UNIVERSITY OF TARTU

FACULITY OF SCIENCE AND TECHNOLOGY

Institute of Chemistry

SIMONA SELBERG

MOLECULAR DESIGN OF INHIBITORS FOR RNA METHYLATION REGULATING ENZYMES

Master´s thesis (Chemistry)

30 EAP

Supervisor: Prof. Mati Karelson

Tartu 2017

Molecular design of inhibitors for RNA methylation regulating enzymes

The RNA m6A methylation plays crucial role in various physiological processes and therefore the development of chemical agents controlling this process has large biological and medical importance. In the present work, the results of the computational design of the inhibitory ligands for the enzymes regulating the RNA adenosine N6-methylation are presented. The structure of the enzyme-ligand complexes was established using the molecular docking and molecular dynamics approaches. The activity of the best predicted RNA m6A methyltransferase inhibitors was confirmed using enzyme inhibition and cell proliferation assays. These compounds are the first known RNA m6A methylation inhibitors and therefore of substantial interest for further biomedical studies.

RNA metüleerimist reguleerivate ensüümide inhibiitorite molekulaardisain

RNA m6A metüleerimisel on suur roll erinevates füsioloogilistes protsessides ning sellest tulenevalt on seda protsessi kontrollivate keemiliste agentide arendamine suure bioloogilise ja meditsiinilise tähtsusega. Käesolevas töös esitatakse RNA adenosiini N6-metüleerimist reguleerivate ensüümide inhibiitorite molekulaardisaini tulemused. Vastavate ensüüm-ligand komplekside struktuurid leiti molekulaarsildamise ja molekulaardünaamika meetodite abil. Parimate RNA m6A metüültransferaasi inhibiitorite aktiivsused leidsid eksperimentaalset kinnitust ensüümi inhibeerimise ja rakkude proliferatsiooni katsetes. Nimetatud ühendid on esimesed teadaolevad RNA m6A metüleerimise inhibiitorid ning pakuvad seetõttu suurt huvi edasistes biomeditsiinilistes uuringutes.

Contents

Abbreviations

^{3}H-SAM – tritiated S-adenosyl-L-methionine

3-meT – 3-methylthymidine

3-meU – 3-methyluracil

5mC – 5-methylcytosine

6mA – N6-methyladenine

AlkBH5 – AlkB family member 5

ATP – adenosine triphosphate

CpG – 5'-Cytosine- phosphate-Guanine-3'

DE – docking efficiency

DMEM – Dulbecco's modified eagle medium

DNA – deoxyribonucleic acid

DNMT – DNA methyltransferases

DTT – dithiothreitol

ECM – extracellular matrix

ELISA – enzyme-linked immunosorbent assay

ESC – embryonic stem cell

FBS – fetal bovine serum

FTO – fat mass and obesity-associated protein

HEK – human embryonic kidney cells 293

iCM – inducing functional cardiomyocyte

IOX1 – 5-carboxy-8-hydroxyquinoline

iPSC – induced pluripotent stem cell

J – pseudouracil

KLF4 – Kruppel-like factor 4

m6A – N6-methyladenosine

Mettl14 – methyltransferase-like 14

Mettl3 – methyltransferase-like 3

mRNA – messenger RNA

MYC – v-myc avian myelocytomatosis viral oncogene homolog

N-CDPCB – N-(5-Chloro-2,4-dihydroxyphenyl)-1-phenylcyclobutane-carboxamide

OCT4 – octamer-binding transcription factor 4

PBS – phosphate buffered saline

PFA – paraformaldehyde solution 4% in PBS

RMSD – root mean square deviation

RNA – ribonucleic acid

RNaseOUT – recombinant RNase Inhibitor

SAM – S-adenosylmethionine

SOX2 – SRY (sex determining region Y)-box 2

ss-RNA – single-stranded *RNA*

TBS – tris-buffered saline

Tris – tris(hydroxymethyl)aminomethane

WST 1 – tetrazolium salt

WTAP – Wilm's tumour-1-associated protein

YTHDC1 – YTH domain-containing protein 1

YTHDC2 – YTH domain-containing protein 2

YTHDF1 – YTH N6-methyladenosine RNA binding protein 1

YTHDF2 – YTH N6-methyladenosine RNA binding protein 2

YTHDF3 – YTH N6-methyladenosine RNA binding protein 3

ZIKV – Zika virus

1. Introduction

The regenerative medicine has today high expectations as related to the reversible epigenetic regulation of the gene expression in diverse tissues of the organism. Differently from the direct gene regulation initiated by the modification of DNA code, the epigenetic change in gene expression does not result from alteration in DNA sequence [1], thus allowing larger flexibility and control of this process. The classic epigenetic mechanisms include DNA methylation, histone modifications, but also chromatin remodeling and regulation mediated by non-coding RNAs. In past few years, the interest in the study of global modification patterns to both coding and noncoding RNA has significantly increased. This new area of research called epitranscriptomics is expected to give essential information for understanding human health and molecular disease pathology. The central mechanism of the epigenetic modification of RNA is the methylation of the nucleic acid bases. The most common RNA base modification is N6-methyladenosine (m6A).

Recently, it has been shown that m6A modifications of RNA affect its splicing, intracellular distribution, translation, and cytoplasmic degradation, playing thus a crucial role in regulating cell differentiation, neuronal signaling, carcinogenesis and immune tolerance [2]. The m6A presence in RNA is regulated by specific enzymes, i.e. the RNA methyltransferases, RNA methylases and RNA reader proteins. The regulation of these enzymes by suitable inhibitors or activators will thus enable the fine tuning of epitranscriptomic processes in biological and medical applications. Targeted development of small-molecule modulators (inhibitors and/or activators) of such enzymes involved in epitranscriptomic modulation would lead to compounds that control over the RNA methylation process, with potential outcomes for traditional and regenerative medicine. [3,4] In this study, we present the results of the computational molecular design of small-molecule inhibitors of enzymes involved in the regulation of RNA m6A methylation.

2. Literature overview

2.1 Stem cells and epigenetic gene regulation

Stem cells that can transdifferentiate to other types of cells have been highlighted as potentially transforming the whole regenerative medicine by enabling the direct repair or substitution of body tissues and organs [5]. However, there are several major obstacles in the way for their widespread use in clinical practice. First, because of their scarcity, it is very difficult to locate and separate the adult stem cells from tissues. Therefore, the embryonic stem cells (ESCs) are mostly used in laboratory and clinical investigations. This makes the research and clinical applications expensive and rises ethical problems with using the human embryos. [6] Furthermore, the differentiation of the stem cells into tissue cells is not strictly controlled and can lead to cancer [7,8].

A major breakthrough in the stem cells research was made by Yamanaka and coworkers [9] who demonstrated that cells with a gene expression profile and developmental potential similar to embryonic stem cells (ESCs) can be generated from mouse fibroblasts by using a cocktail of four transcription factors. These stem cells were called induced pluripotent stem cells (iPSCs), and the transcription factors were named "Yamanaka factors" (OCT4, SOX2, KLF4 and MYC). It was soon proven that using iPSCs technology, the skin fibroblasts and terminally differentiated cells of other adult tissues can be transformed into cells similar to ESCs. [10,11] Moreover, it appeared that it is also possible to transform the adult tissue cells into cells of other lineages by passing the phase of pluripotency. This possibility is schematically illustrated on the modified Waddington model of the cell reprogramming (Figure 1). The tissue specific defined culture can transform skin cells to become trophoblast, heart valve cells, photoreceptor cells, immune cells, melanocytes, and so forth. Extracellular matrix complexation with iPSCs enables generation of tissue organoids for organs of the body. Transdifferentiation has been reported for various cell types such as pancreatic beta cells, neurons, hepatocyte-like cells, and haematopoietic progenitor cells. The functional cardiomyocytes (iCMs) have been directly transformed from fibroblasts. [12]
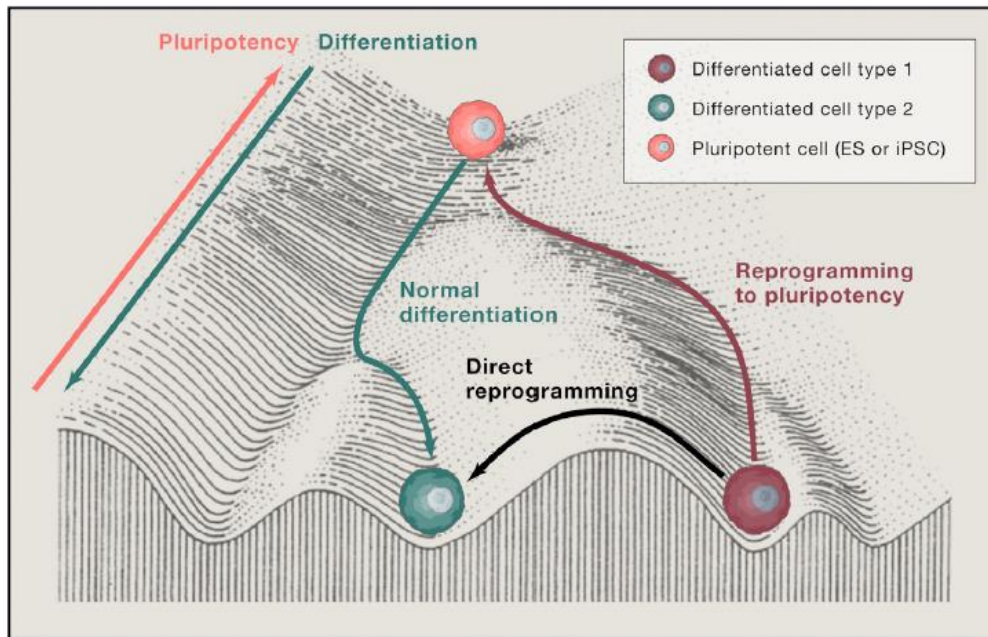
Figure 1. Modified Waddington Model for Cellular Reprogramming (from Srivastava *et al*) [13].

