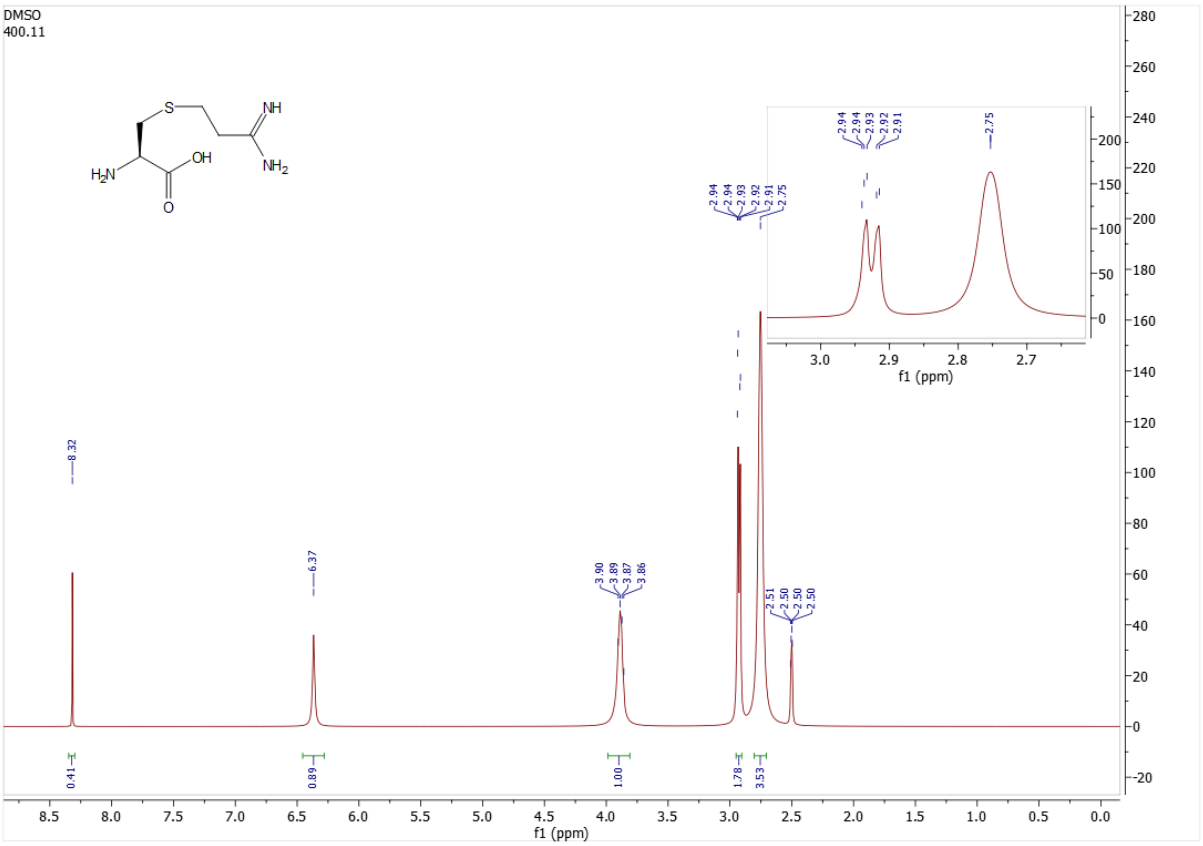


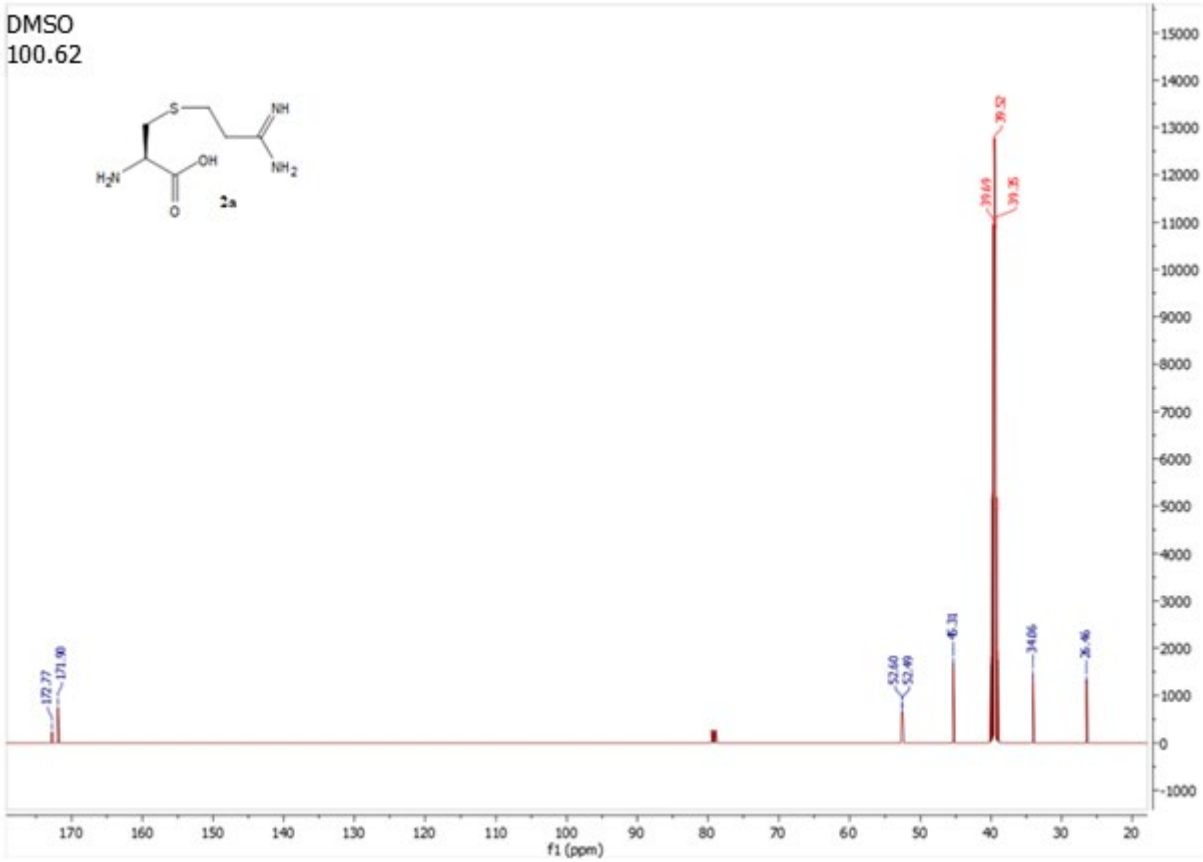
8.2.3 NMR analysis results of compound 1c



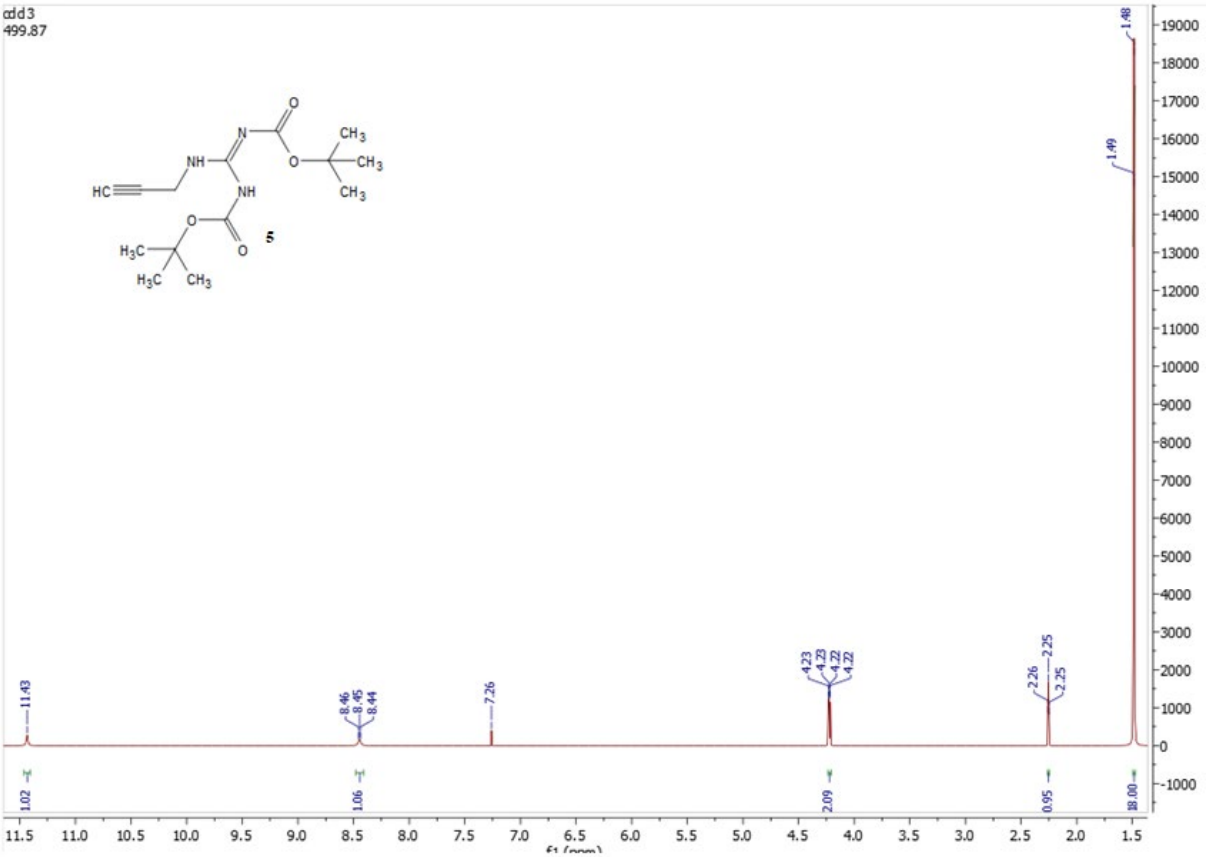


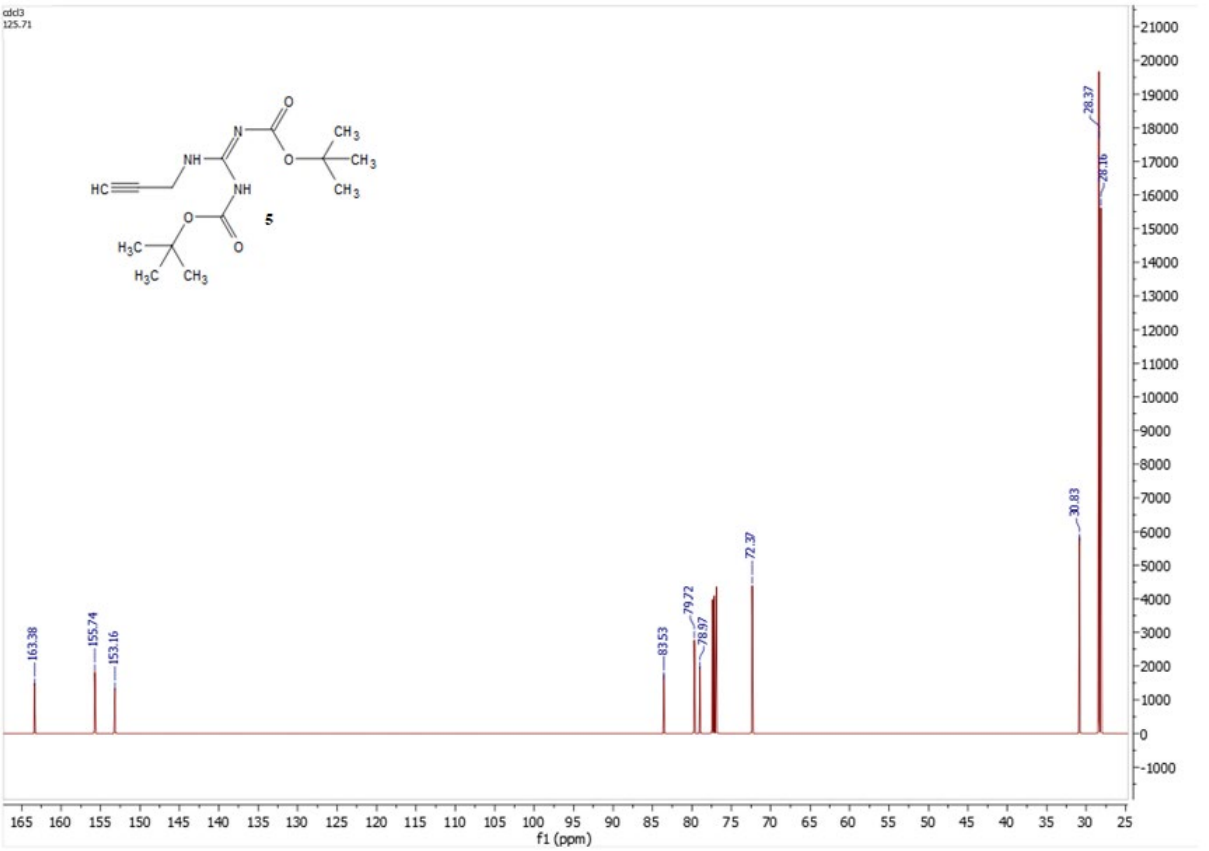
8.2.4 NMR analysis results of compound 2a



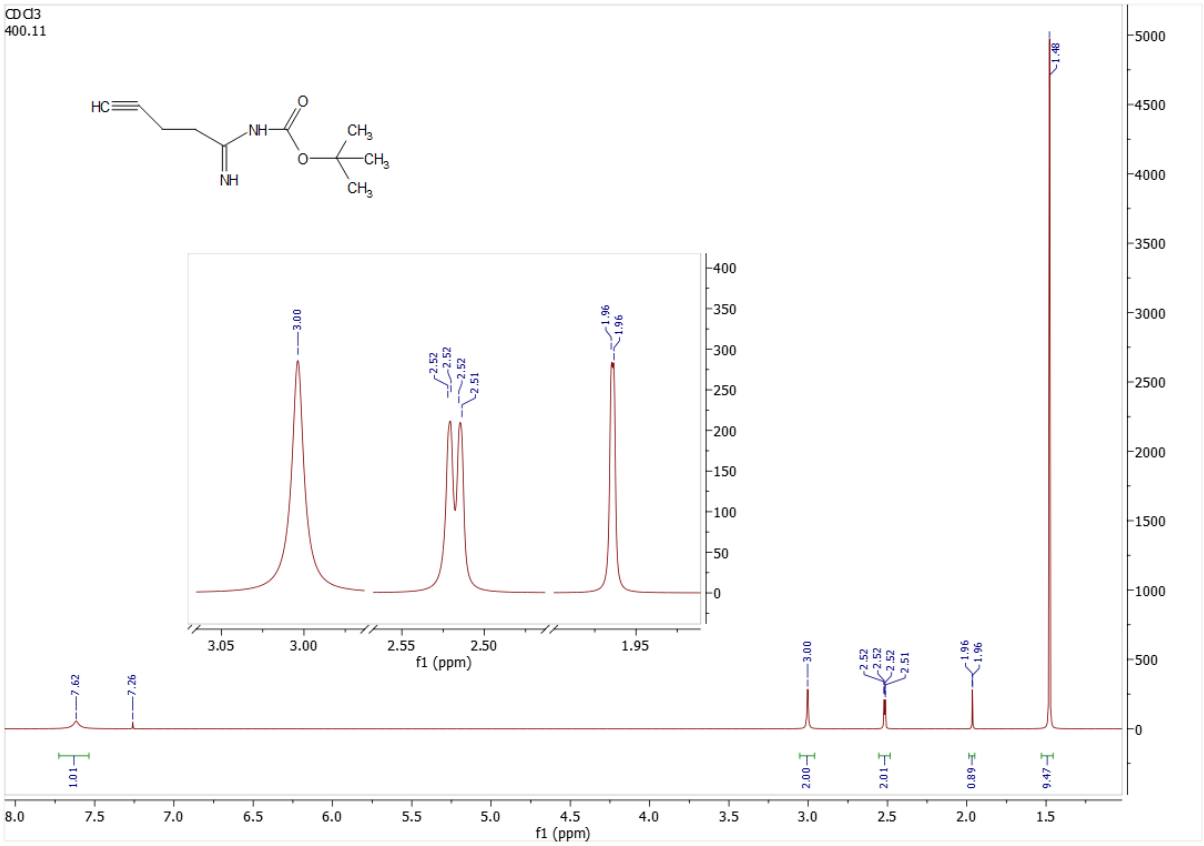


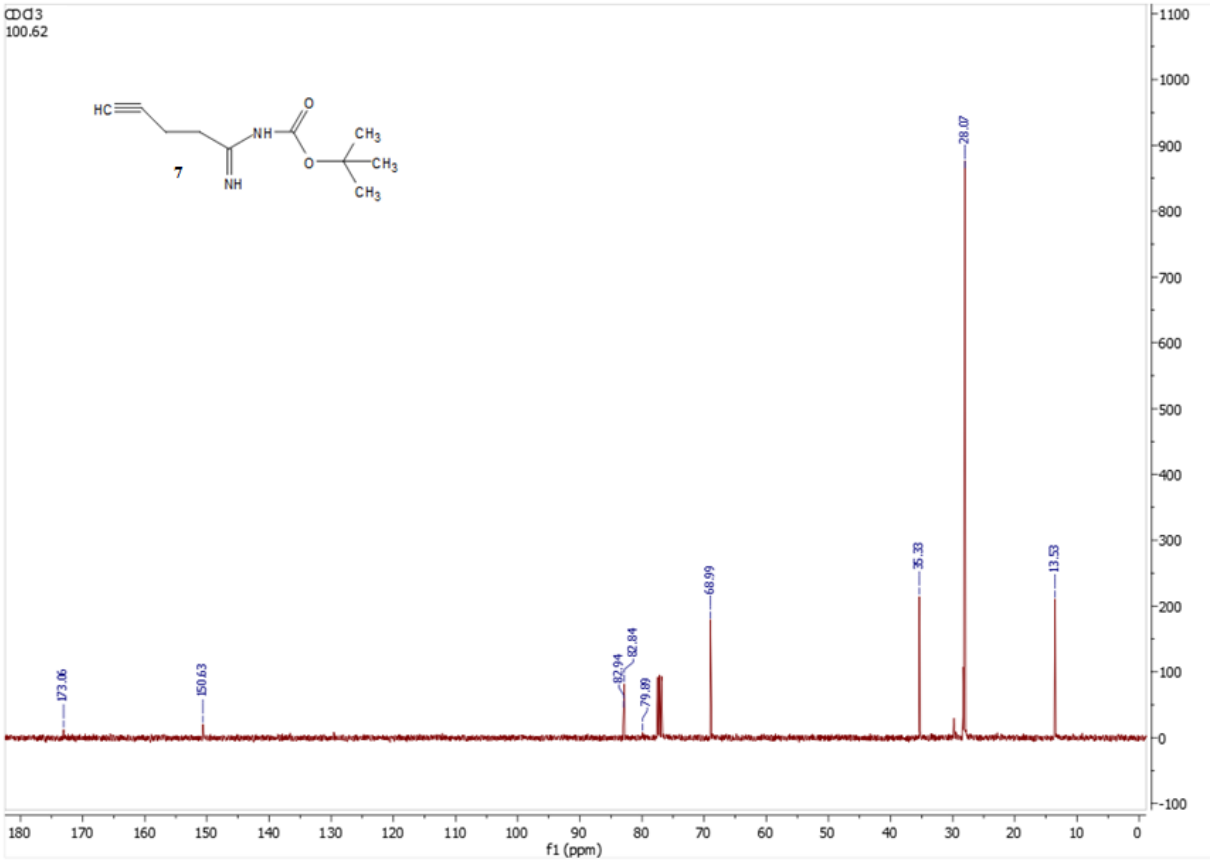
8.2.5 NMR analysis results of compound 5



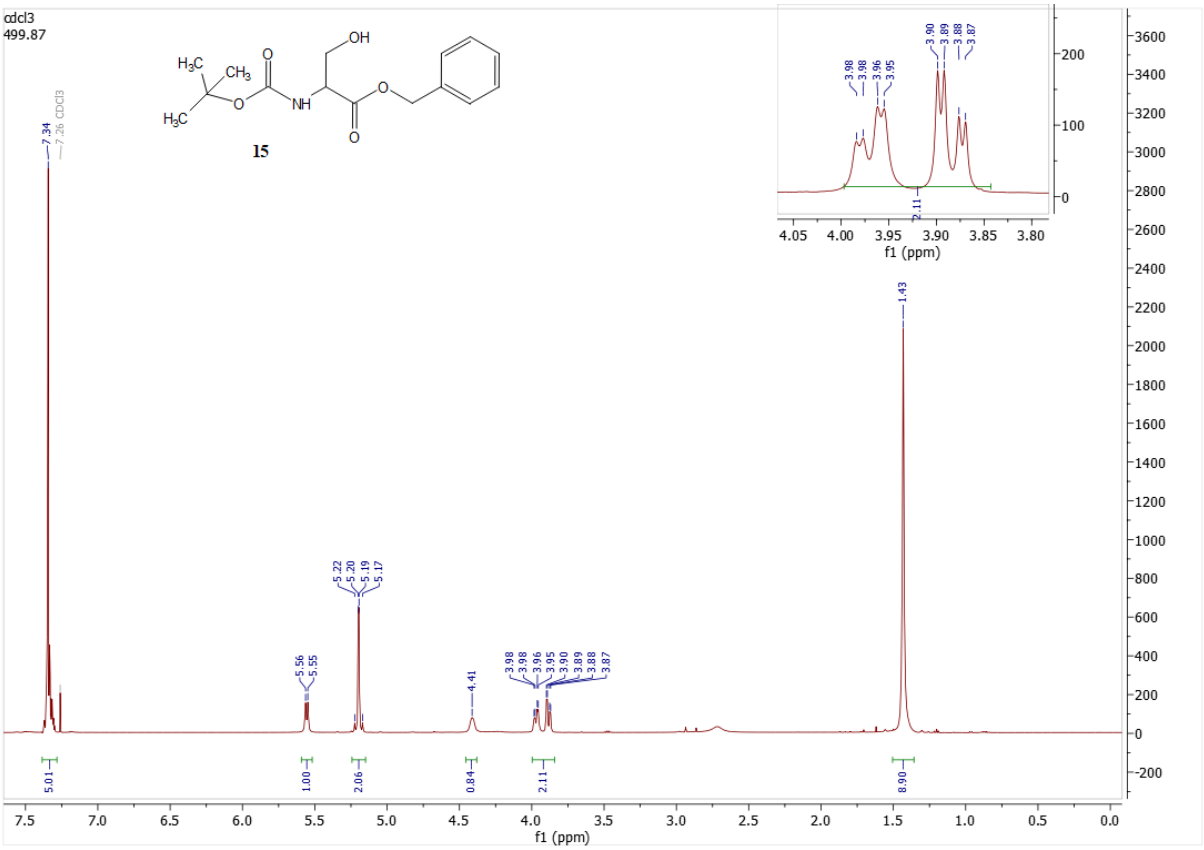


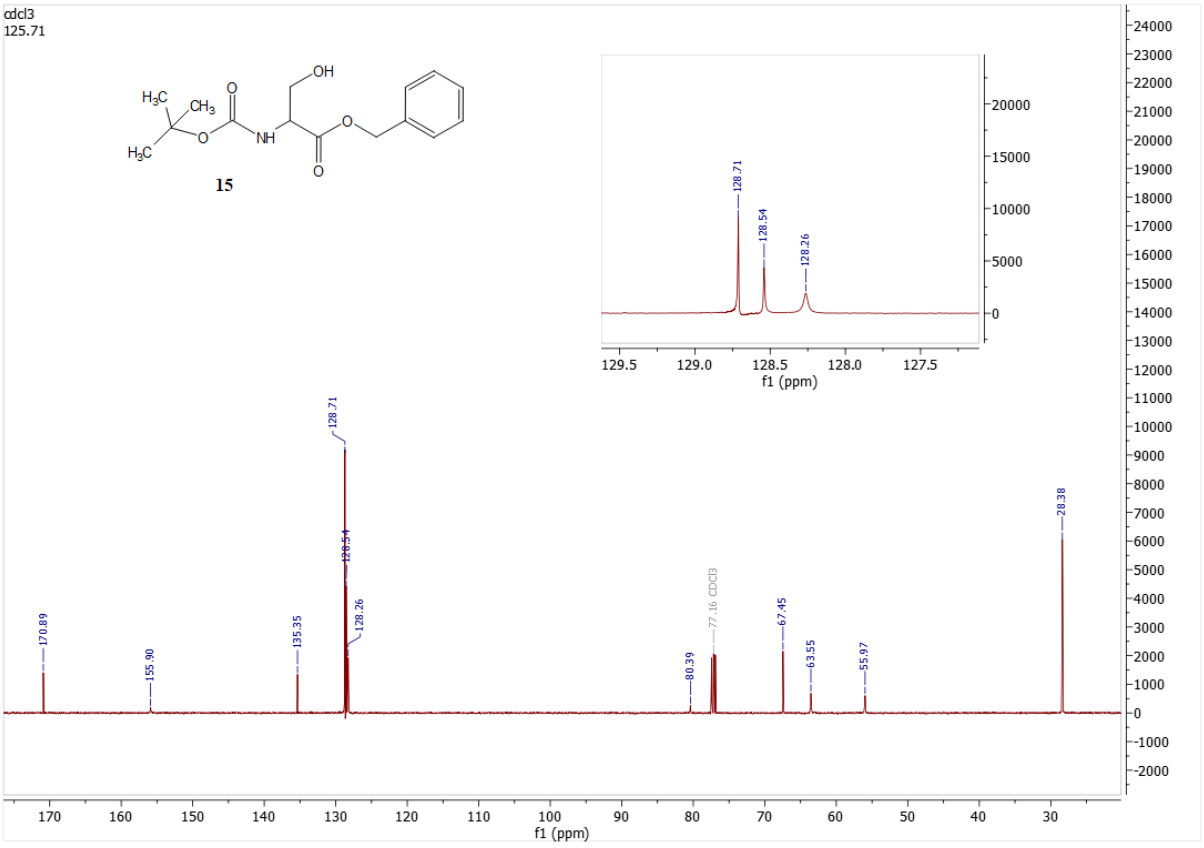
8.2.6 NMR analysis results of compound 7



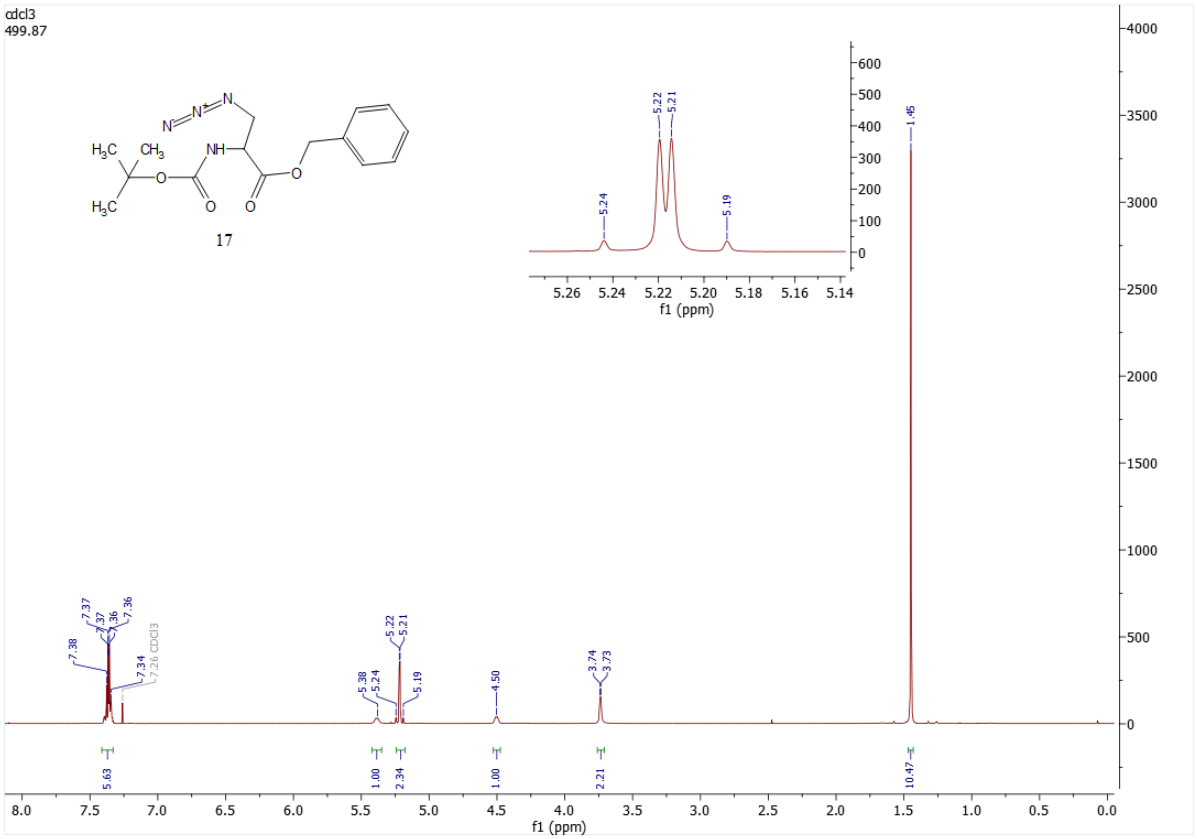


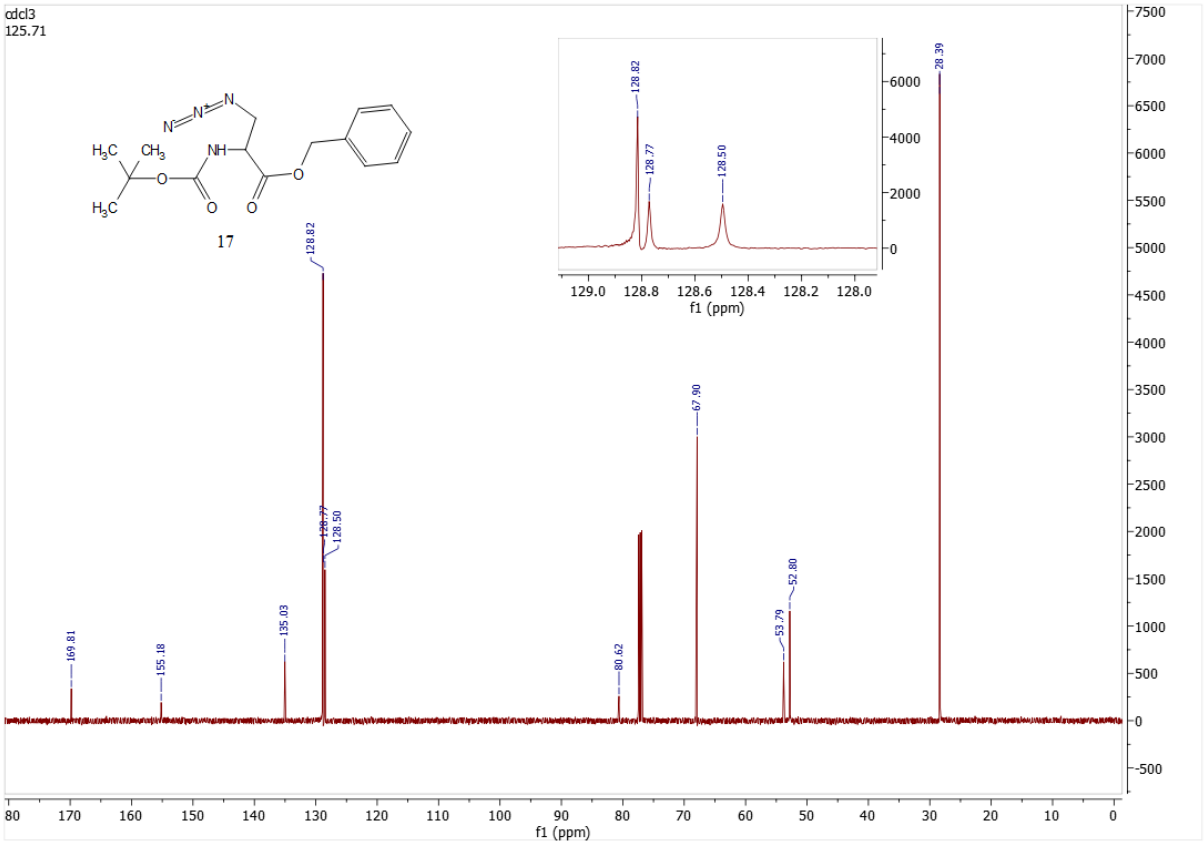
8.2.7 NMR analysis results of compound 15