Important factors in the gene expression that regulate the cell pluripotency, differentiation or transdifferentiation are simple chemical modifications of DNA and histones in chromatin. These are known as the "epigenetic code" (Figure 2). [14] The main epigenetic modification of DNA is featured by methylation of cytosine residues in 5'-Cytosine-phosphate-Guanine-3' (CpG) sequence. The N-terminal tails of histone proteins are subject to a wide range of different modifications, including acetylation, methylation, phosphorylation and ubiquitylation. All of these chemical changes seem to have a substantial influence on chromatin structure and gene function, which differs depending on the type and location of the modification. [15] The DNA and histone lysine methylation essentially initiates the cell differentiation and development [16].

Whereas the epigenetic chemical modifications of DNA and histones have been relatively well investigated, the study of similar chemical modifications of RNA is still in its infancy [15]. Nevertheless, RNA chemical modifications have recently been identified to have an increasing, unprecedented and species-wide conserved impact on several critical cellular functions, such as proliferation, survival and differentiation, mostly through regulation of RNA stability [17,18,19]. The most abundant modification in eukaryotic messenger RNA is N6-methyladenosine (m6A). In addition to mRNA, this modification also appears in long non-coding RNAs [18] and microRNAs [19], thus covering the whole epitranscriptome. The m6A

modifications in RNA are essential for the development and functions of several tissues including liver, kidney and brain. In the adult brain, m6A modifications have been shown to affect neuronal functions and maturation. Given the major role of RNA modifications as controllers of a wide array of biological processes, it is not surprising that they have been associated with various pathologies such as several developmental and immunological disorders, obesity, diabetes, cancer as well as with cardiovascular, neuronal and infectious diseases. [20]
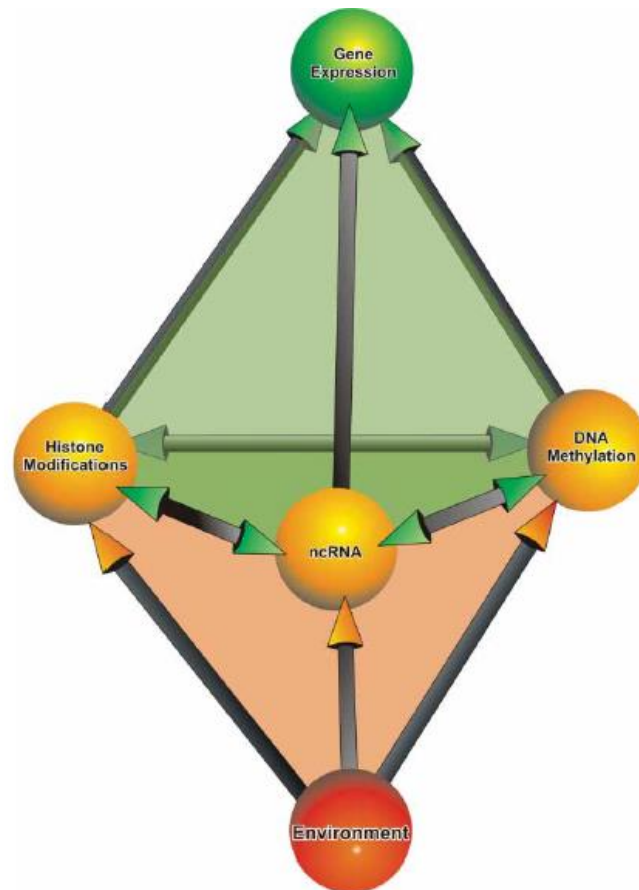

Figure 2. Epigenetic regulation in eukaryotes [14].

2.2 Methylated modifications of nucleic acids

The nucleic acid modifications play important biological regulatory roles, especially in the regulation of the gene expression. The most essential DNA, messenger RNA and long non-coding RNA modifications are the 5-methylcytidine with its oxidative derivatives and N6-methyladenine.
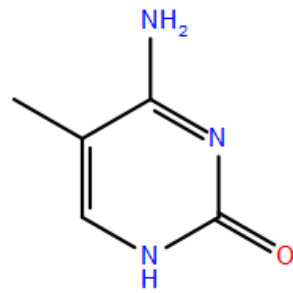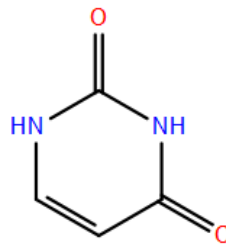
5-methylcytidine
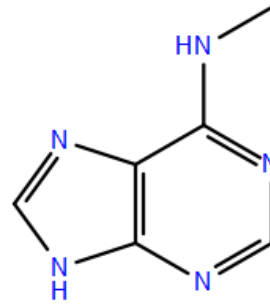
N6-methyladenine

Another important RNA modification is pseudouracil (J). [21]

Pseudouracil

The DNA methylation is generally associated with the regulation of gene expression. The chromatin structure is influenced by the methylation, but the effect may be dual. Namely, the DNA methylation in the promoter region of the nucleic acid is directly leading to the repression of transcription. DNA methylation may also attract the specific proteins ("readers") acting as repressors of the transcription. On the other hand, the DNA methylation in the gene body area has positive correlation with gene expression. [21]

The RNA chemical modifications do not affect nucleoside sequence in most cases, but are much more diverse and functionally versatile. This is due to the more complex post-transcriptional processing of the mRNA that involves alternatively joining introns and removing exons. Notably, chemical modification of RNA requires significant amounts of energy. The RNA methylation in cell is carried out by using S-adenosylmethionine (SAM) as the methyl donor. The synthesis of a SAM molecule requires the energy equal to hydrolyzing 12 to 13 ATP molecules. [22]

2.3 RNA methylation/demethylation

Today, more than 140 types of post-transcriptional chemical modifications of RNAs have been identified. The most abundant modification is m6A that accounts for about 50% of total methylated ribonucleotides and is present in 0.1%–0.4% of all adenosines in total cellular RNAs. [23] The N-methylation of the adenosine is a reversible process, catalyzed by specific enzymes (Figure 4). Those include the RNA methyltransferase enzyme complex Mettl3/Mettl14/WTAP consisting of three components: Mettl3 (methyltransferase-like 3), Mettl14 (methyltransferase-like 14), and WTAP (Wilm's tumour-1-associated protein), called also the "writer" enzyme; the RNA demethylases FTO (fat mass and obesity-associated protein) and AlkBH5 (AlkB family member 5), called "erasers". The fate of the RNA in post-transcriptomic processes is also directed by the "reader" enzymes that recognize specific m6A methylation in RNA. Several RNA reader enzymes have been identified, i.e. YTHDF1 (YTH N6-Methyladenosine RNA Binding Protein 1), YTHDF2 (YTH N6-Methyladenosine RNA Binding Protein 2) YTHDF3 (YTH N6-Methyladenosine RNA Binding Protein 3), YTHDC1 (YTH domain-containing protein 1) and YTHDC2 (YTH domain-containing protein 2). [24]

These three types of enzymes collectively coordinate the m6A RNA methylome in the eukaryotic cell.
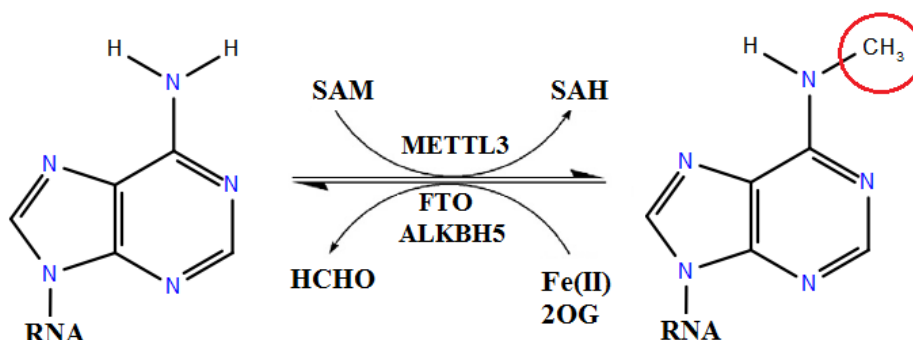

Figure 4. Dynamic and reversible m6A methylation in RNA (SAM – *S*-adenosyl-L-methionine; SAH – S-adenosyl-L-homocystein) [25].