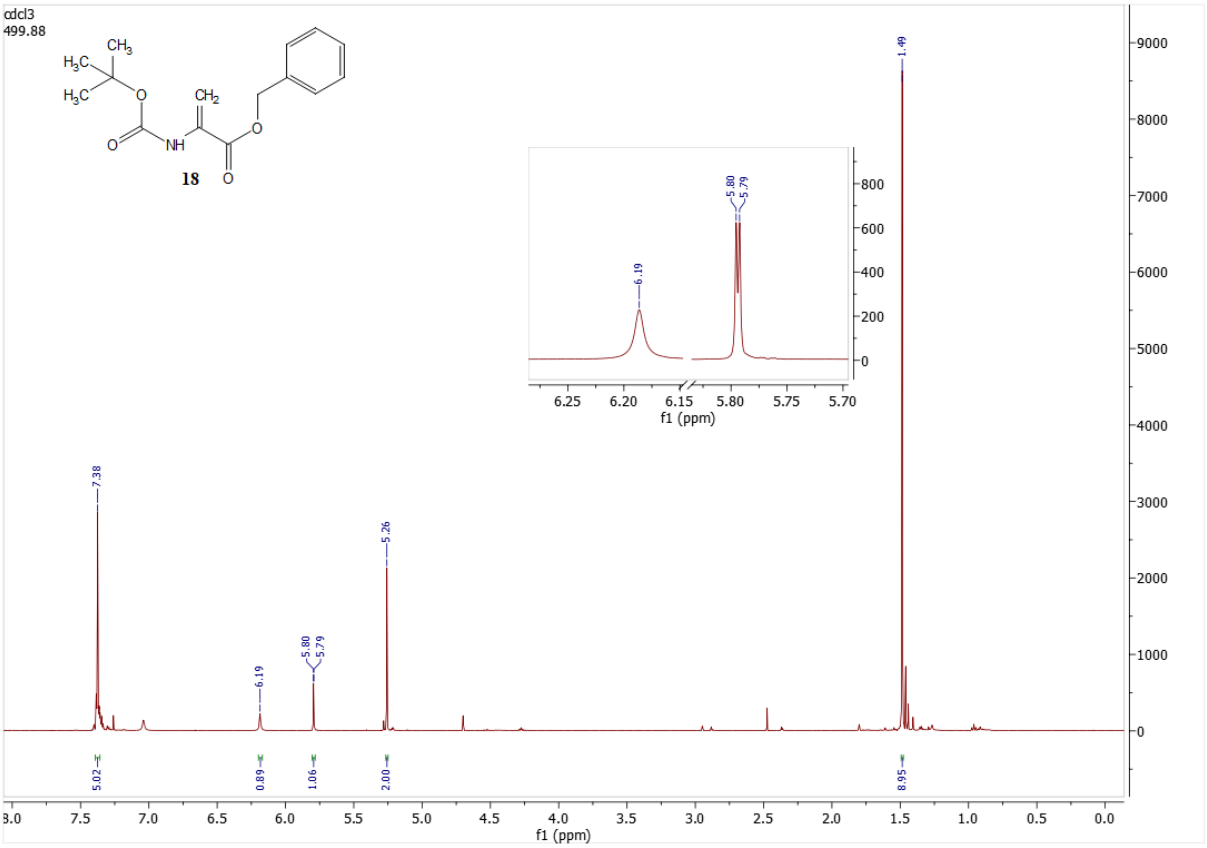


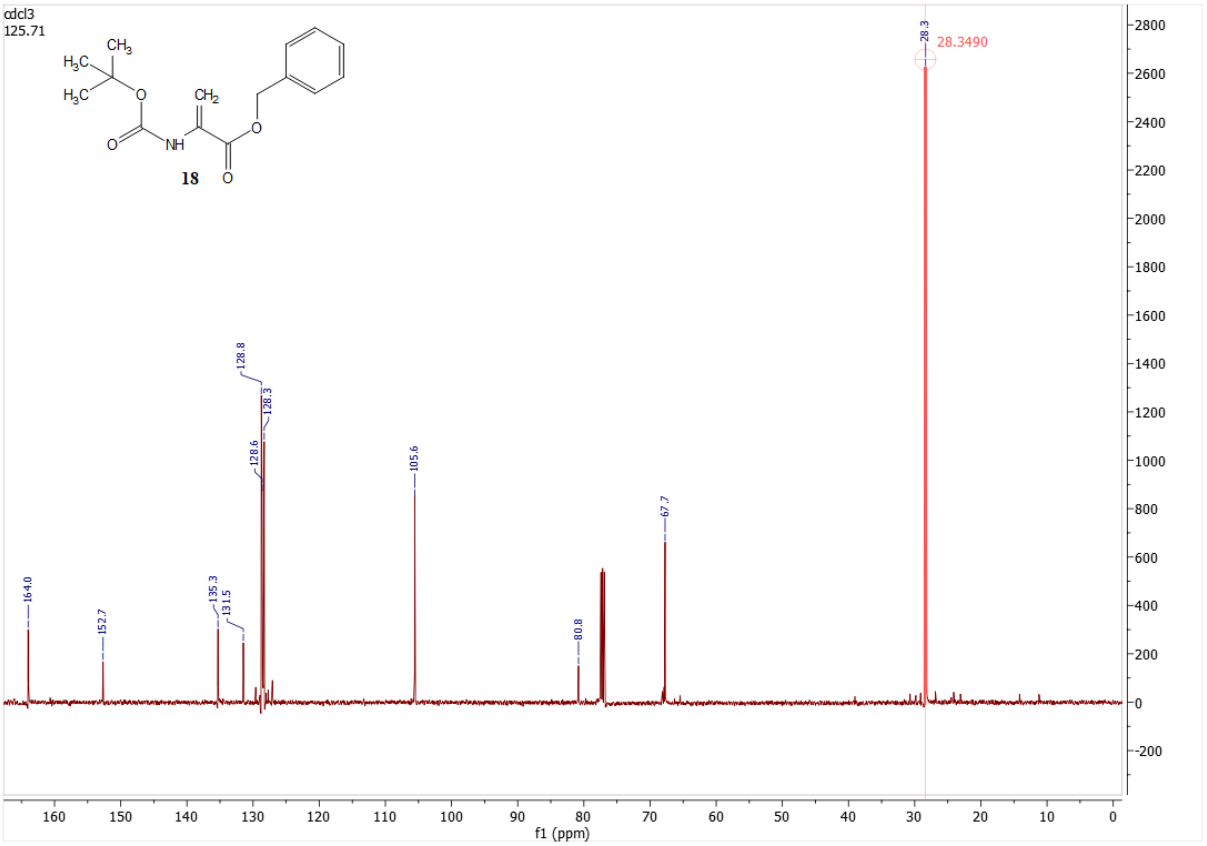
8.2.8 NMR analysis results of compound 17



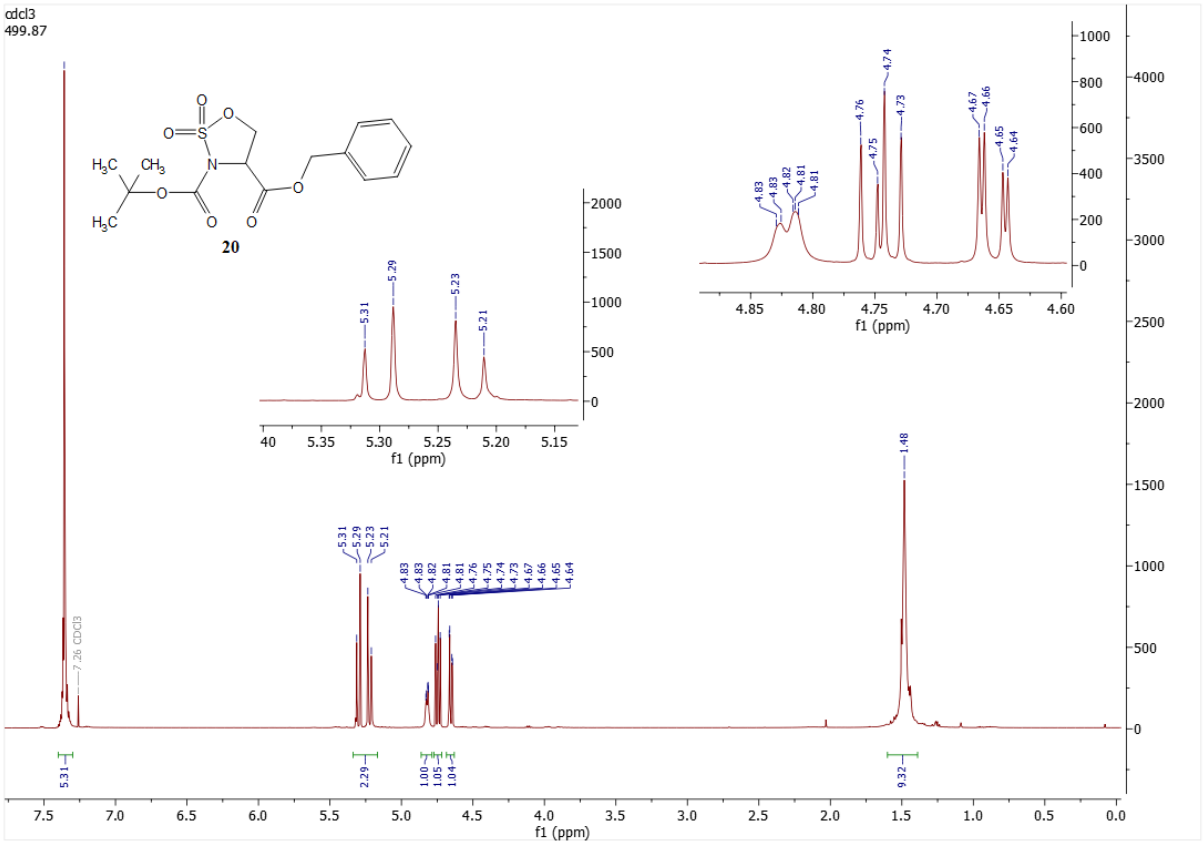


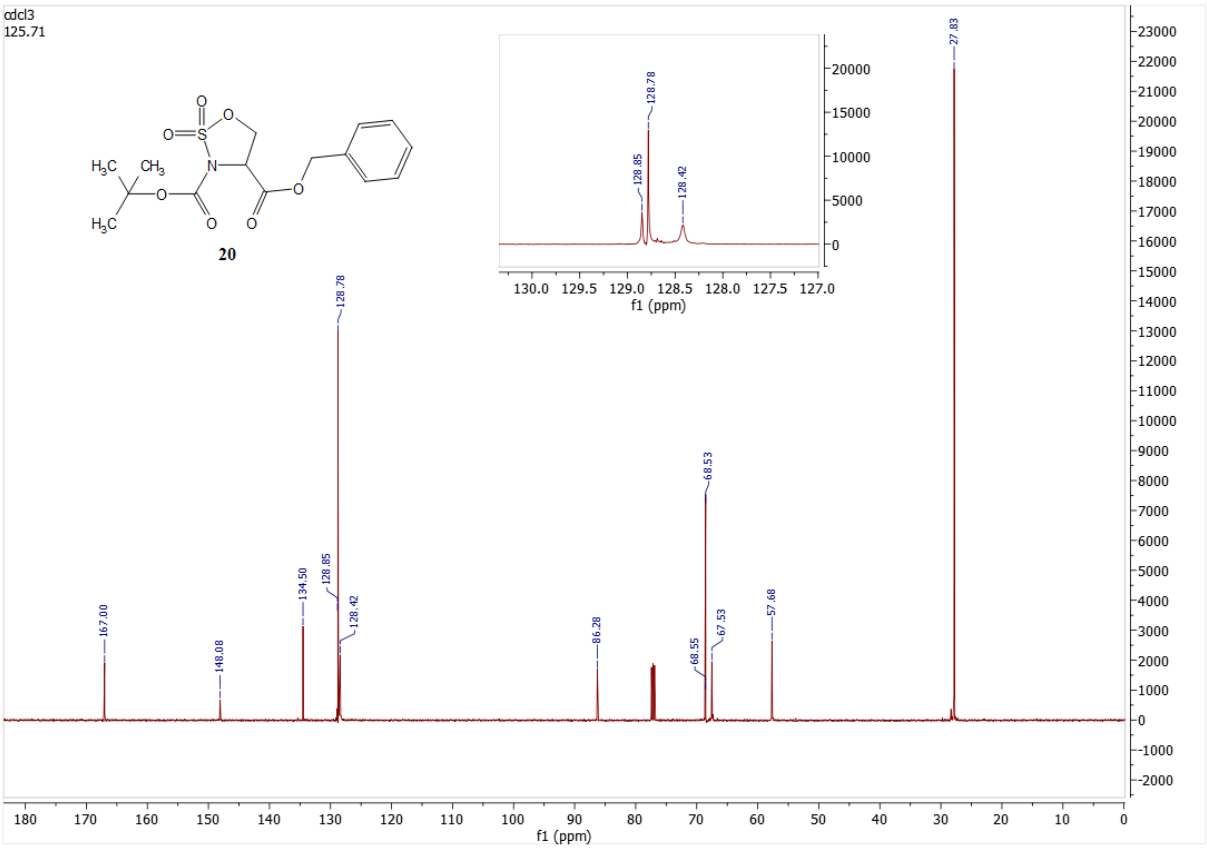
8.2.9 NMR analysis results of compound 18



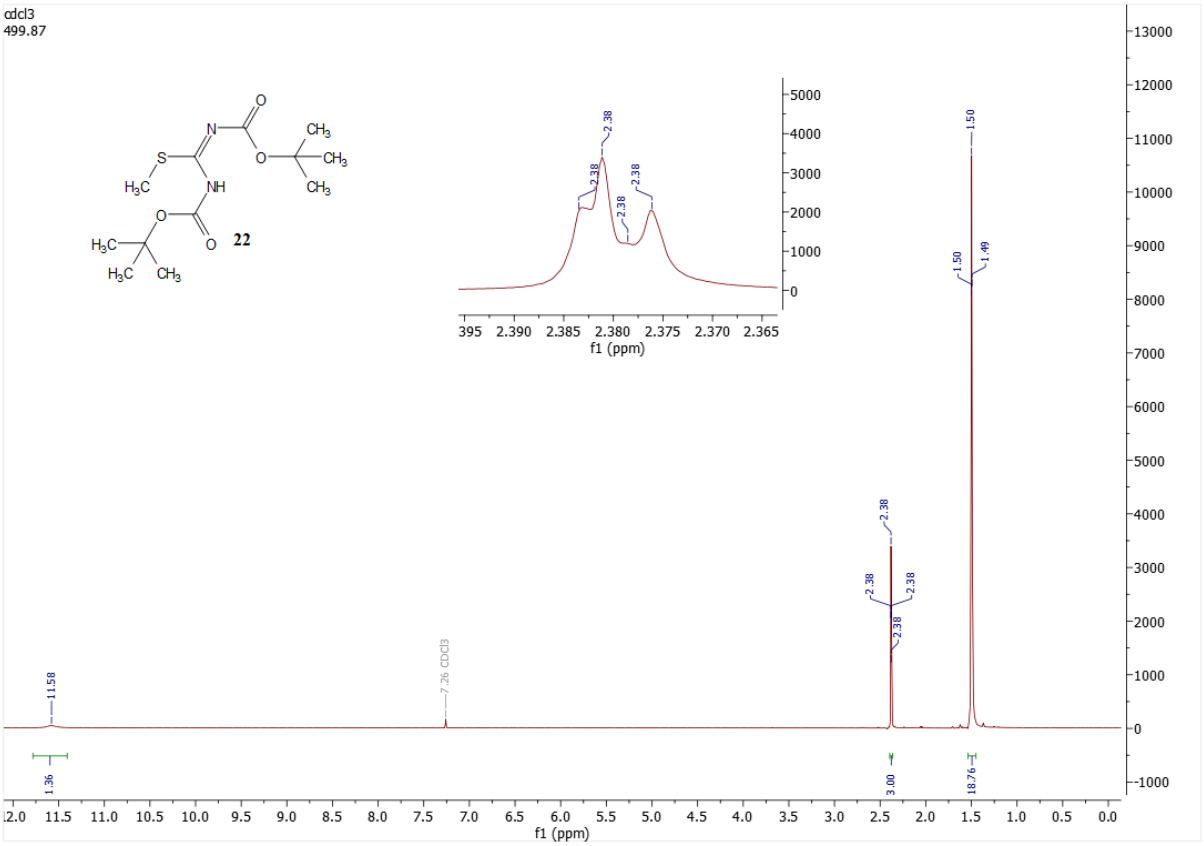


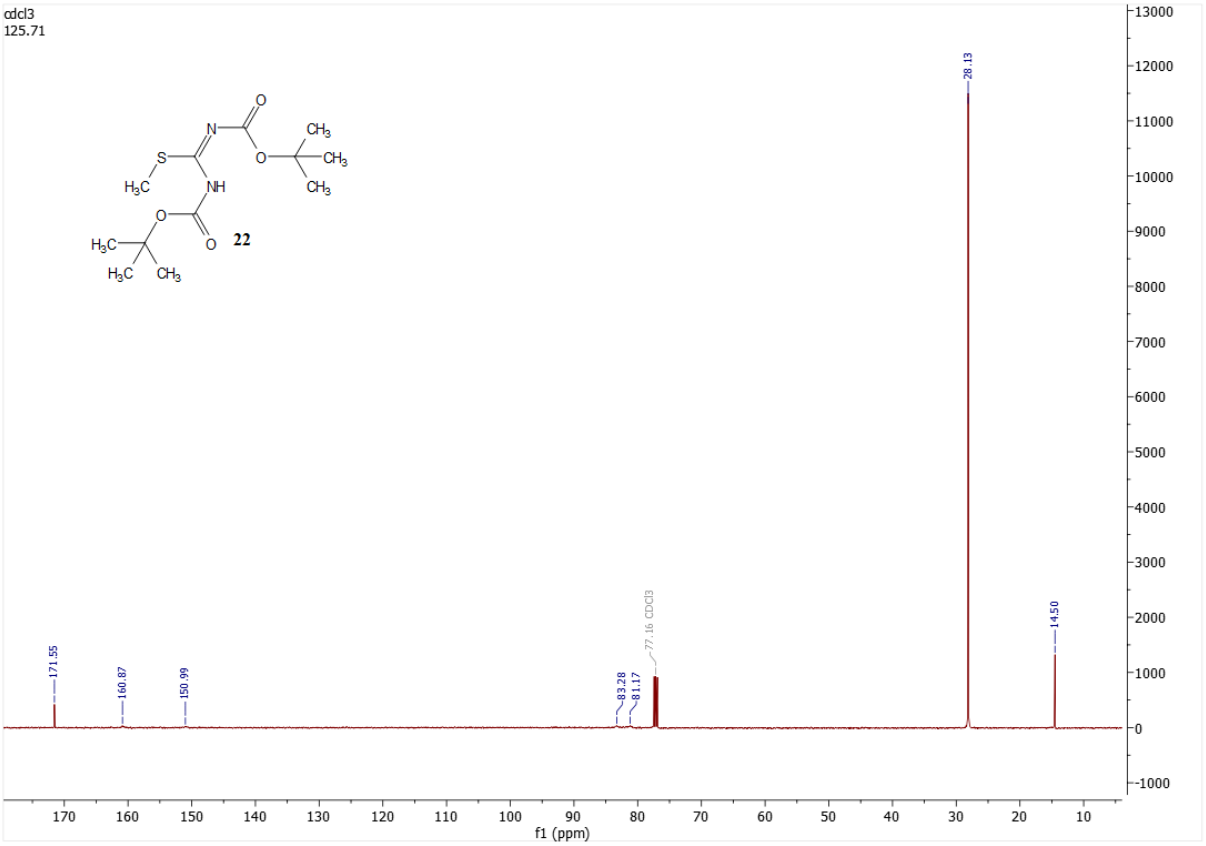
8.2.10 NMR analysis results of compound 20



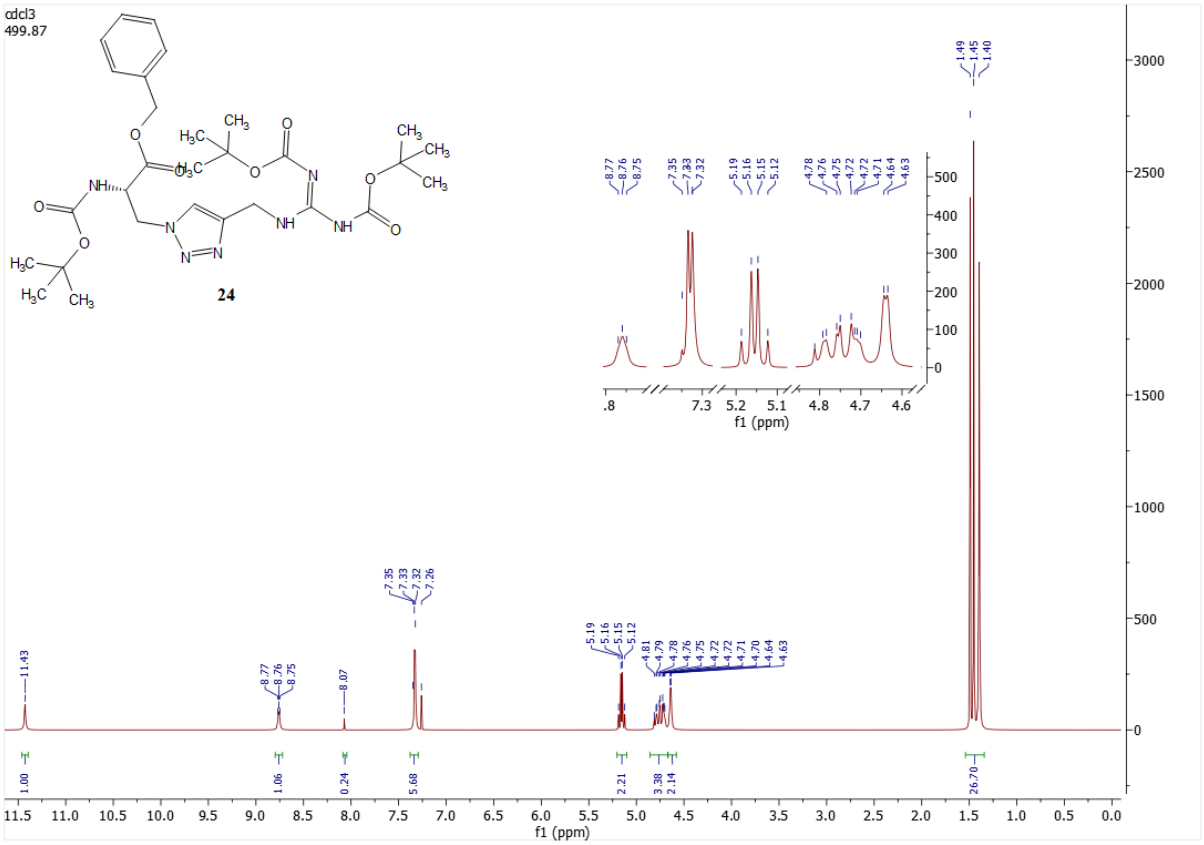


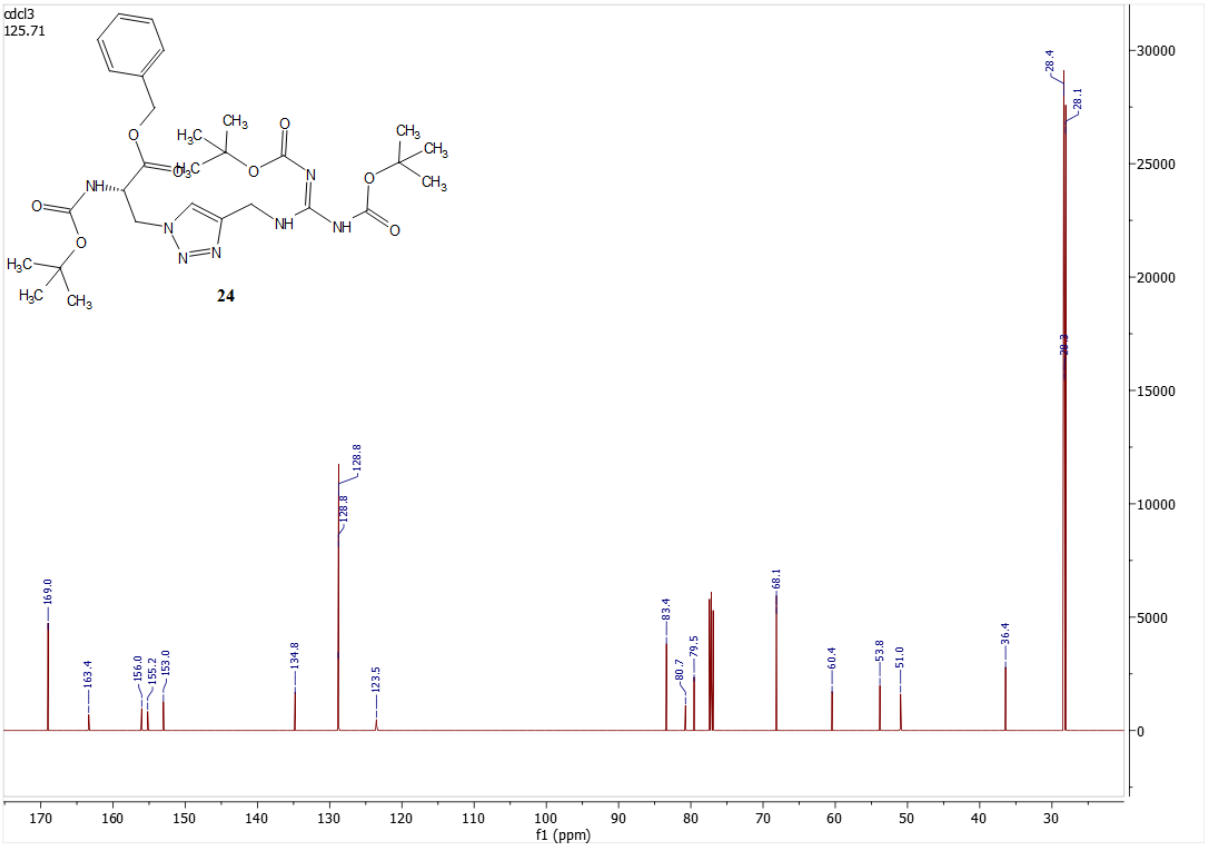
8.2.11 NMR analysis results of compound 22

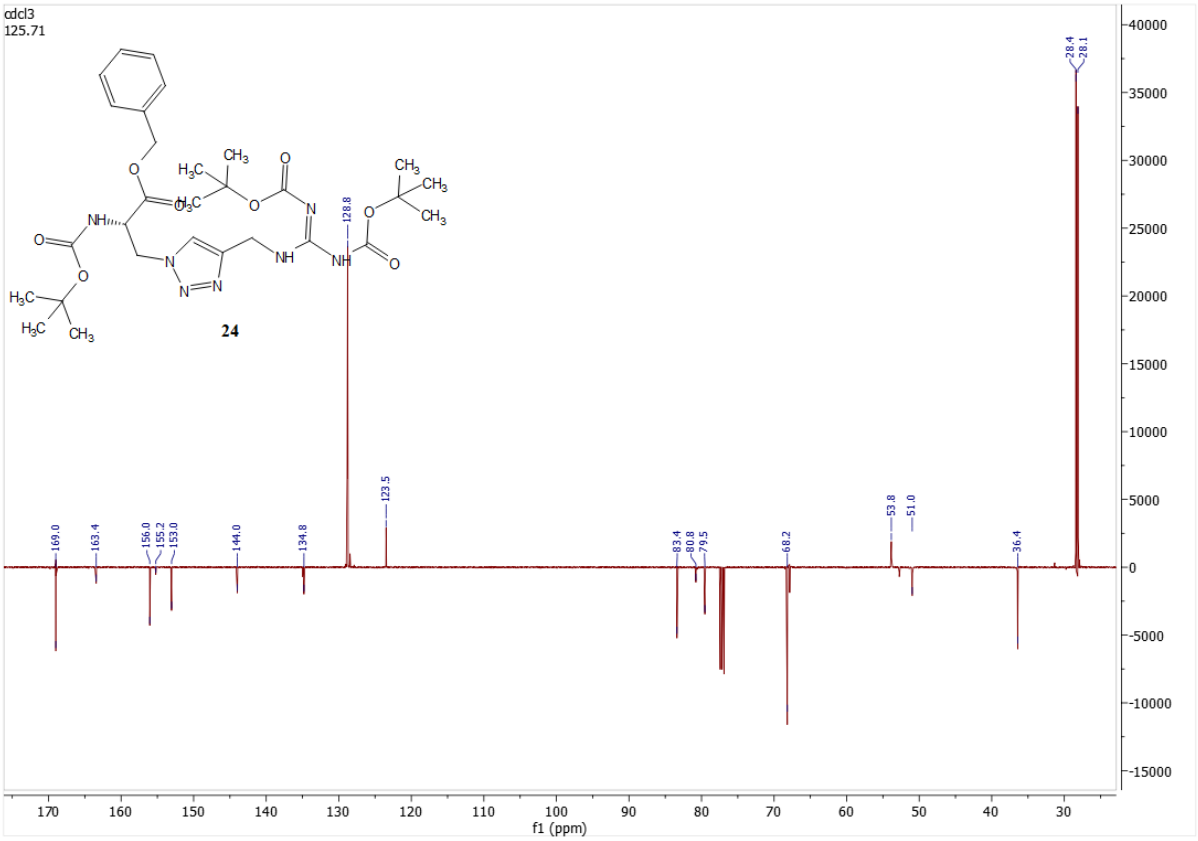




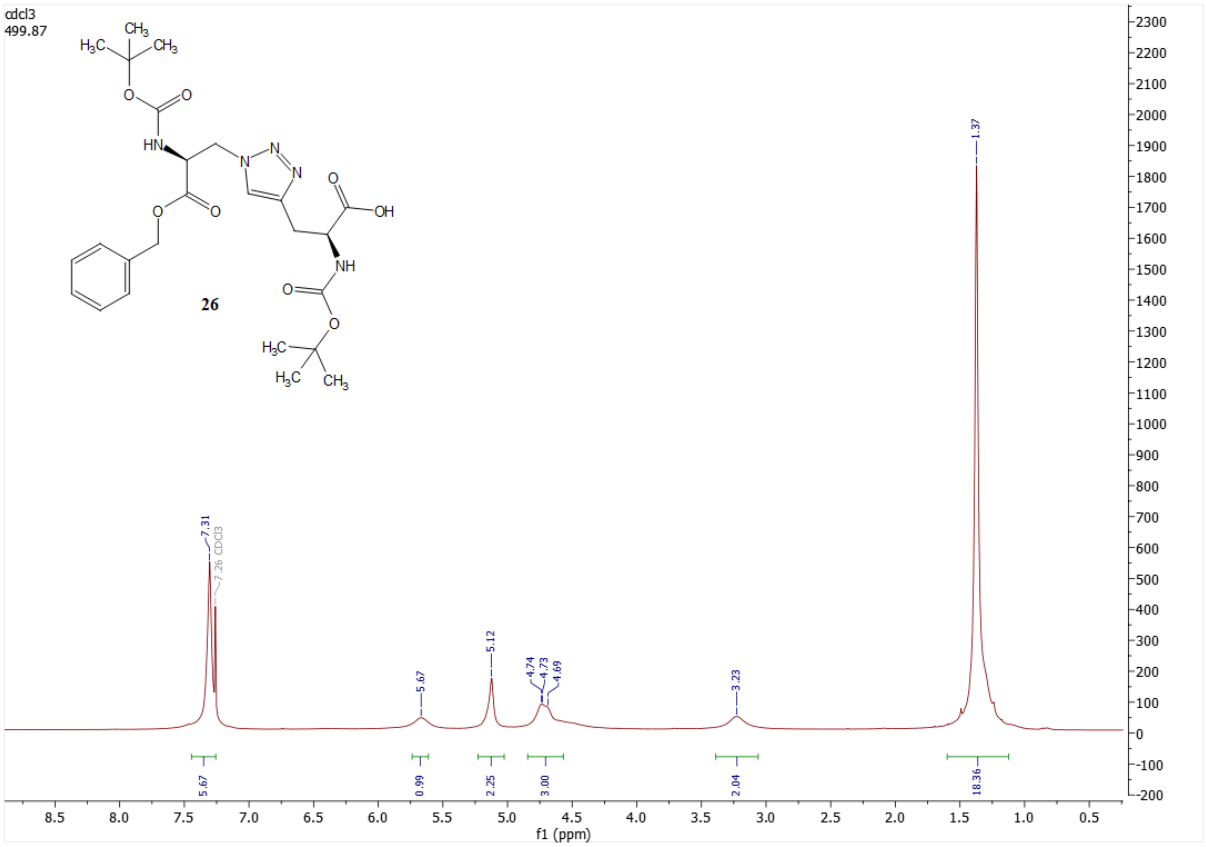
8.2.12 NMR analysis results of compound 24