2.3.1 Writers

The RNA methyltransferase enzyme complex Mettl3/Mettl14/WTAP is located in the nuclear speckle, and catalyzes the transfer of methyl groups from *S*-adenosyl-L-methionine (SAM) to

the specific adenines on the target RNAs. The reaction product is S-adenosyl-L-homocystein (SAH). [26]

Notably, Mettl3 and Mettl14 were demonstrated to be the catalytic subunits, whereas WTAP controls the process [26]. The hypothetic chemical mechanism on the nucleic acid m6A methylation has been suggested in the case of DNA methyltransferase DNMT1 (Figure 5).
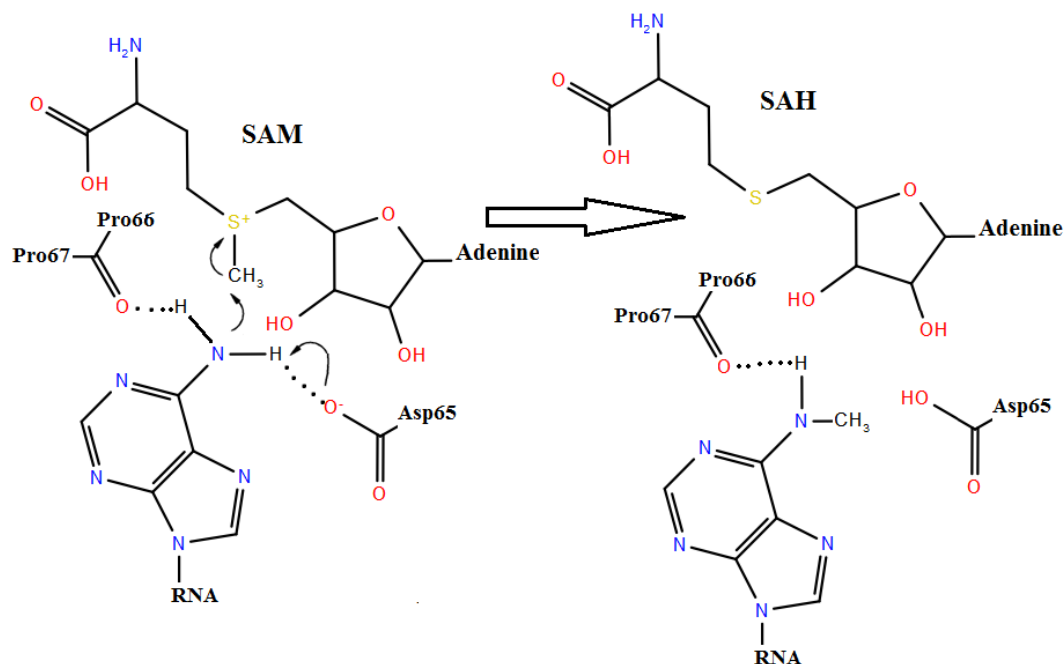

Figure 5. Hypothetic mechanism of the m6A methylation of DNA by DNA methyltransferase DNMT1. [27]

The catalytic centre of DNMT1 involves Asp65 and the peptide bond carbonyl oxygen between Pro66 and Pro67 residues [27]. From the principle of rationality in nature, it is feasible to assume a similar m6A methylation mechanism in the case of RNA (of course, with the applicable amino acid residues).

The crystal structures of the Mettl3/Mettl14 protein complex together with SAM and SAH ligands, respectively, have been published recently by several groups [26,28,29,30] (Figure 6). From these structures, it was established that in Mettl3 the catalytic cavity might be fenced in by three major loops: the two loops near the perimeter of the SAM binding site, and a third, larger loop that makes extensive contact with Mettl14 [29].

These crystal structures can be a basis for the rational target-based design of efficient small-molecule ligands, acting as the inhibitors of the RNA methylation.
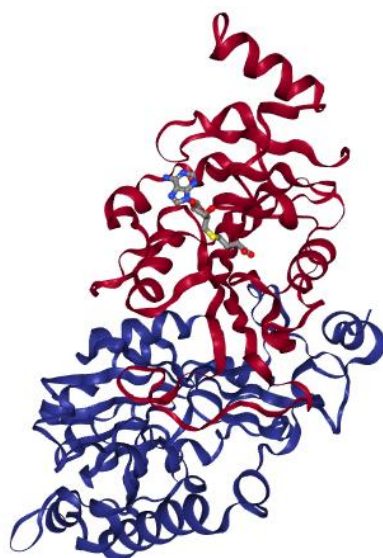
Figure 6. Crystal structure of the catalytic domain of Mettl3/Mettl14 complex with SAH by chain (pdb:5K7W) [29].

2.3.2 Erasers

Two enzymes have been identified as responsible for the RNA m6A demethylation. The first of these is the fat mass and obesity-associated protein FTO. It was shown that FTO demethylates m6A containing RNA efficiently *in vitro* [31]. The crystal structure of the FTO in complex with 3-methylthymidine (3-meT) was first reported by H*an* and colleagues in 2010 (pdb:3LFM) [32]. The structural analysis indicated that the hydrogen-bonding interaction between FTO and two carbonyl oxygen atoms in 3-meT or 3-methyluracil (3-meU) enables FTO to distinguish 3-meT or 3-meU from other nucleotides. Moreover, FTO contains an amino-terminal AlkB-like domain and a carboxyl-terminal domain with a novel fold, which interact with each other and involve in FTO catalytic activity. The difference between FTO and other AlkB family members is that FTO has an extra loop covering one side of the conservative jelly-roll motif, which can selectively compete with the unmethylated strand of the DNA duplex for binding to FTO. [31,32]

Thereafter, a series of crystal structures FTO in complex with different inhibitory ligands have been published [33,34,35] that enable the target-based computational design of novel efficient inhibitors. An exemplary crystal structure of the complex between FTO and its inhibitor meclofenamic acid is given on Figure 7.
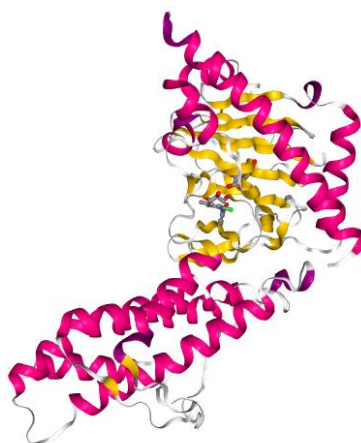
Figure 7. Crystal structure of FTO bound to a selective inhibitors meclofenamic acid, N-oxalylglycine and glycerol by secondary structure (pdb:4QKN) [35].

Recently the crystal structures of another RNA demethylase AlkBH5 with various ligands were published [36]. The structure of AlkBH5 in complex with α-ketoglutarate is presented on Figure 8.
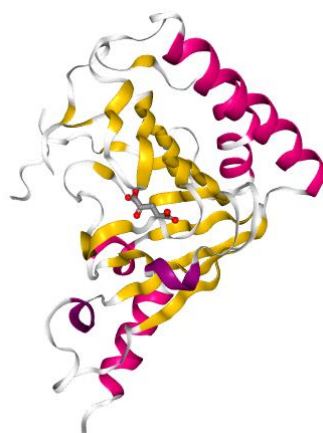
Figure 8. Crystal structure of human ALKBH5 in complex with α-ketoglutarate by secondary structure (pdb:4NRO) [36].

2.3.3 Readers

Decoding of the information engraved in the methyl code is executed by a special category of proteins called m6A readers, which can function in several ways. The presence of m6A in a target RNA alters the secondary or tertiary structure of its specific domain, which promotes binding of the m6A reader. This, in turn, locally remodels the RNA-protein interface to create another binding site for a second protein. Binding of the latter component (which might be involved in mRNA biogenesis, splicing, or decay) may thereby alter the fate of the target RNA.

Alternatively, the altered local secondary/tertiary structure in the target RNA induced by m6A may remodel the already bound protein(s) at that site to facilitate the binding of the m6A reader. Either of these outcomes can subsequently influence the processes that modulate the fate of a target mRNA, such as its export, translation, or stability. [2]

The crystal structures of all three known RNA m6A readers (YTHDF1 [37], YTHDF2 [38], YTHDC1 [39]) together with various ligands and/or RNA have been published (Figure 9).

(a)   (b)   (c)
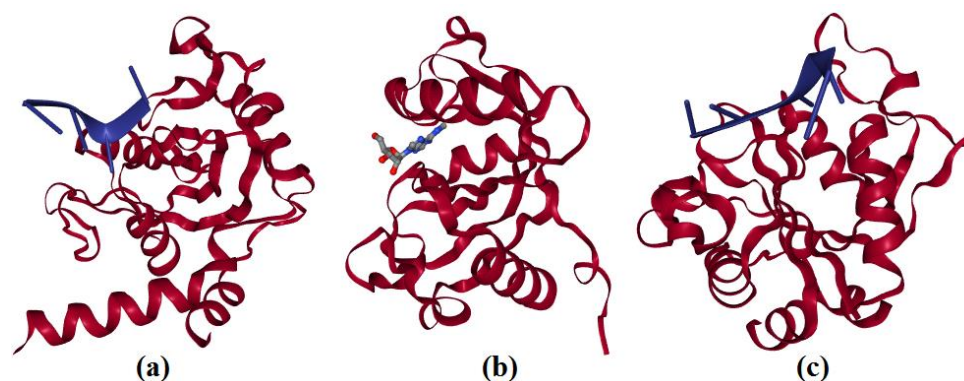
Figure 9. Crystal structure of RNA m6A readers with ligands. Crystal structure of YTHDF1 YTH domain in complex with 5mer m6A RNA (pdb:4RCJ) [37] (a); Crystal structure of YTH-YTHDF2 in complex with m6A (pdb:4RND) [38] (b); Crystal structure of YTHDC1 YTH domain in complex with 5mer m6A RNA (pdb:4R3I) [39] (c).