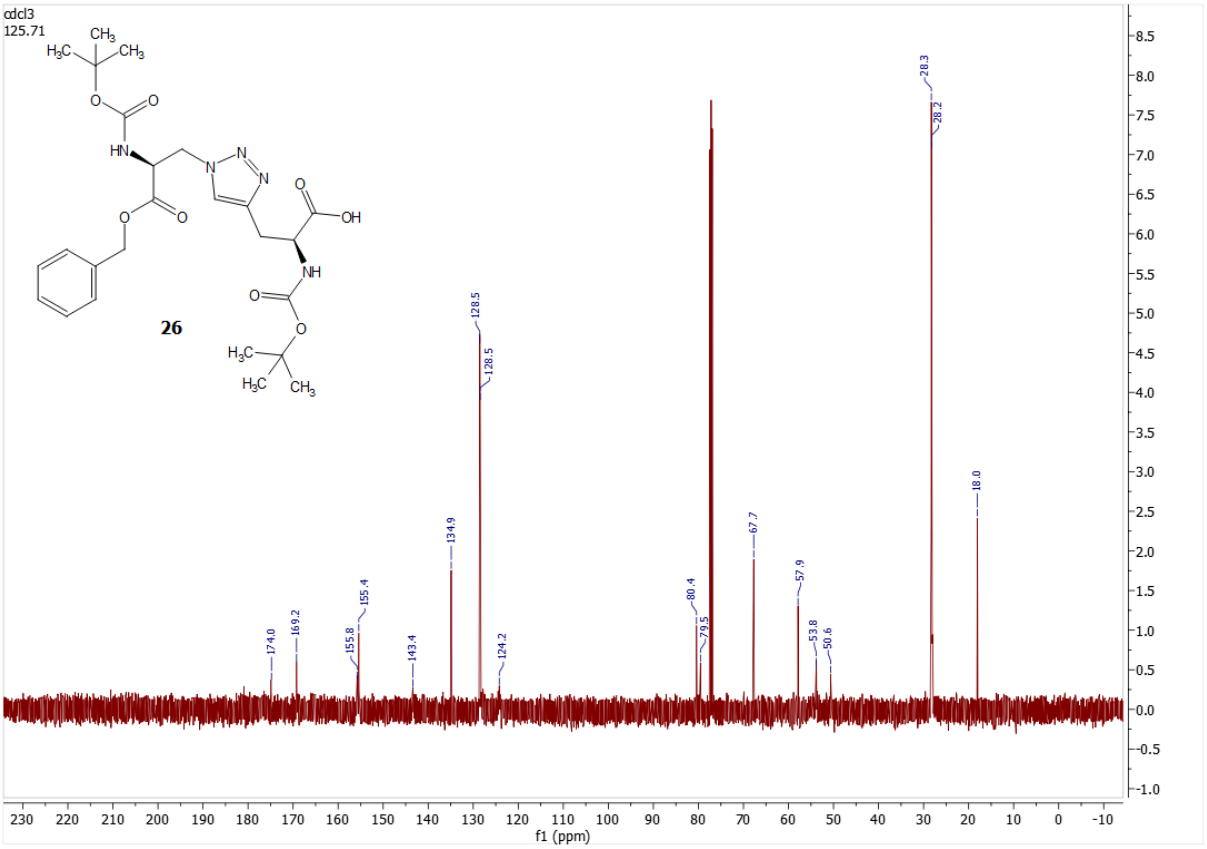




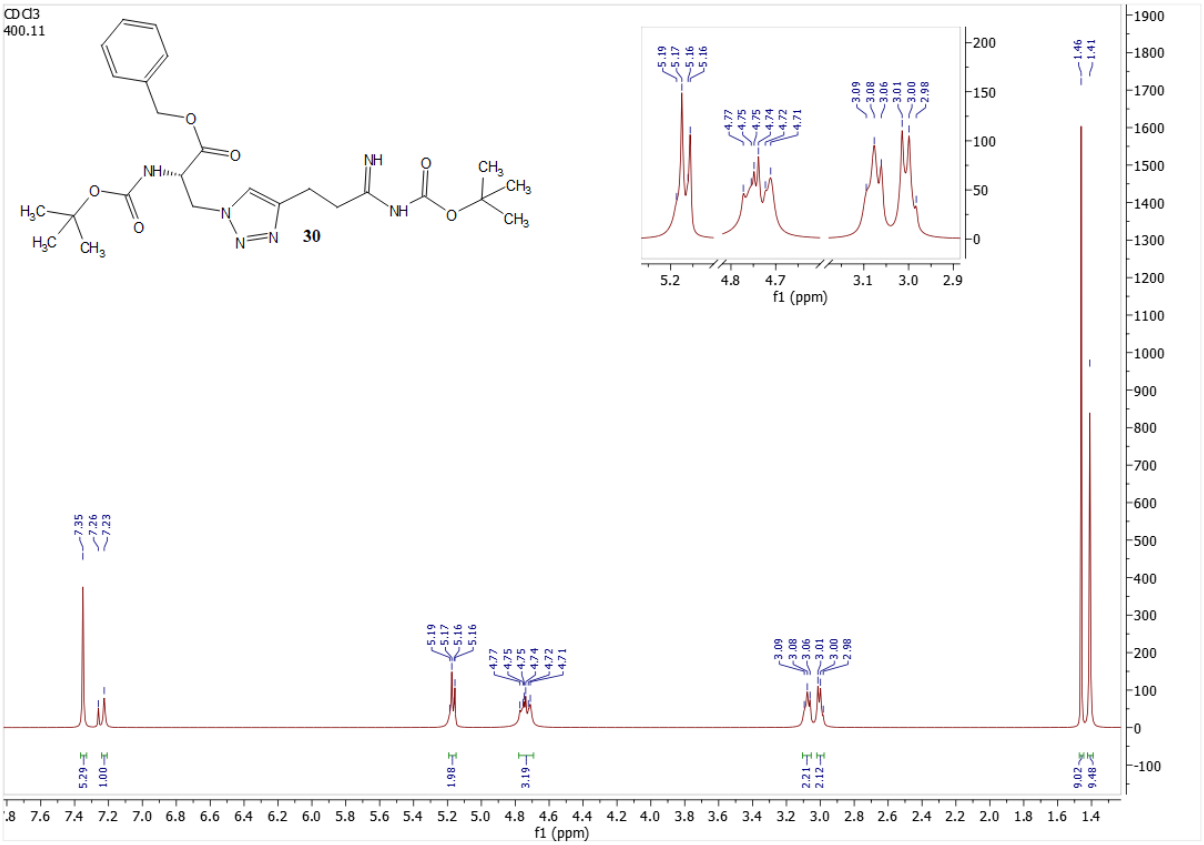

cdcl3
125.71
24
169.0
163.4
156.0
155.2
153.0
144.0
134.8
128.8
123.5
83.4
80.8
79.5
68.2
53.8
51.0
36.4
28.4
28.1
f1 (ppm)

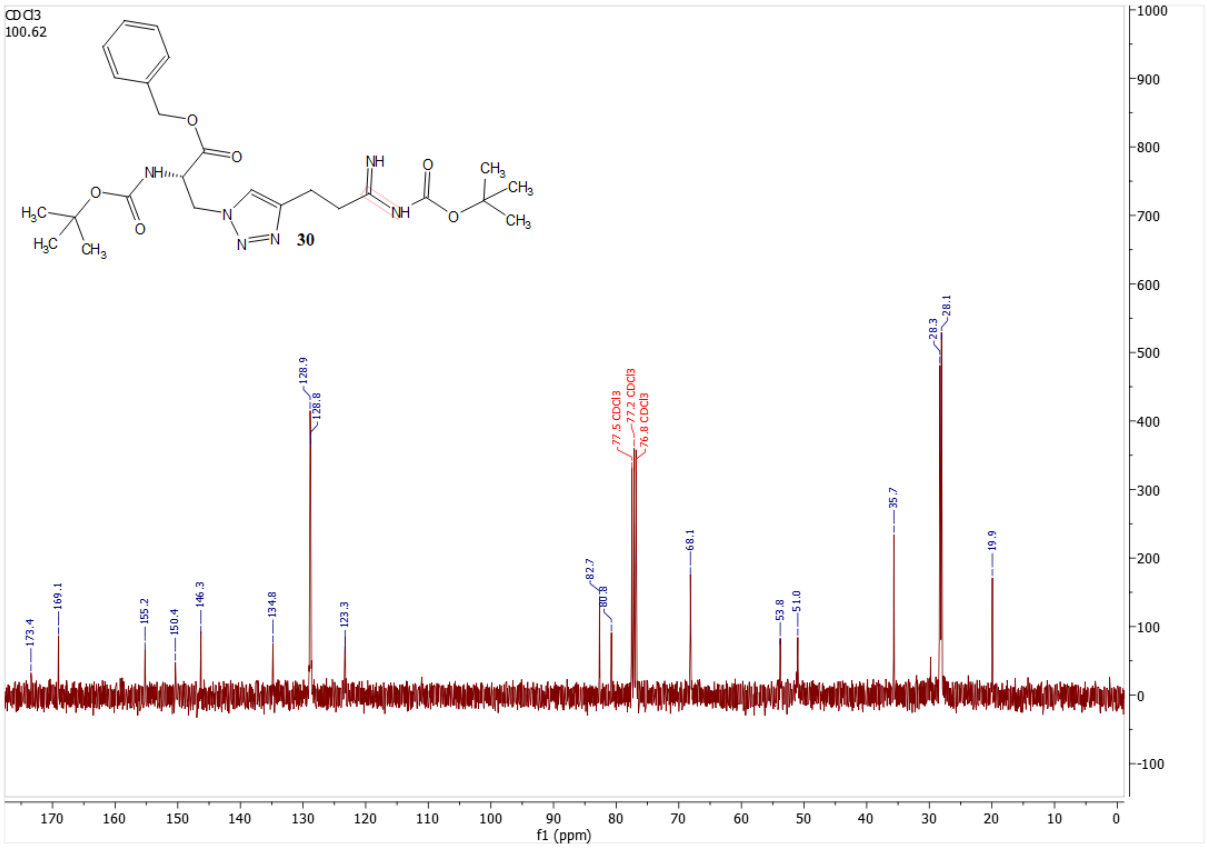
8.2.13 NMR analysis results of compound 26





8.2.14 NMR analysis results of compound 30





8.3 Docking

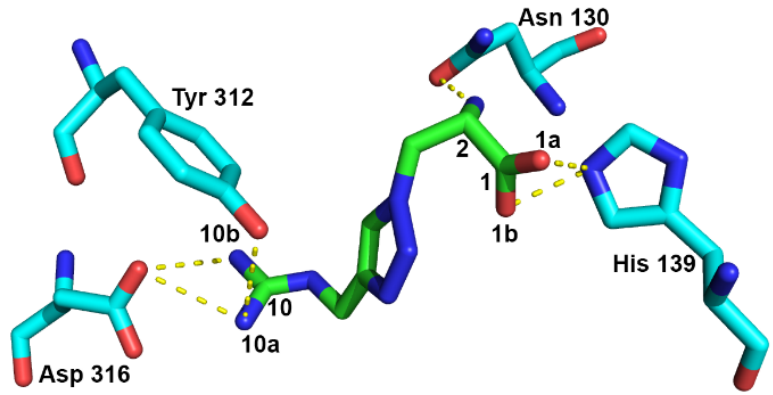
Represented docking results are for calculations with flexible amino acids residues within search box. Whenever docked structure docked out of position (Compound structure generally aligns with the Arg position in the PDB crystal structure but position of the carboxylic acid deviates strongly.) or out of binding pocket (There is no alignment between the docked structure and Arg position in the PDB crystal structure.) results of calculations for rigid receptor structure are presented. If rigid receptor structure also resulted in compound binding out of position/out of binding pocket, position in relation to the position of Arg position in the PDB crystal structure is represented.

Measured distances come from the docking experiment.

Applied Colors:

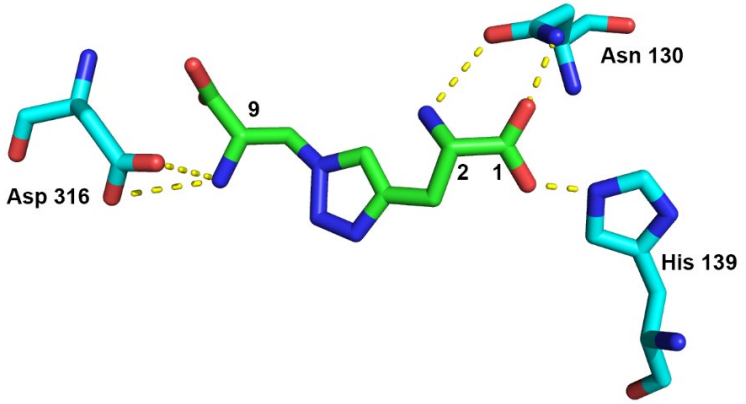
Carbon (ligand)	=	green
Carbon (protein residue)	=	cyan
Nitrogen	=	dark blue
Oxygen	=	red
Hydrogen-bond	=	yellow

8.3.1 Analysis of docking performed for compound 1a



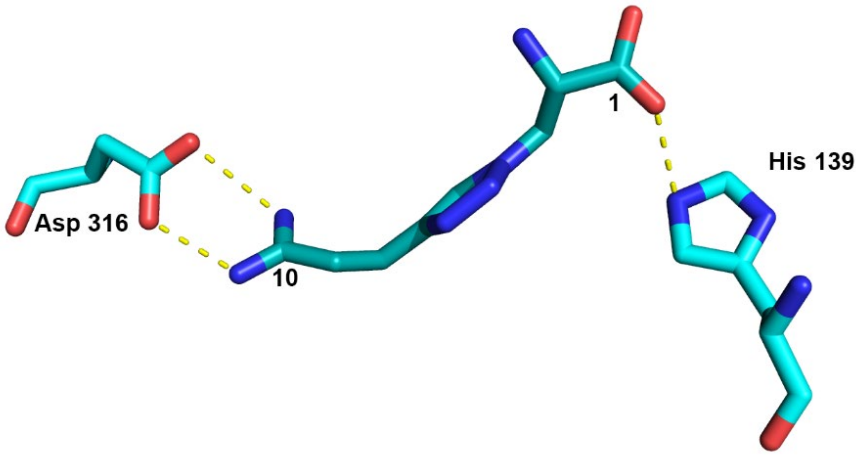
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(N2a)H	2.4
(His139)NH	(C1)O1a	2.3
(His139)NH	(C1)O1b	2.3
(Tyr312)O^{4}H	(N11a)H	3.4
(Asp316)O^{g1}H	(N11a)H	3.1
(Asp316)O^{g1}H	(N11b)H	3.2

8.3.2 Analysis of docking performed for compound 1b



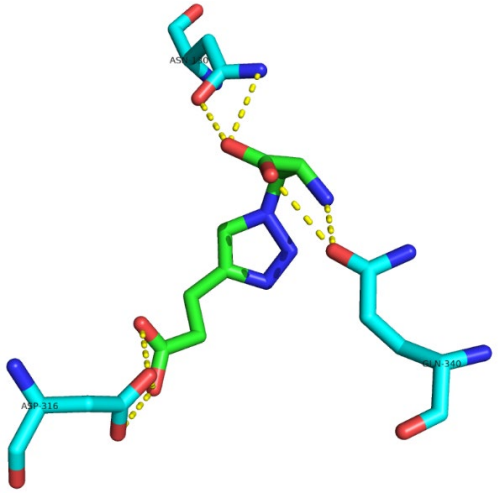
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(N2)H	3.2
(Asn130)N^{g}	(C1)O^{1}	2.1
(His139)NH	(C1)O^{2}	2.0
(Asp316)O^{g1}H	(N9)H	3.0
(Asp316)O^{g2}H	(N9)H	3.2

8.3.3 Analysis of docking performed for compound 1c



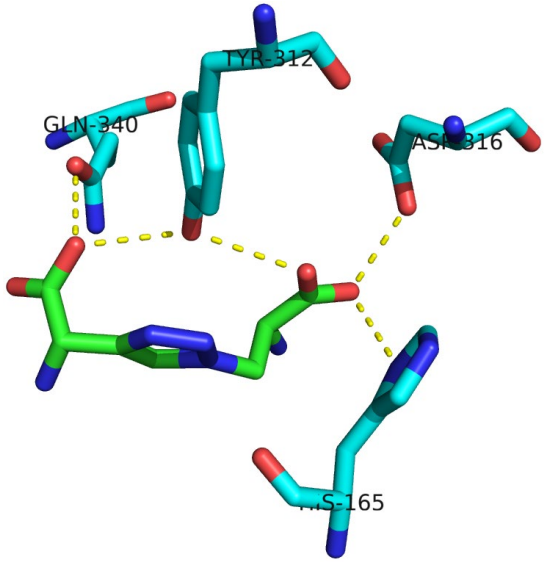
Residue	Atomic group	Distance (Å)
(His139)NH	(C1)O1a	2.6
(Asp316)$O^{g1}H$	(N10a)H	3.1
(Asp316)$O^{g2}H$	(N10b)H	2.7

8.3.4 Analysis of docking performed for compound 1d



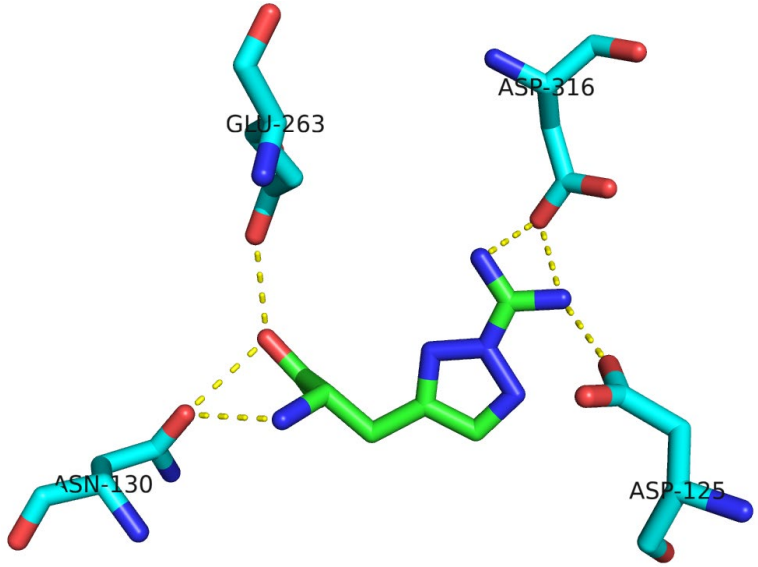
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(C1)$O^{1}H$	2.9
(Asn130)N^{g}	(C1)$O^{1}H$	3.0
(Asp316)$O^{g1}H$	(C9)O^{1}	2.6
(Asp316)$O^{g1}H$	(C9)O^{2}	2.6
(Asp316)$O^{g2}H$	(C9)O^{2}	2.9
(Gln340)O^{g1}	(C1) $O^{2}H$	2.6