2.4 RNA methylation/demethylation inhibitors and activators

Very little is presently known about the inhibitors or activators of the RNA methylation or demethylation enzymes.

Using a structure-based virtual screening, Chen *et al* [40] found that rhein is an inhibitor of FTO.
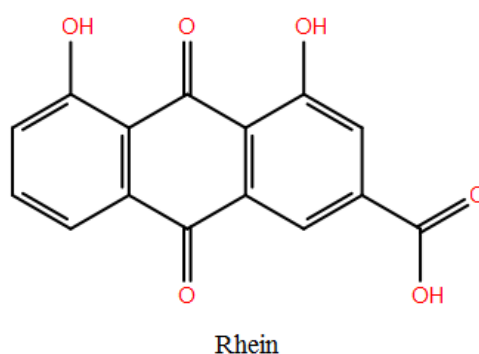
Rhein

Rhein can increase the cellular contents of m6A in mRNA by binding to the FTO active site and preventing FTO from binding to m6A substrate [40]. Rhein displays some selectivity on AlkB subfamily too. Meclofenamic acid is another highly selective inhibitor of FTO by competition for FTO binding with dm6A-containing substrate. However, MA does not bind to AlkBH5 in vitro and it cannot inhibit AlkBH5 demethylation activity, suggesting its target selection on FTO over AlkBH5 [41]. Moreover, citrate that adopts different binding patterns with FTO and AlkBH5 can act as an inhibitor for these human m6A demethylases [42]. Furthermore, IOX3 that covalently binds to Cys200 residue positioned outside of the active site of AlkBH5 also can be used as a small molecule inhibitor of AlkBH5 [43]. N-(5-Chloro-2,4-dihydroxyphenyl)-1-phenylcyclobutane-carboxamide (N-CDPCB) is found to be another inhibitor of the FTO. The crystal structure of human FTO with N-CDPCB reveals a novel binding site for the FTO inhibitor and defines the molecular basis for recognition by FTO of the inhibitor. The identification of the new binding site offers new opportunities for further development of selective and potent inhibitors of FTO. [44]

Up to date, no RNA methyltransferase or reader enzyme inhibitors have been reported. Therefore, the main objective of our present research has been to discover such inhibitors and/or enzyme activators.

3. Methodology

3.1 Molecular docking

Three most important enzymes involved in the RNA m6A methylation/demethylation were chosen for finding their efficient ligands using molecular modeling. Those were the RNA methyltransferase complex Mettl3/Mettl14, RNA demethylase FTO and RNA m6A reader YTHDF1.

Several crystal structures of the Mettl3/Mettl14 complex are available at the Protein Data Bank [28], reported by different groups [26,28,29,30]. We selected the structure of the Mettl3/Mettl14 complex with the S-adenosyl-L-homocysteine (SAH) as describing the potential target binding site for a small-molecule inhibitor. The crystal structure of this complex (pdb:5K7W) had been measured by X-ray diffraction with resolution 1.65 Å. It has been suggested that Mettl3 may be the only active methyltransferase in the heterodimeric complex. [29]

The crystal structure of the RNA demethylase FTO has been published as for the pure protein [32] and together with different ligands [33,34,35]. Crystal structure of FTO in complex with 5-carboxy-8-hydroxyquinoline (pdb:4IE4) was chosen for further molecular docking modeling. The structure had been determined using X-ray diffraction with the resolution 2.5 Å. [33] The choice of this RNA demethylase was made on the basis of the availability of the ligand compounds.

In the case of RNA m6A readers, most experimental information is known for the YTHDF1 enzyme. Therefore, the crystal structure of YTHDF1 together with pentamer m6A RNA was used for the molecular docking modeling (pdb:4RCJ). The X-ray diffraction structure of the protein had been measured with the resolution 1.6 Å. [37]

The raw crystal structures were corrected and hydrogen atoms were automatically added to the protein using Schrödinger's Protein Preparation Wizard of Maestro 10.7 [45].

AutoDock Vina 1.1.2 [46] was used for the docking studies to find out binding modes and binding energies of ligands to the receptor. The number of rotatable bonds of ligand was set by default by AutoDock Tools 1.5.6 [47]. However, if the number was greater than 6, then some of rotatable bonds were made as non-rotatable, otherwise calculations can be inaccurate. The active site, was surrounded with a grid-box sized $65 \times 65 \times 65$ points with spacing of 1.0 Å.

The AutoDock 4.2 force field [47] was used in all molecular docking simulations. The docking efficiencies (DE) were calculated as follows

$$DE = \frac{\Delta G_{dock}}{N} \tag{1}$$

where ΔG_{dock} is the docking free energy and N – the number of non-hydrogen ("heavy") atoms in the ligand molecule.

The structure of ligand molecules was optimized using the density functional theory B3LYP method [48] with 6-31G basis set.

The molecular dynamics simulations were carried out using Desmond simulation package of Schrödinger LLC. [49] The NPT ensemble with the temperature 300 K and pressure 1 bar was applied in all runs. The simulation lengths were 10 ns and 50 ns with relaxation time 1 ps. The OPLS_2005 force field parameters were used in all simulations [50]. The long range electrostatic interactions were calculated using the Particle Mesh Ewald method [51]. The cutoff radius in Coloumb´ interactions was 9.0 Å. The water molecules were described using SPC (simple point charge) model. [52] The Martyna-Tuckerman-Klein chain coupling scheme [53] with a coupling constant of 2.0 ps was used for the pressure control and the Nosé-Hoover chain coupling scheme [53] for the temperature control. Non-bonded forces were calculated using an r-RESPA integrator where the short range forces were updated every step and the long range forces were updated every three steps. The trajectories were saved at 4.8 ps intervals for analysis. The behavior and interactions between the ligands and enzyme were analyzed using the Simulation Interaction Diagram tool implemented in Desmond molecular dynamics package. The stability of molecular dynamics simulations was monitored by looking on the root mean square deviation (RMSD) of the ligand and protein atom positions in time. For a system involving N atoms, RMSD is defined as follows:

$$RMSD = \sqrt{\frac{1}{N}\sum_{a=1}^{N}\left(\vec{r}_{aj} - \vec{r}_{ai}\right)^2} \tag{2}$$

where $\vec{r}_{ai}$ and $\vec{r}_{aj}$ are the position vectors of atom *a* at the consecutive time moments *i* and *j*, respectively.

3.2. Experimental

3.2.1. Cell culture

HEK-293 (ATCC®) cells were cultured in DMEM (Gibco®), supplemented with 10% FBS (Gibco®) and penicillin-streptomycin (Gibco®) at 37°C and 5% CO_2. For experiments conducted in 96-well plate format, the surface was coated with fetal bovine fibronectin (Thermo Fisher) 10 µg/ml.

3.2.2. Protein production for the enzymatic assay

Mettl3/Mettl14 complex was produced in HEK-293 cells, plated in 10 cm culture plates. Approximately 2.7 million cells were plated per culture plate and incubated until reaching 70% confluency. Then the cells were co-transfected with 25 µg of pcDNA3/Flag-Mettl3 (Addgene) and 25 µg of pcDNA3/Flag-Mettl14 (Addgene) vectors using Lipofectamine® 2000 (Invitrogen). Cells were then incubated for 48 h.

The Mettl3/Mettl14 complex was purified using the ANTI-FLAG® M2 Affinity Gel (Sigma-Aldrich). The complex was eluted with 150 ng/ml 3x FLAG® peptide (Sigma-Aldrich).

3.2.3. Mettl3/Mettl14 enzyme inhibition reaction

All enzyme inhibition experiments were conducted in reaction buffer (20 mM Tris pH 7.5, 1 mM DTT, 0,01% Triton™ X-100, 40 U/100 µl buffer RNaseOUT™ (Invitrogen)). The reaction mixture contained 200 nM unmethylated N^6-adenine ss-RNA probe with a biotin tag (Integrated DNA Technologies), 500 nM ^{3}H-SAM (Perkin Elmer) and 50 mM Mettl3/Mettl14 complex.

Reaction was incubated for 20 h at 21°C on shaker. Then the reaction mixture was transferred to wells in streptavidin-coated 96-well plate (Perkin Elmer) and incubated 1 h at room temperature. After that, the plate was washed with sterile TBS 3x, and then 100 µl/well of scintillation fluid was added. The plate was incubated 30 min at room temperature and the results were obtained using 1450 MicroBeta® liquid scintillation counter (Wallack). The scintillation counts were proportional to amount of methylated RNA. [54]

3.2.4. Tetrazolium salt WST-1 assay

Tetrazolium salt WST-1 (Roche®) assay was used to quantify the cell proliferation. The cells were cultured in 96-well plates, coated with fibronectin, with 10 000 cells seeded per well and incubated for overnight. Small molecules were introduced to the fresh culture media on the next day and incubated for 24h before the samples were collected for assays.

The tetrazolium salt WST-1 is cleaved by mitochondrial dehydrogenases to form formazan in viable cells. The amount of the formazan product produced following the addition of WST-1 is directly related to the number of viable, metabolically active cells, enabling thus to quantify the cell proliferation in time.

The WST-1 reagent was added in 1:10 dilution to the media (10 µl to 100 µl media) of cells treated with small molecules for 24 h. After addition of WST-1, reagent cells were placed into cell culture incubator at 37°C, 5% CO_2 for 2 h. The spectroscopic readouts on the amount of formazan were taken at 1 h and 2 h time points with an ELISA reader at the wavelength 450 nm, and with a reference at 620 nm.

3.2.5. Imaging

After cell proliferation assay, HEK-293 cells were fixed with 4% PFA for 20 min, washed with PBS five times and stored at 4°C in PBS. Imaging was done with AxioCam ERc5S (Zeiss) camera at 40x magnification.

4. Results and Discussion

4.1 Molecular docking

4.1.1. Mettl3/Mettl14

The crystal structure of the Mettl3/Mettl14 complex with SAH (pdb:5K7W) [29] was chosen for the molecular modeling of potential enzyme inhibitors. The binding site of SAH was selected as target area for potential Mettl3/Mettl14 RNA methyltransferase complex inhibitors. As reported by Wang *et al* [29], there are several distinct regions of probable interactions between the ligand and enzyme. As confirmed by our molecular docking calculations, the amino group of the adenosyl fragment of SAH is hydrogen bonded with Asp377 of the Mettl3 (cf. Figure 10). The binding is further supported by another bond between the adenine N1 atom and an adjacent peptide bond NH group. The adenine ring is sandwiched between Phe534 and Asn549, while many polar contacts help to hold the hydroxyl groups on the ribose as well as the amino and carboxyl groups of SAH [28]. The terminal amino group of SAH is acting as hydrogen bond donor to the Asp395 of the catalytic center of enzyme.
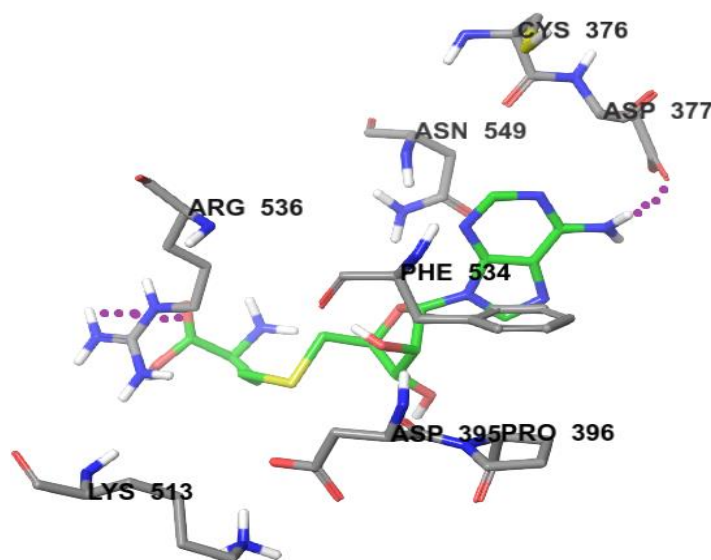

Figure 10. The binding site of SAH.

Based on this structure we proceeded with the search of effectively bound small molecule fragments. The first group of fragments consisted of various substituted purines. In variance to the crystal structure of SAH itself, these compounds tend to be bound to the region of the Mettl3 protein involving Asp395, Phe534, Arg536 and Asn539 (Figure 11).
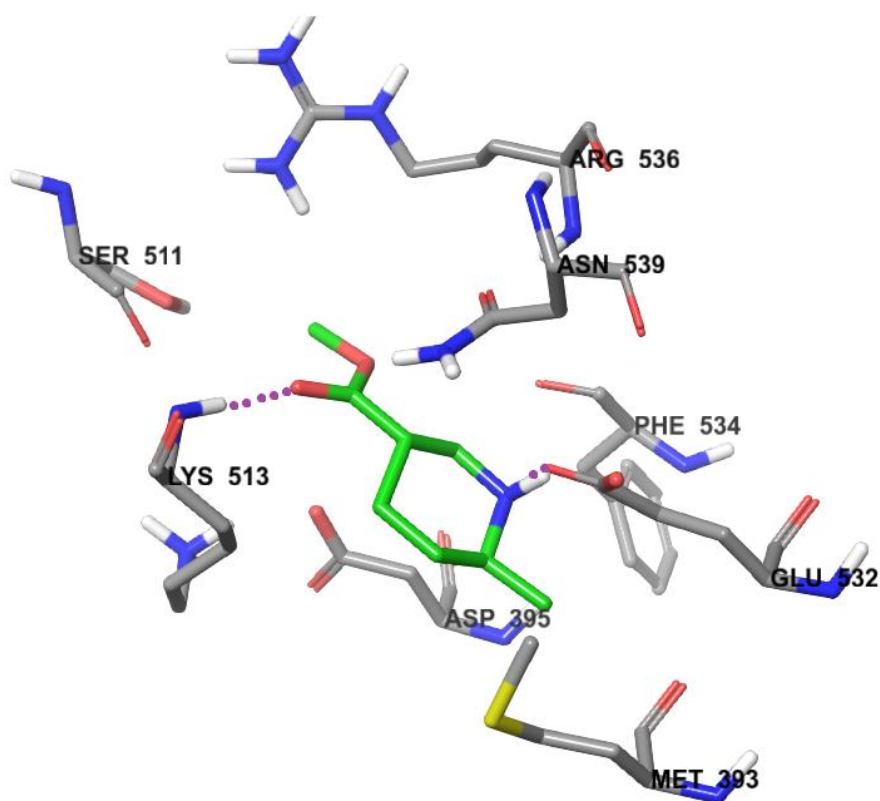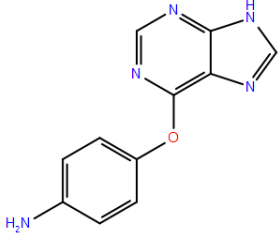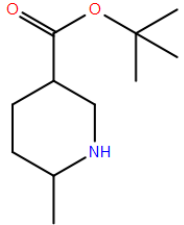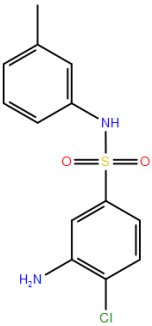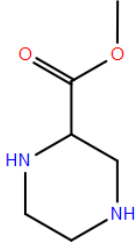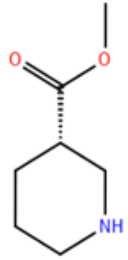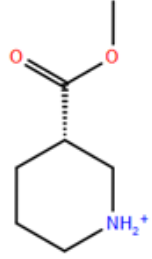
Figure 11. The binding site of the compound **1**.

The second group of fragments were selected based on the configuration of the protein residues, hydrogen bonded to the tail part of SAH. A virtual screening on ZINC [55] and DrugBank 4.0 [56] databases was carried out using nitrogen-containing heterocycles as base structures. Remarkably, we found a series of fragments with piperidine and piperazine rings having exceptionally high docking efficiencies. The docking free energies and docking efficiencies of the best fragment compounds are given in Table 1.

Table 1. The compounds with the highest docking efficiencies to Mettl3/Mettl14 complex.