8.3.5 Analysis of docking performed for compound 1e



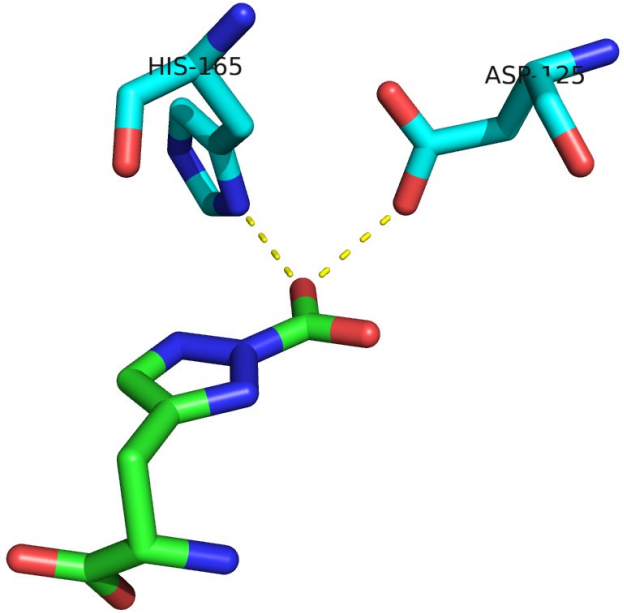
Residue	Atomic group	Distance (Å)
(His165)N^{π}	(C1)$O^{1}H$	2.6
(Tyr312)$O^{g1}H$	(C1)O^{2}	2.7
(Tyr312)$O^{g1}H$	(C8)O^{1}	2.6
(Asp316)$O^{g1}H$	(C1)O^{1}	2.8
(Gln340)O^{g1}	(C8) $O^{2}H$	2.9

8.3.6 Analysis of docking performed for compound 1f



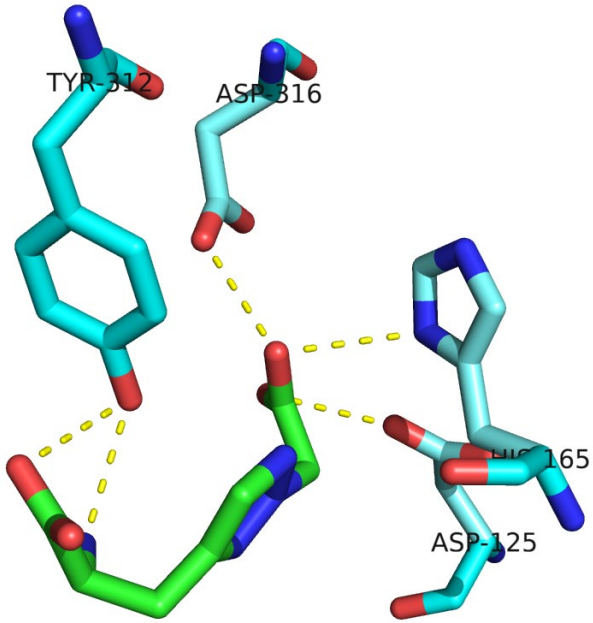
Residue	Atomic group	Distance (Å)
(Asp125)$O^{g}H$	(N8a)H	2.9
(Asn130)O^{g}	(C1)$O^{1}H$	2.8
(Asn130)O^{g}	(N2)H	3.0
(Glu263)$O^{\delta1}H$	(C1)O^{1}	3.0
(Asp316)$O^{g2}H$	(N8a)	3.0
(Asp316)$O^{g2}H$	(N8b)	2.7

8.3.7 Analysis of docking performed for compound 1g



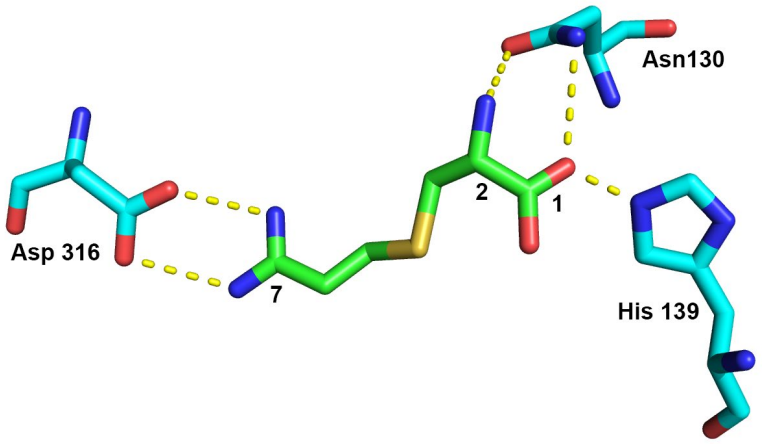
Residue	Atomic group	Distance (Å)
(Asp125) $O^{g2}H$	$(C1)O^{1}$	3.0
(His139)N	$(C1)O^{1}H$	2.6

8.3.8 Analysis of docking performed for compound 1h



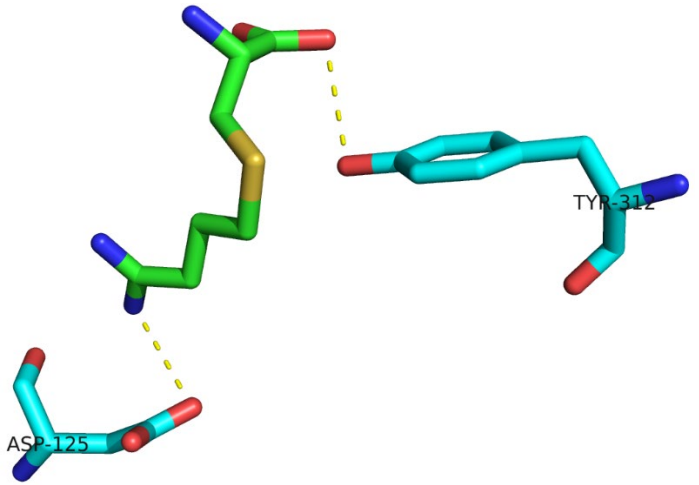
Residue	Atomic group	Distance (Å)
(Asp125)$O^{g}H$	$(C9)O^{1}$	2.6
(His165)N^{π}	$(C9)O^{2}H$	2.7
(Tyr312)$O^{g1}H$	$(C1)O^{1}$	2.7
(Tyr312)$O^{g1}H$	(N2)H	3.0
(Asp316)$O^{g2}H$	$(C9)O^{2}$	3.0

8.3.9 Analysis of docking performed for compound 2a



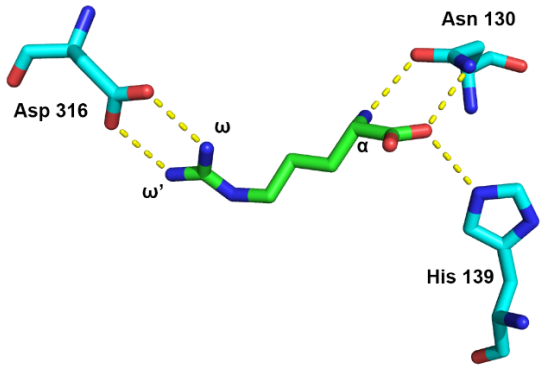
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(N2)H	2.7
(Asn130)N^{g}	(C1)O^{1}	3.3
(His139)NH	(C1)O^{1}	2.1
(Asp316)O^{g1}H	(N7a)H	3.1
(Asp316)O^{g2}H	(N7b)H	3.0

8.3.10 Analysis of docking performed for compound 2b



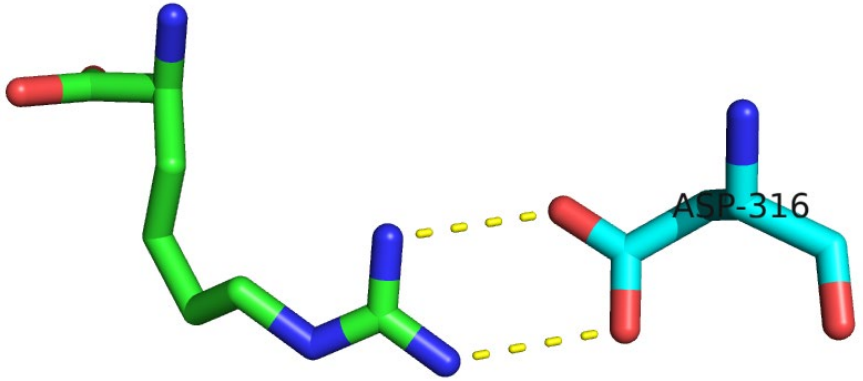
Residue	Atomic group	Distance (Å)
(Asp125)O^{g1}	(N8)H	3.2
(Tyr312)O^{g1}H	(C1)O^{1}	3.1

8.3.11 Analysis of docking performed for compound 2c



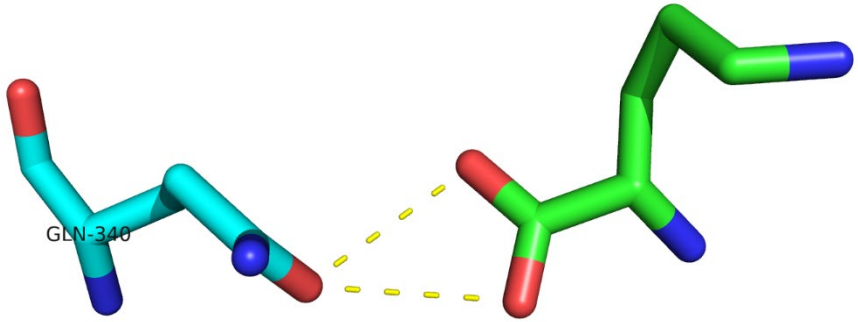
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(Nα)H	2.9
(Asn130)N^{g}	(C1)O^{1}	2.8
(His139)NH	(C1)O^{1}	3.1
(Asp316)O^{g1}H	(Nω)H	3.1
(Asp316)O^{g1}H	(Nω')H	3.0

8.3.12 Analysis of docking performed for compound 2d

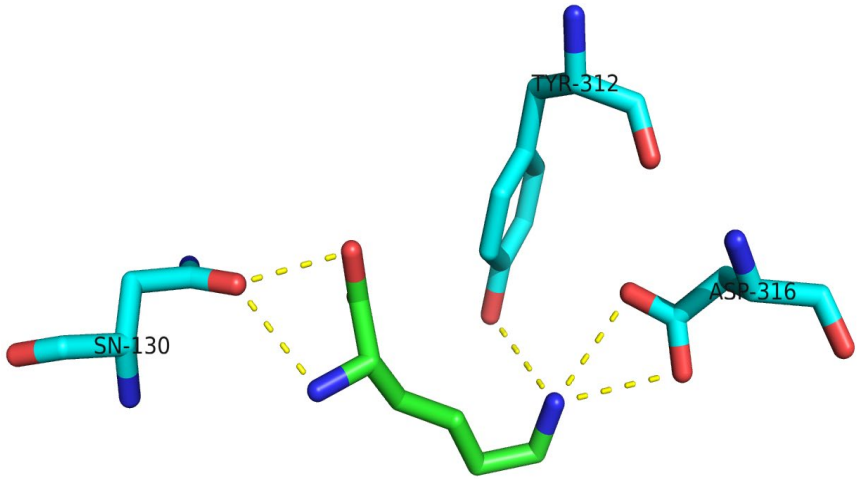


Residue	Atomic group	Distance (Å)
(Asp316)O^{g1}H	(N8a)H	3.0
(Asp316)O^{g2}H	(N8b)H	3.0

8.3.13 Analysis of docking performed for compound 2e



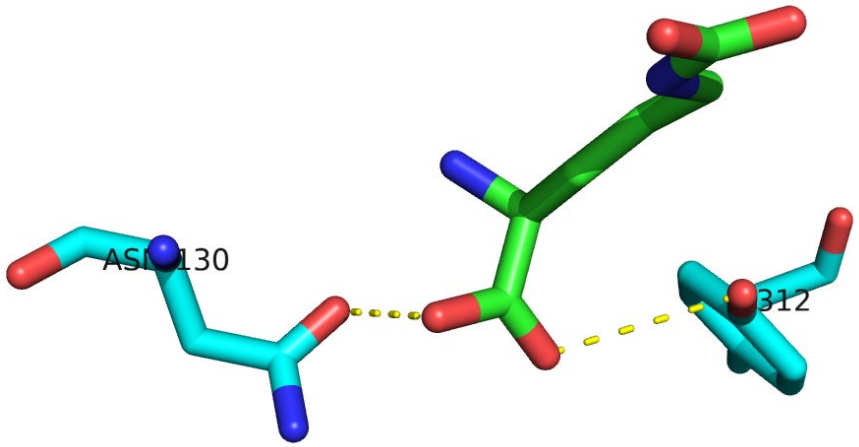
Residue	Atomic group	Distance (Å)
(Gln340)O^{g1}	(C1) $O^{2}H$	3.1
(Gln340)O^{g1}	(C1) $O^{1}H$	3.0



Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(N2)H	2.8
(Asn130)O^{g}	(C1)$O^{1}H$	3.0
(Tyr312)$O^{g1}H$	(N6)H	3.1
(Asp316)$O^{g1}H$	(N6)H	3.1
(Asp316)$O^{g2}H$	(N6)H	3.2

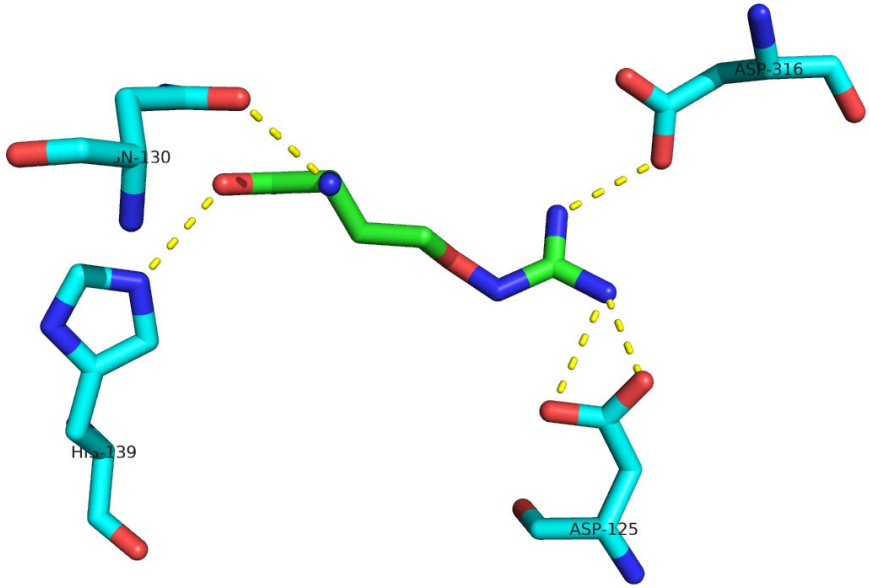
8.3.14 Analysis of docking performed for compound 2f

8.3.15 Analysis of docking performed for compound 2g



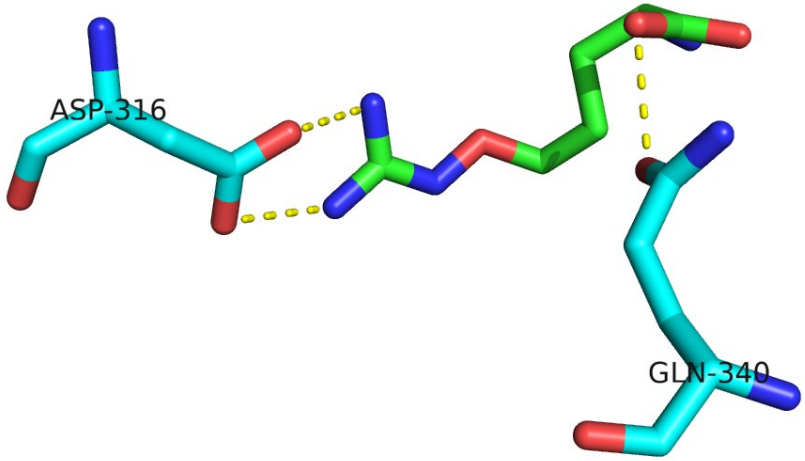
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(C1)$O^{1}H$	3.1
(Tyr312)$O^{g1}H$	(C1)O^{2}	3.2

8.3.16 Analysis of docking performed for compound 2h



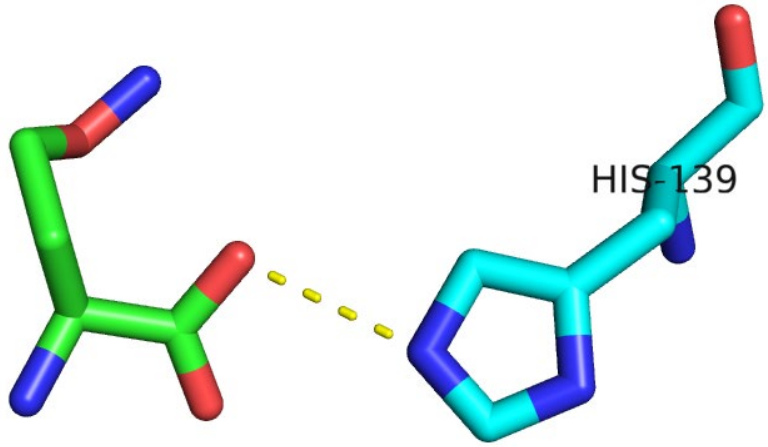
Residue	Atomic group	Distance (Å)
(Asp125)O^{g1}	(N7a)H	3.0
(Asp125)O^{g2}	(N7a)H	2.4
(Asn130)O^{g}	(N2)H	2.9
(His139)NH	(C1)O^{1}	2.9
(Asp316)$O^{g1}H$	(N7b)H	2.8