No.	Compound structure	ΔG (kcal/mol)	DE
1		-6.98	0.64

2		-7.96	0.47
3		-8.67	0.62
4		-9.53	0.50
5		-5.82	0.58
6		-6.92	0.69
7		-7.41	0.67

4.1.2 FTO

The crystal structure of the FTO in complex with 5-carboxy-8-hydroxyquinoline IOX1 (pdb:4IE4) [33] was chosen for the prediction of potential efficient ligands using molecular docking modeling. The ligand IOX1 was removed from the complex in order to proceed with the search of novel ligands. The catalytic centre of the protein involves bivalent transition metal ion, either Mn^{2+}, Fe^{2+}, Ni^{2+} or Zn^{2+}. In our molecular docking simulations Zn^{2+} was used. The amino acid residues of the protein Asp233, Tyr295, Arg316 and Ser318, Arg322 were responsible for specific interactions between the protein and ligand (Figure 12).
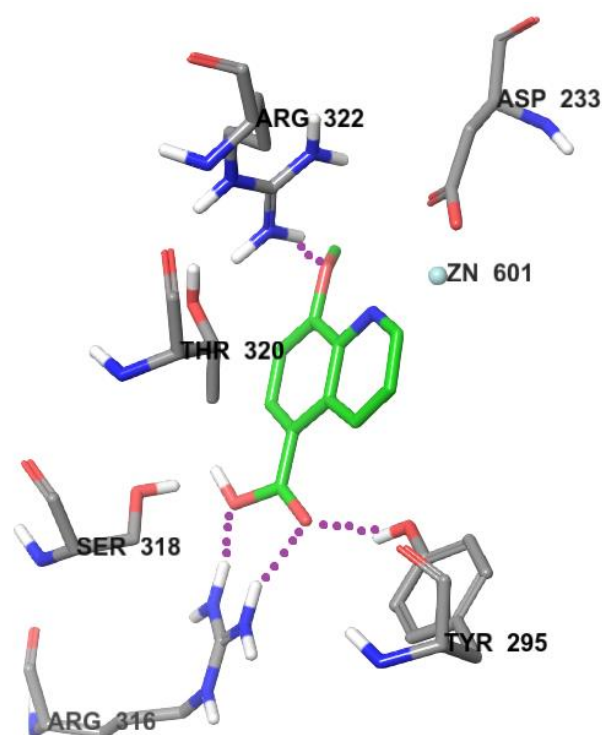
Figure 12. The binding site of the compound **10**.

A virtual screening on ZINC [55] database was carried out using the best known FTO inhibitors (I – V) as templates [57].

I   II   III
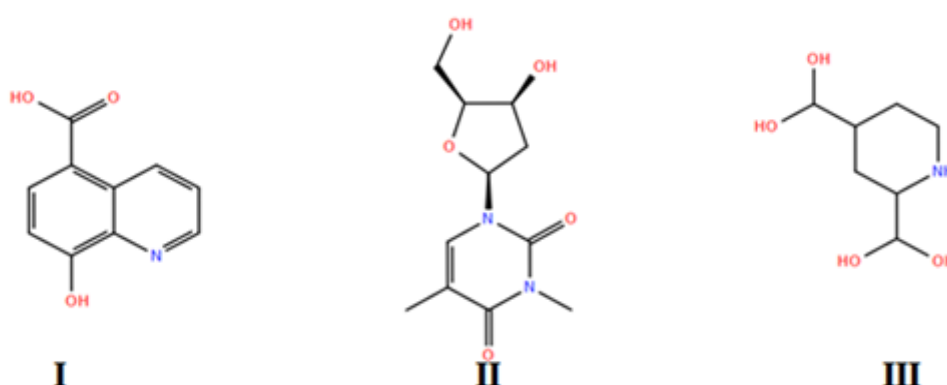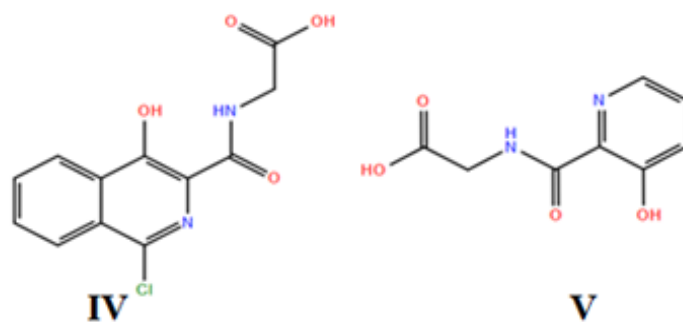
The docking free energies and docking efficiencies of the best predicted inhibitors are given in Table 2.

Table 2. The compounds with the highest docking efficiencies to FTO protein.

No.	Compound structure	ΔG (kcal/mol)	DE
8		-8.78	0.49
9		-7.87	0.44
10		-7.17	0.48
11		-7.37	0.53
12		-9.45	0.47

It has been shown that knockout silencing of RNA demethylases FTO and AlkBH5 leads to the decrease of the Zika virus (ZIKV) production in cells [58]. Therefore, we suggest that apart from the study of inhibitory activity of the predicted compounds in FTO enzymatic and cellular assays, the measurements of the ZIKV inhibition could be also of significant interest.

4.1.3 YTHDF1

In order to develop efficient RNA m6A reader protein YTHDF1 ligands using molecular docking modelling, we proceeded from the crystal structure of YTHDF1 together with pentamer m6A RNA (pdb:4RCJ) [37]. The RNA pentamer was removed in all simulations. A virtual screening on ZINC [55] and DrugBank 4.0 [56] databases was carried out using N-6-methyladenine (VI) as template.

VI

The active centre of the enzyme that embeds the methylated adenosine is located in a cavity surrounded by the following amino acid residues: Lys395, Ser396, Tyr397 and Asp507. The purine ring of potential inhibitors is sandwiched between the Tyr470 and Tyr411 residues of the protein (Figure 13).
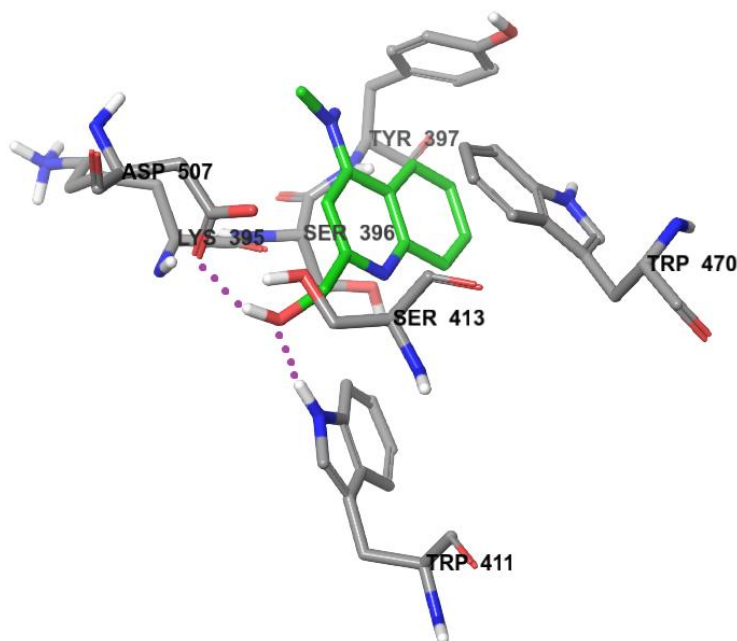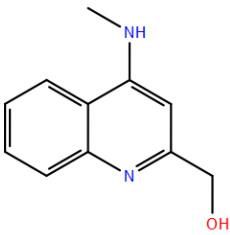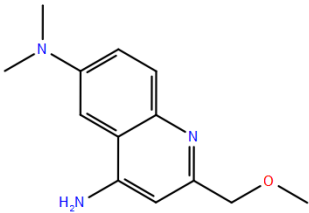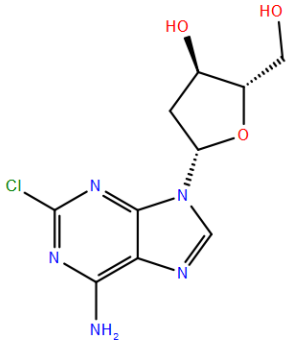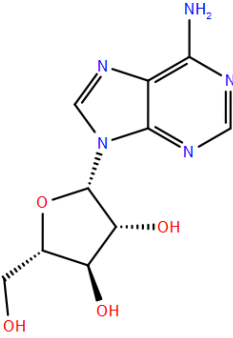
Figure 13. The binding site of the compound **13**.

The best predicted RNA m6A reader YTHDF1 inhibitors (Table 3) will be experimentally tested in enzymatic and cellular assays. Notably, compound **15** (Cladribine) is known as an antiproliferative agent against hairy cell leukemia [59]. Therefore it would be highly interesting to establish if its mode of action is the YTHDF1 inhibition. This would enable efficient structure based design of novel anti-leukemia drugs.

Table 3. The compounds with the highest docking efficiencies to the RNA m6A reader YTHDF1.

No.	Compound structure	ΔG (kcal/mol)	DE
13		-8.01	0.57
14		-6.61	0.39
15		-6.21	0.33
16		-5.92	0.31

4.2 Molecular dynamics

The molecular dynamics simulations were thereafter carried out for two compounds, representing the two different promising scaffolds, the compounds **2** (4-(9H-purin-6-yloxy)aniline) and **6** (piperidine-3-carboxylate), respectively.

In the case of compound **6** three simulations with the lengths 10 ns, 20 ns and 50 ns were carried out. The 50 ns simulation involved the compound in protonated form **7** that is present at physiological pH values. In this simulation, after 30 ns ligand an instability was observed in the ligand position RMSD (Figure 14a). Therefore only the first 25 ns were taken into account in the further trajectory and energy analysis (Figure 14b).
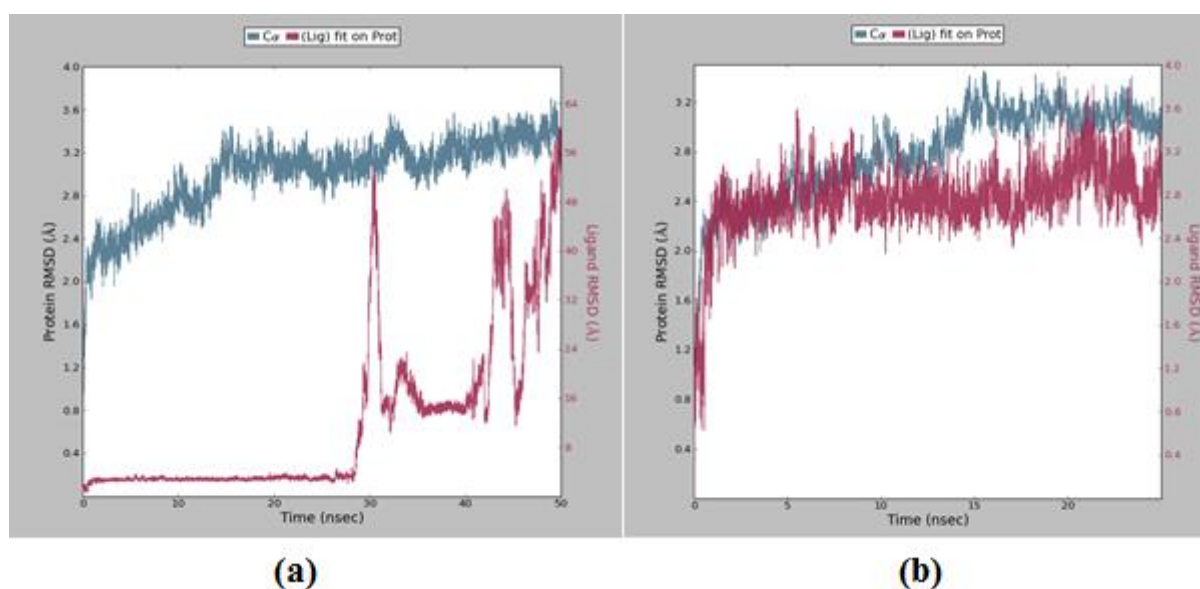
Figure 14. The protein and ligand position root mean square deviation (RMSD) plot against time in the case of the Mettl3/Mettl14 complex with compound **7** for 50 ns (a) and 25 ns (b) runs.

The simulation interactions diagram (Figure 15) indicates that the most important interactions are hydrogen bonds between the ligand and Asp395 and Lys513 residues of the Mettl3/Mettl14 protein. Another hydrogen bond is detected between the ligand and the carbonyl oxygen of the peptide bond between Leu533 and Phe534 (Figure 16).

Figure 15. 2D profile data for the compound **7** binding to Mettl3/Mettl14 complex: (a) interactions occurring more than 30 % of the simulation time; (b) timeline of interactions and contacts.
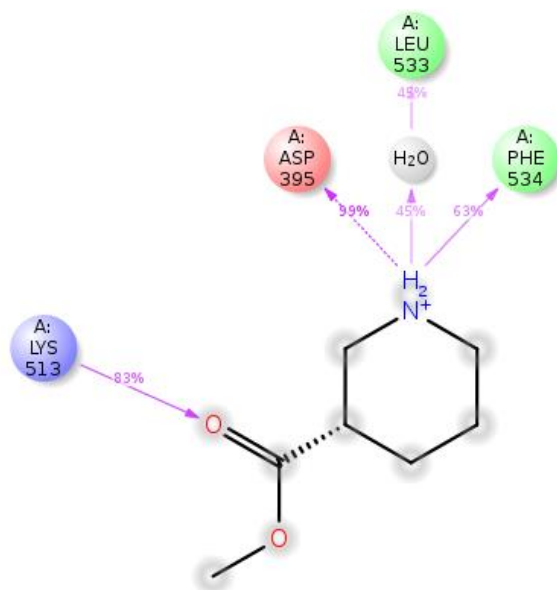
Figure 16. Interaction diagram between the compound **7** and Mettl3/Mettl14 complex.

In the case of compound **2** the simulation with the length 50 ns was carried out. The rather constant plot of position RMSD for both the protein and ligand after 10 ns indicates that the system has achieved steady state.
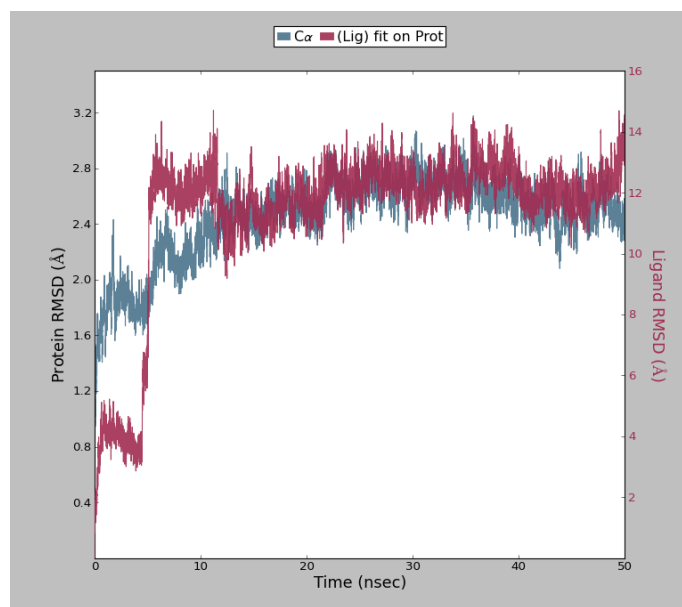
Figure 17. The protein and ligand position root mean square deviation (RMSD) plot against time in the case of the Mettl3/Mettl14 complex with compound **2** for 50 ns.

The protein-ligand interaction profile for the compound **2** (Figure 18) indicates that the main interactions with Mettl3/Mettl14 involve the Arg379, Glu413 and Asn549 residues of the protein. Interestingly, the interaction of Arg379 with ligand involves the contributions from the hydrogen bonding as well as from the ionic and hydrophobic interactions. The Glu413 and Asn549 residues are involved in hydrogen bonding as acceptors.

(a)

(b)

Figure 18. 2D profile data for the compound **2** binding to Mettl3/Mettl14 complex. Interactions occurring more than 30 % of the simulation time (a). Timeline of interactions and contacts (b).

These observations are depicted in the interaction diagram given in Figure 19.
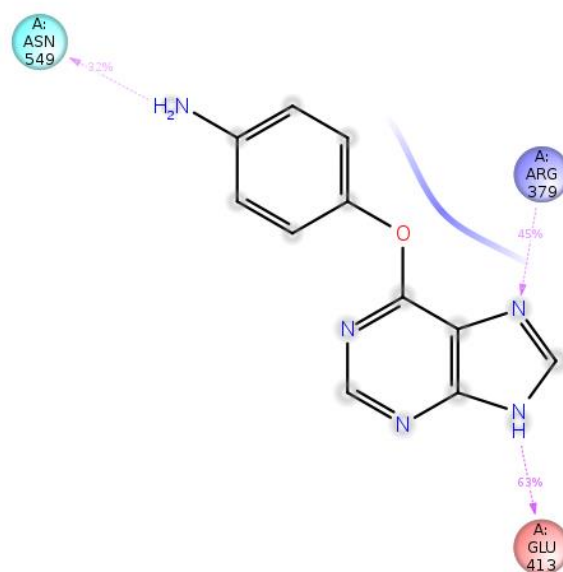
Figure 19. Interaction diagram between the compound **2** and Mettl3/Mettl14 complex.

In order to further specify the nature of the ligand-protein interactions, the Molecular Mechanics/Poisson–Boltzmann Surface Area (MMPBSA) [60] binding energy calculations were carried out using data from molecular dynamics simulations. In MM/PBSA, the free energy of a state (ligand or enzyme) is estimated from the following sum

$$G = E_{bnd} + E_{el} + E_{vdW} + G_{pol} + G_{np} - TS \qquad (3)$$

where the first three terms are standard molecular mechanics energy terms from bonded (bond, angle and dihedral), electrostatic and van der Waals interactions. G_{pol} and G_{np} are the polar and non-polar contributions to the solvation free energies. G_{pol} is typically obtained by solving the Poisson-Bolzmann equation or by using the generalized Born (GB) model (giving the MM/GBSA approach), whereas the non-polar term is estimated from a linear relation to the solvent accessible surface area. The last term in above equation is the absolute temperature, T, multiplied by the entropy, S, estimated by a normal-mode analysis of the vibrational frequencies. The results for the lowest binding energy states are given in Table 4.

Table 4. The MMPBSA energies for the Mettl3 complexes with the compounds **2** and **7** (kcal/mol).

Energy term	**2**	**7**
ΔE_{vdW}	-23.72	-8.48
ΔE_{el}	-16.74	-25.26
ΔE_{pol}	22.53	25.56
ΔE_{np}	-8.36	-7.37
ΔG_{bind}	-27.95	-15.51

The total binding energies are large for both compounds and predict significant biological effects. However, it is interesting to compare the contributions arising from different physical interactions to the binding energies. For the compound **7**, the electrostatic interaction energy between the ligand and enzyme and the polar (de)solvation energy almost exactly cancel each other. The binding energy is thus primarily determined by large van der Waals and nonpolar solvation terms. Thus, this relatively small compound is strongly bound to the enzyme due to specifically oriented van der Waals and nonpolar solvation (lipophilic) interactions.