8.3.17 Analysis of docking performed for compound 2i



Residue	Atomic group	Distance (Å)
(Asp316)$O^{g1}H$	(N8a)H	3.1
(Asp316)$O^{g2}H$	(N8b)H	3.1
(Gln340)O^{g1}	(C1) $O^{1}H$	3.0

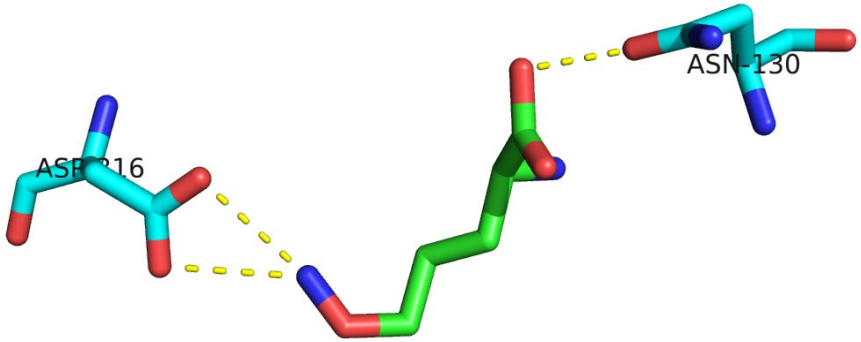
8.3.18 Analysis of docking performed for compound 2j



Residue	Atomic group	Distance (Å)
(His139)NH	(C1)O^{1}	2.9

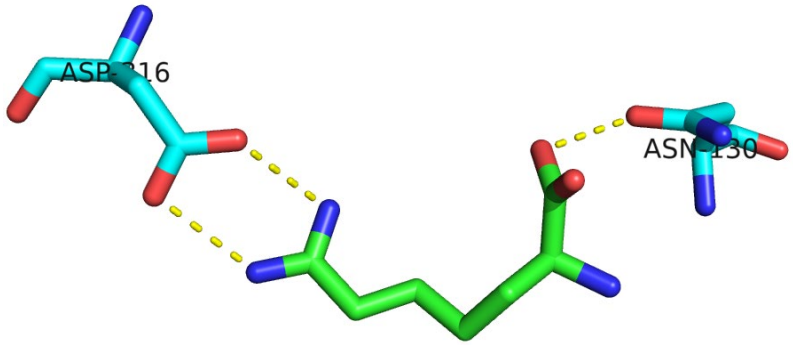
Rigid receptor structure, compound docked out of position in flexible mode.

8.3.19 Analysis of docking performed for compound 2k



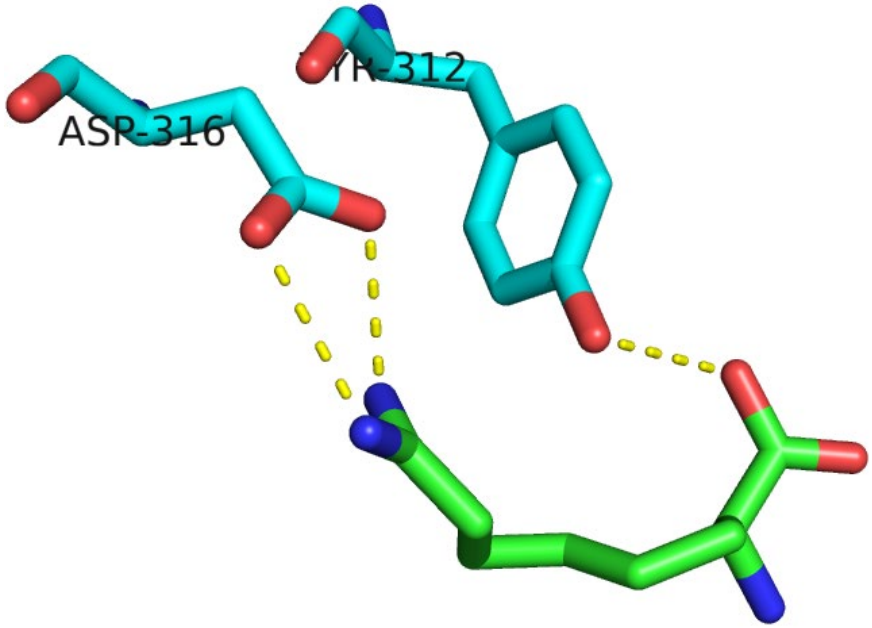
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(C1)O^{1}	2.9
(Asp316)O^{g1}H	(N6)H	3.1
(Asp316)O^{g2}H	(N6)H	3.2

8.3.20 Analysis of docking performed for compound 2l



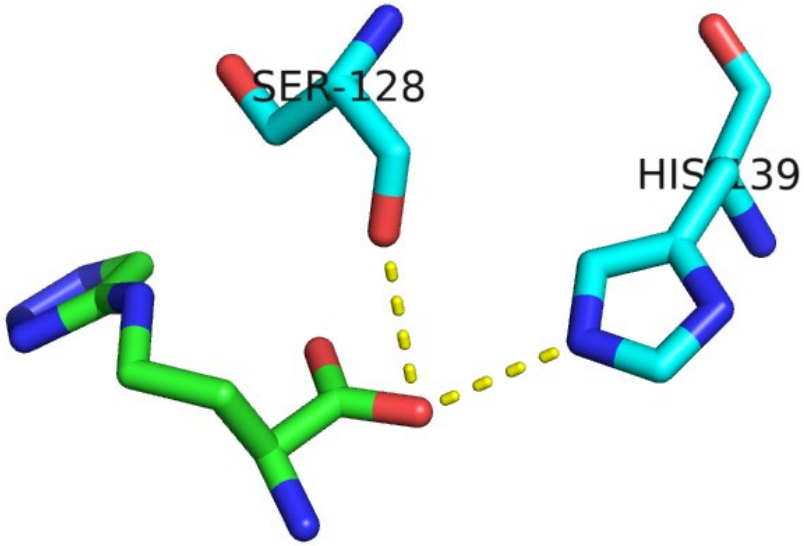
Residue	Atomic group	Distance (Å)
(Asn130)O^{g}	(C1)O^{1}	3.0
(Asp316)O^{g1}H	(N7a)H	3.1
(Asp316)O^{g2}H	(N7b)H	3.1

8.3.21 Analysis of docking performed for compound 2m



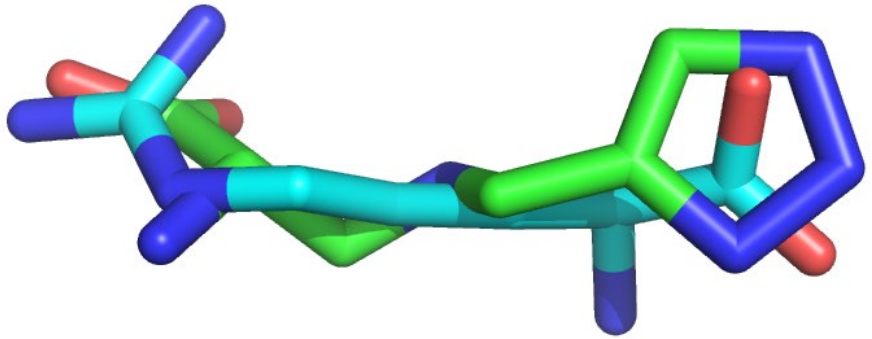
Residue	Atomic group	Distance (Å)
(Tyr312)$O^{g1}H$	(C1)O^{1}	3.2
(Asp316)$O^{g1}H$	(N8a)H	2.9
(Asp316)$O^{g2}H$	(N8b)H	3.2

8.3.22 Analysis of docking performed for compound 3a



Residue	Atomic group	Distance (Å)
(Ser128)O^{β}	(C1)O^{1}	3.2
(His139)NH	(C1)O^{1}	3.0

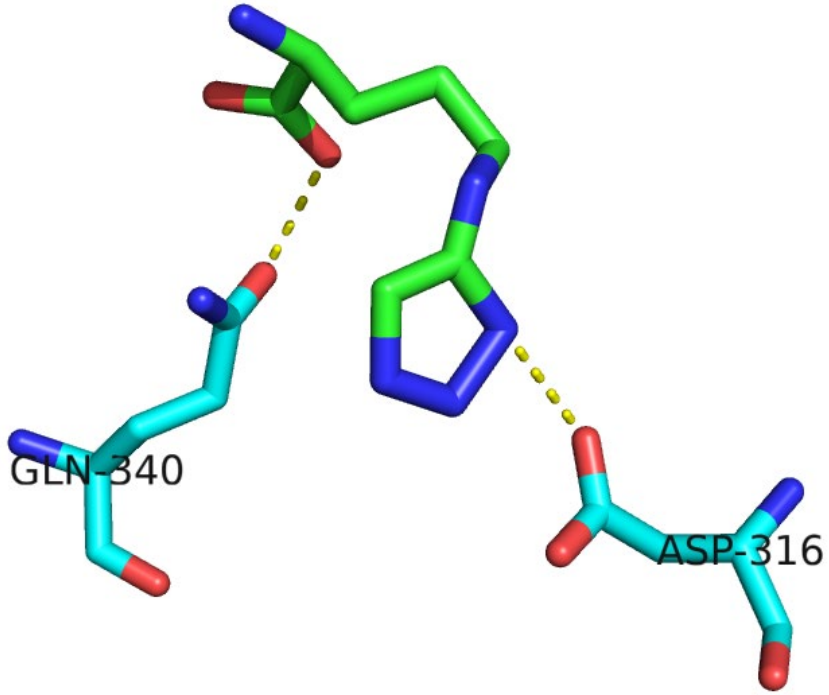
8.3.23 Analysis of docking performed for compound 3b



Residue	Atomic group	Distance (Å)

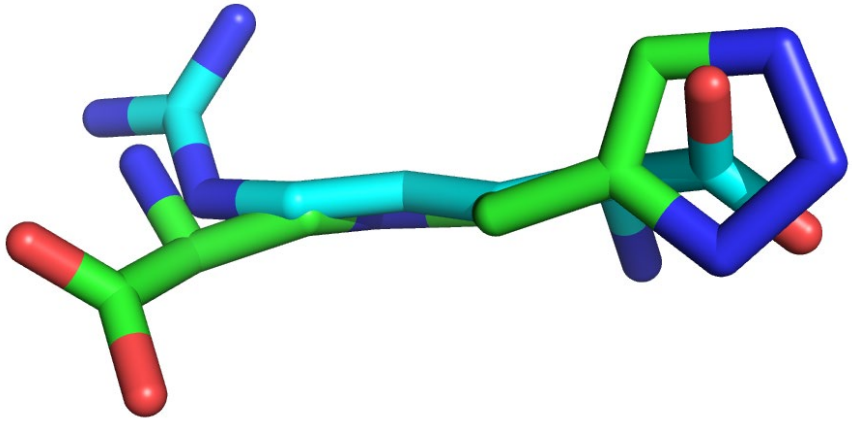
Structure docked out of position

8.3.24 Analysis of docking performed for compound 3c



Residue	Atomic group	Distance (Å)
(Asp316)O^{g1}	(N10)H	3.2
(Gln340)O^{g}	(C1)O^{1}	3.0

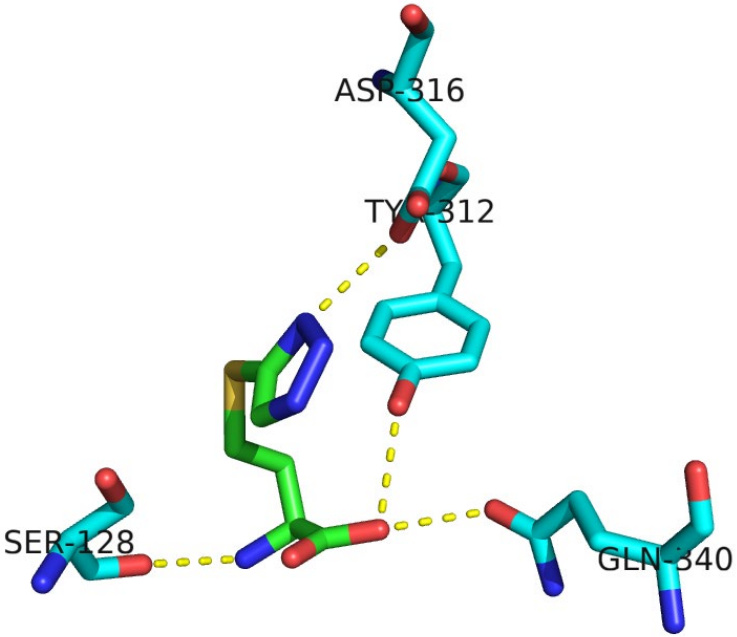
8.3.25 Analysis of docking performed for compound 3d



Residue	Atomic group	Distance (Å)

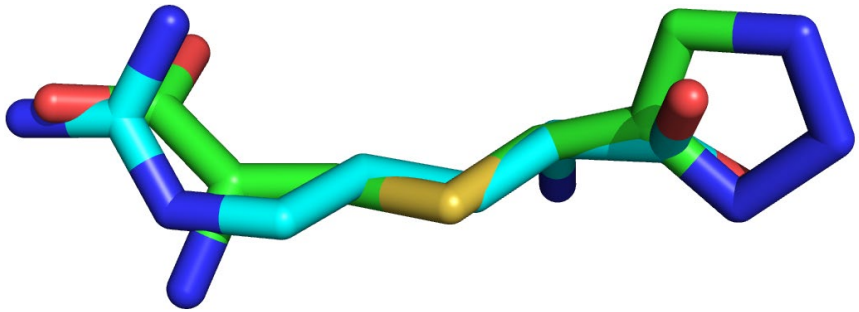
Structure docked out of position

8.3.26 Analysis of docking performed for compound 3e



Residue	Atomic group	Distance (Å)
(Ser128)O^{β}	(N2)H	2.8
(Tyr312)O^{g1}H	(C1)O^{1}	2.9
(Asp316)O^{g1}H	(N10)	3.1
(Gln340)O^{g}	(C1)O^{1}H	2.9

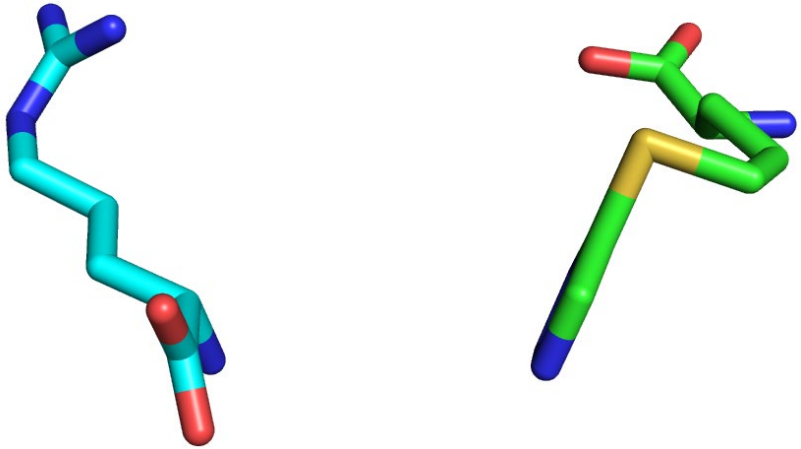
8.3.27 Analysis of docking performed for compound 3f



Residue	Atomic group	Distance (Å)

Structure docked out of position

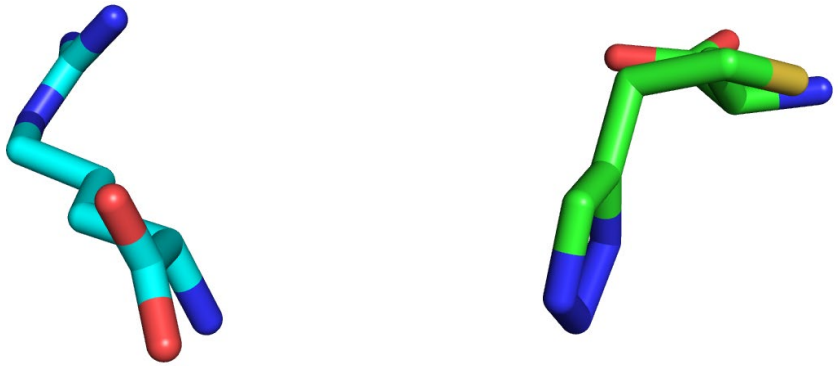
8.3.28 Analysis of docking performed for compound 3g



Residue	Atomic group	Distance (Å)

Structure docked out of binding pocket

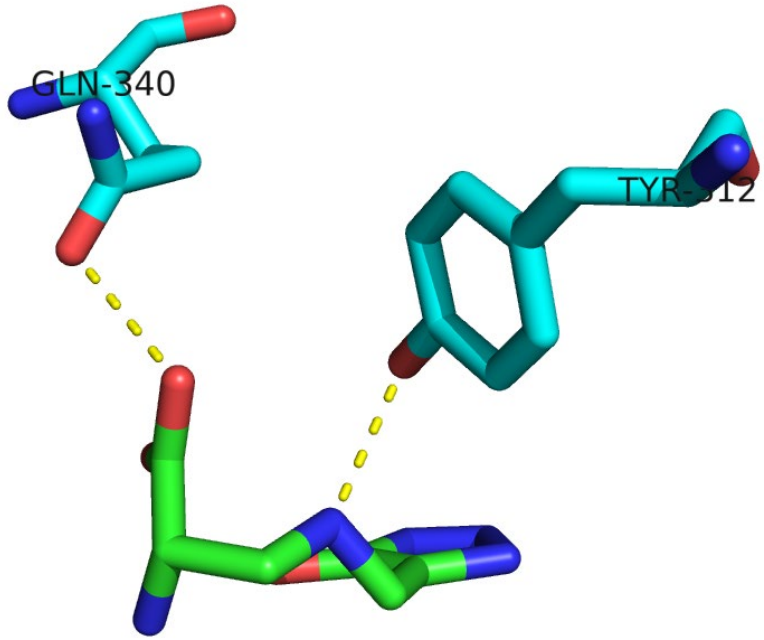
8.3.29 Analysis of docking performed for compound 3h



Residue	Atomic group	Distance (Å)

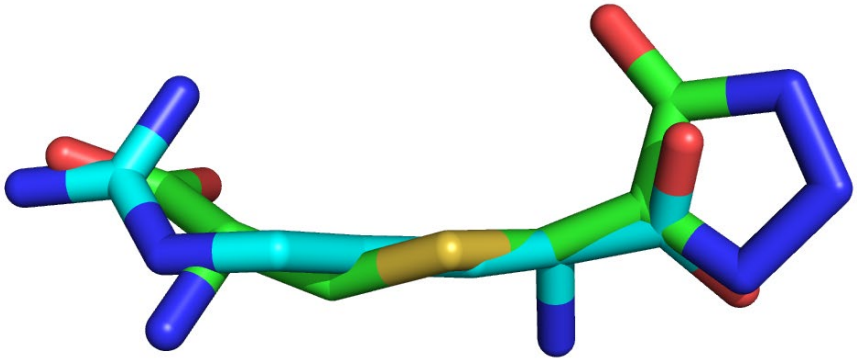
Structure docked out of binding pocket

8.3.30 Analysis of docking performed for compound 3i



Residue	Atomic group	Distance (Å)
(Tyr312)$O^{g1}H$	(N4)H	3.1
(Gln340)O^{g}	(C1)$O^{1}H$	3.0

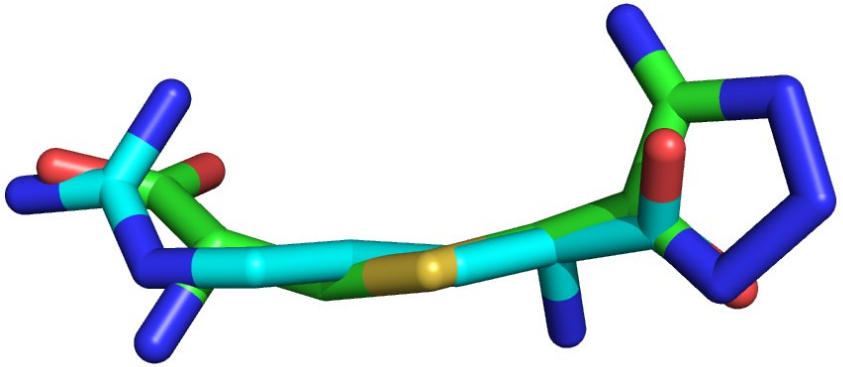
8.3.31 Analysis of docking performed for compound 3k



Residue	Atomic group	Distance (Å)

Structure docked out of position

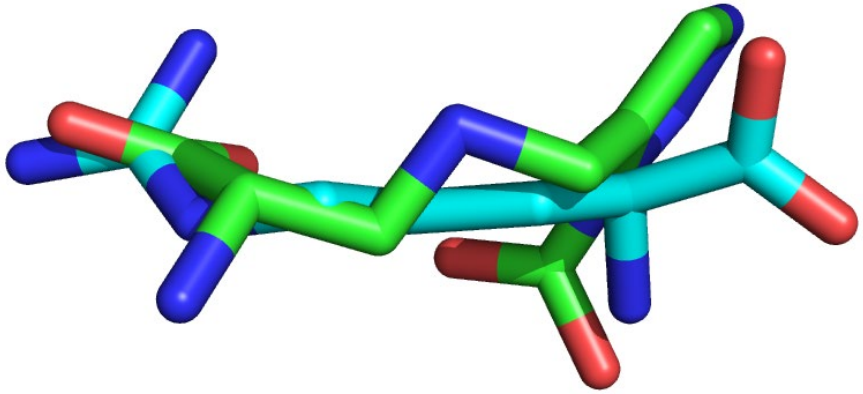
8.3.32 Analysis of docking performed for compound 3l



Residue	Atomic group	Distance (Å)

Structure docked out of position

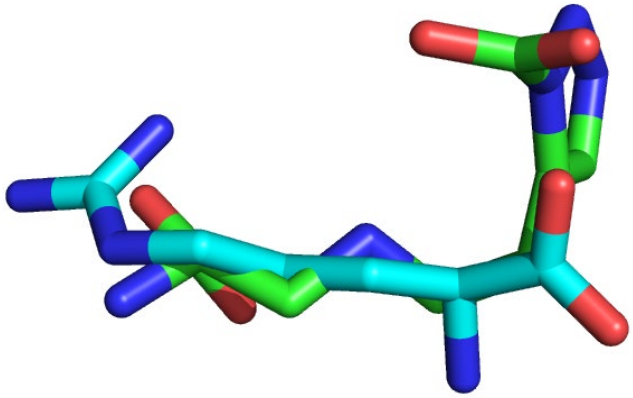
8.3.33 Analysis of docking performed for compound 3m



Residue	Atomic group	Distance (Å)

Structure docked out of position

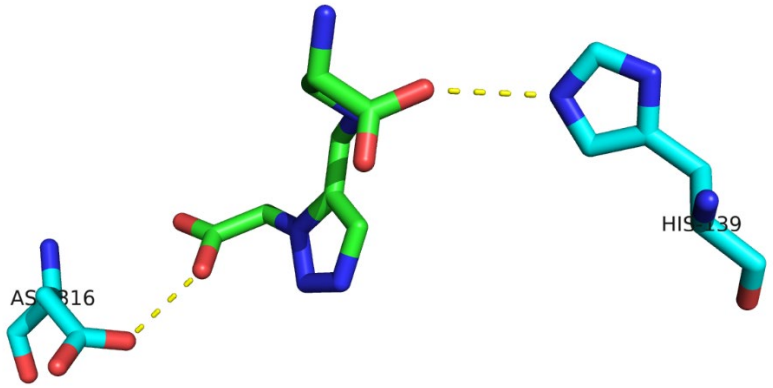
8.3.34 Analysis of docking performed for compound 3n



Residue	Atomic group	Distance (Å)

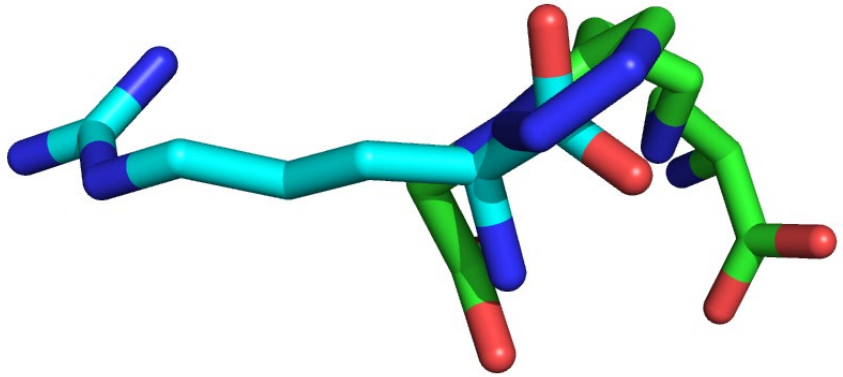
Structure docked out of position

8.3.35 Analysis of docking performed for compound 3o



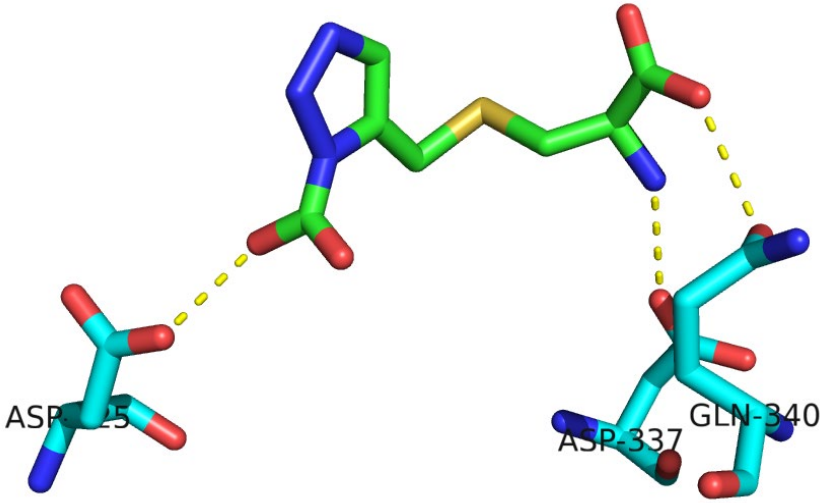
Residue	Atomic group	Distance (Å)
(His139)NH	(C1)O^{1}	3.0
(Asp316)O^{g1}	(12)O^{1}	2.8

8.3.36 Analysis of docking performed for compound 3p



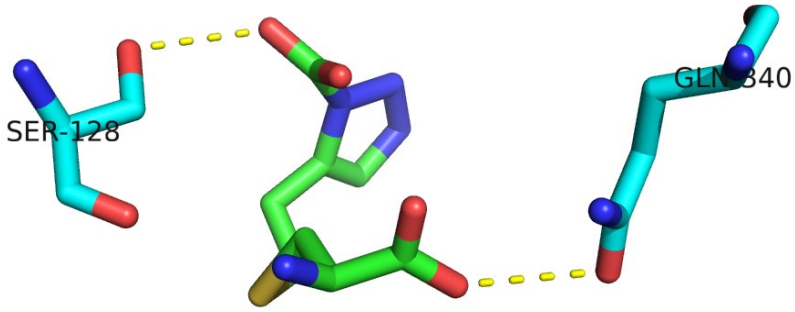
Residue	Atomic group	Distance (Å)
Structure docked out of binding pocket		

8.3.37 Analysis of docking performed for compound 3q



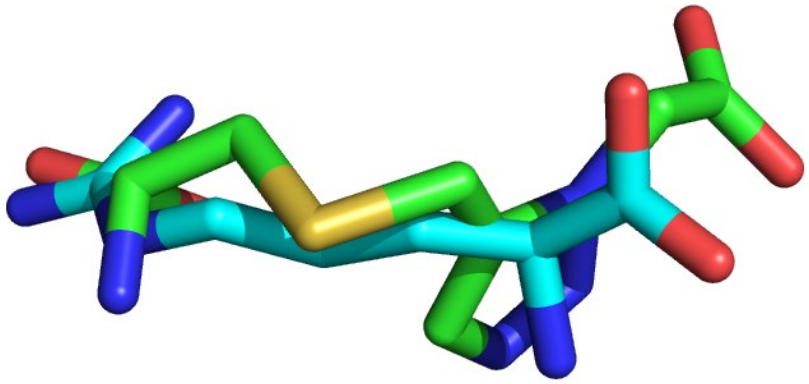
Residue	Atomic group	Distance (Å)
(Asp125)$O^{g1}H$	(C11)O^{1}	2.9
(Asp337)$O^{g2}H$	(N2)H	2.9
(Gln340)O^{g}	(C1)$O^{1}H$	3.0

8.3.38 Analysis of docking performed for compound 3r



Residue	Atomic group	Distance (Å)
(Ser128)O^{β}	(C12)O^{1}	3.0
(Gln340)O^{g}	(C1)$O^{1}H$	3.0

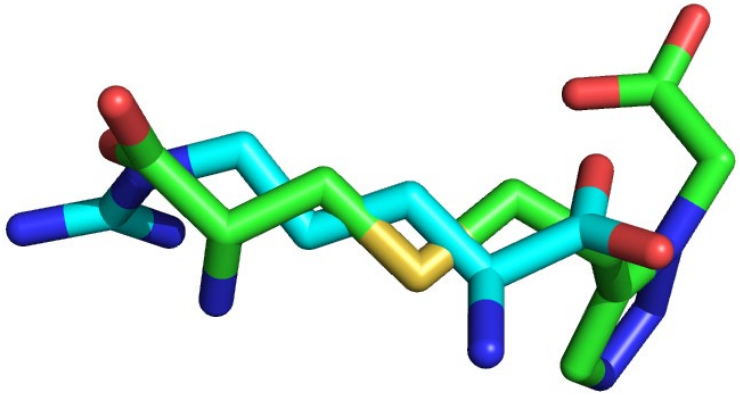
8.3.39 Analysis of docking performed for compound 3s



Residue	Atomic group	Distance (Å)

Structure docked out of position

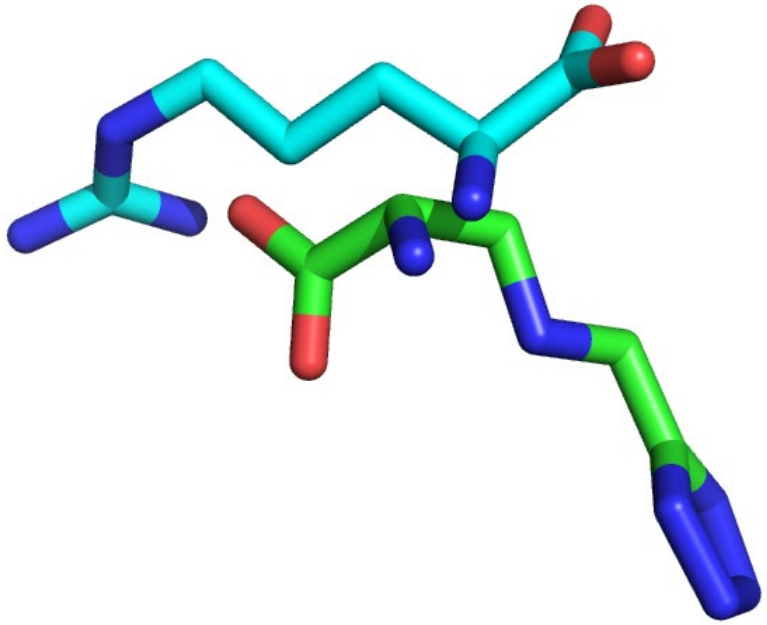
8.3.40 Analysis of docking performed for compound 3t



Residue	Atomic group	Distance (Å)

Structure docked out of position

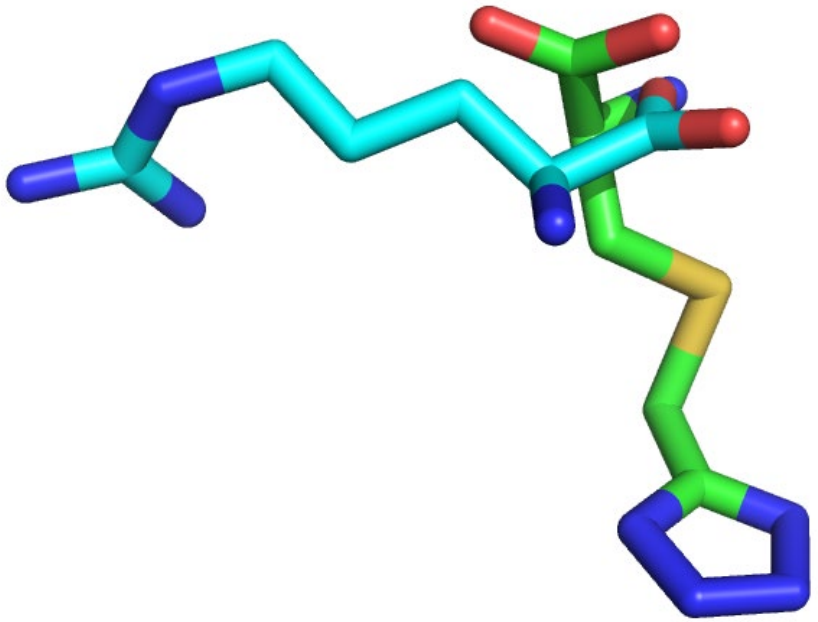
8.3.41 Analysis of docking performed for compound 3u



Residue	Atomic group	Distance (Å)

Structure docked out of binding pocket

8.3.42 Analysis of docking performed for compound 3v



Residue	Atomic group	Distance (Å)

Structure docked out of binding pocket