In the case of the compound **2**, the main contributions to the binding free energy with protein are the electrostatic energy and the van der Waals interaction energy (Table 4). The latter is considerably higher than in the case of smaller ligand **7**. Notably, the polar (de)solvation energy and the nonpolar solvation energies are similar for both compounds studied.

4.3. Mettl3/Mettl14 complex inhibition assay results

4.3.1 Enzymatic assay results

The enzymatic assay studies on Mettl3/Mettl14 complex were carried out using the tritiated ^{3}H-SAM (Perkin Elmer) reaction with the unmethylated RNA. The preliminary results were obtained for the studied inhibitory compounds at 10nM and 10μM concentrations. The estimated IC50 values are given in Table 5.

Table 5. The inhibition of Mettl3/Mettl14 complex in enzymatic assay.

Compound	%inh (10nM)	IC50 (nM)
1	16.7	30
2	24.0	21
3	N/A[a]	N/A
4	4.1	122
5	15.5	32

[a] Toxic compound

These results reveal the inhibitory activity of compounds **1**, **2** and **5** at nanomolar level and make them prospective leads for the further structural optimization of the RNA methyltransferase Mettl3/Mettl14 inhibitors.

4.3.2 Cellular assay results

Within this work, the activity of the designed Mettl3/Mettl14 inhibitors on the cell proliferation was studied using WST-1 assay. It has been shown that the genetic inactivation or depletion of mouse and human Mettl3 enzyme promotes self-renewal and hinders differentiation of the embryonic stem cells [61]. Therefore, it is expected that the inhibition of the Mettl3 in somatic cells would lead to the increase of the proliferation. In Figure 20, the concentration dependence of the HEK-293 cell proliferation extent is given for the compounds **1**-**5** as measured by the absorbance in the WST-1 assay.

The compound **1** (methyl 6-methylpiperidine-3-carboxylate) has caused significant increase of the proliferation of cells at all measured concentration (100 nM, 1 nM and 1 pM) as compared to non-treated control (Figure 20). The amount of the proliferated cells increases in time but remains nearly constant after two hours. A notable increase of the cell proliferation at just 1 pM inhibitor concentration reflects very high efficiency of this compound and can be taken as the lead compound for the further structure optimization.

A similar behavior was observed in the case of compound **2** (Figure 21). As this compound belongs to the different scaffold, it can be also used as a lead compound for the further optimization of the Mettl3/Mettl14 complex inhibitors.

The compound **3** (tert-butyl 6-methylpiperidine-3-carboxylate) has unusual concentration dependence of the WST-1 assay readout, having a maximum at the 1 nM concentration (Figure

22). Whereas this compound is also an efficient proliferation enhancer, it is obviously toxic at higher concentrations. The compounds **4** (3-Amino-4-chloro-N-(m-tolyl)benzenesulfonamide) and **5** (methyl piperazine-2-carboxylate) have little effect on the HEK-293 cell proliferation at all concentrations studied (Figure 23 and Figure 24).

Figure 20. The proliferation efficiency of methyl 6-methylpiperidine-3-carboxylate **1** on HEK-293 cells.
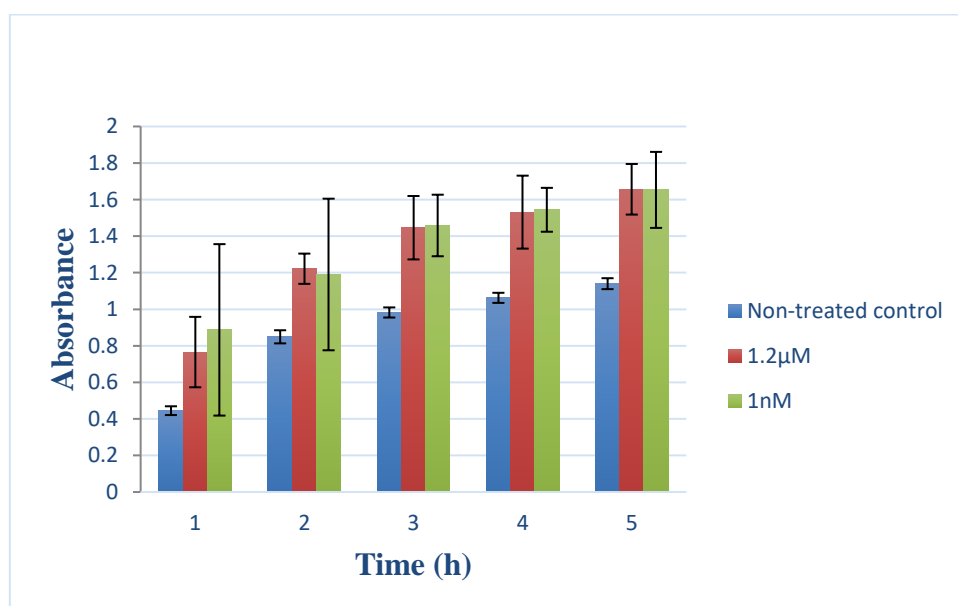
Figure 21. The proliferation efficiency of 4-(9H-purin-6-yloxy)aniline) **2** on HEK-293 cells.
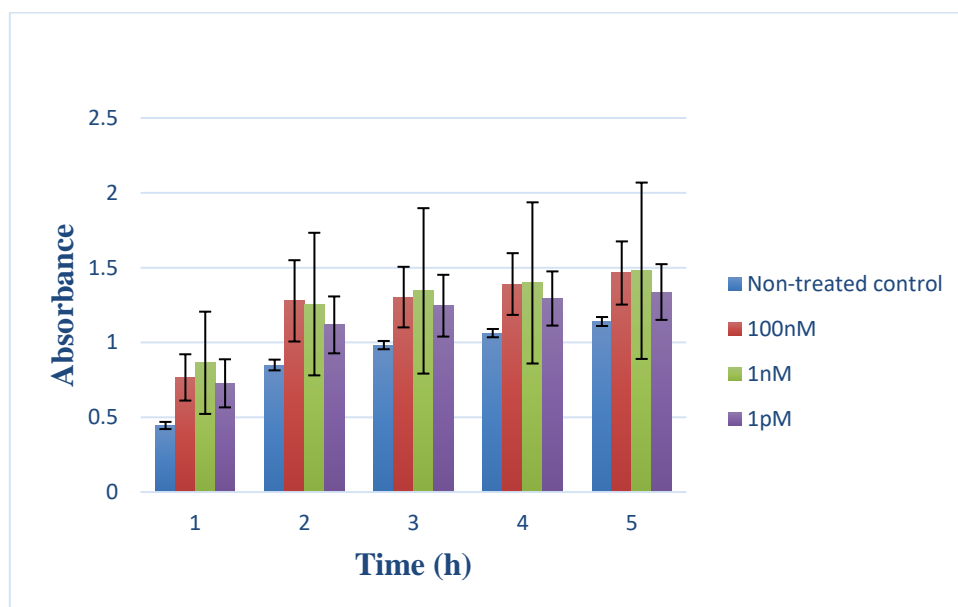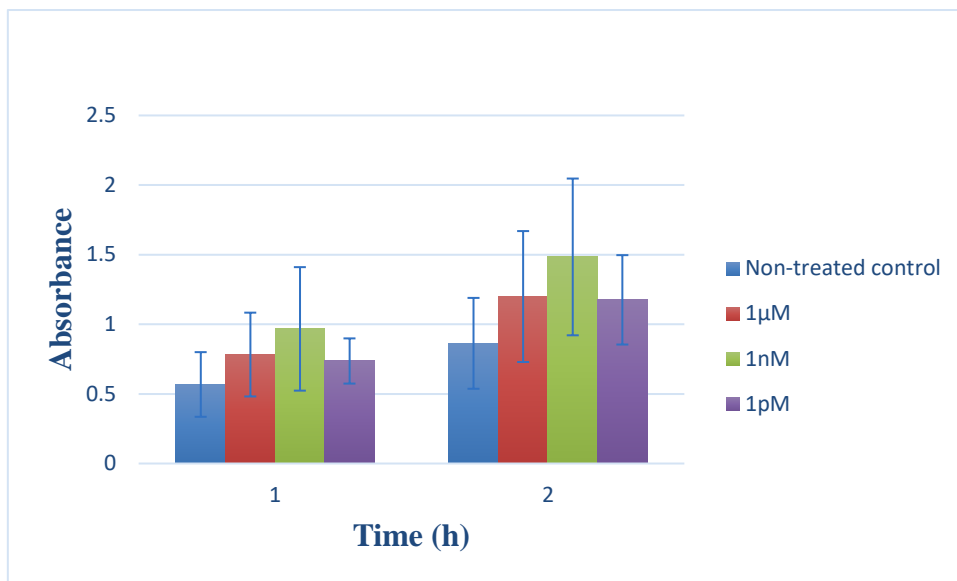
Figure 22. The proliferation efficiency of tert-butyl 6-methylpiperidine-3-carboxylate **3** on HEK-293 cells.
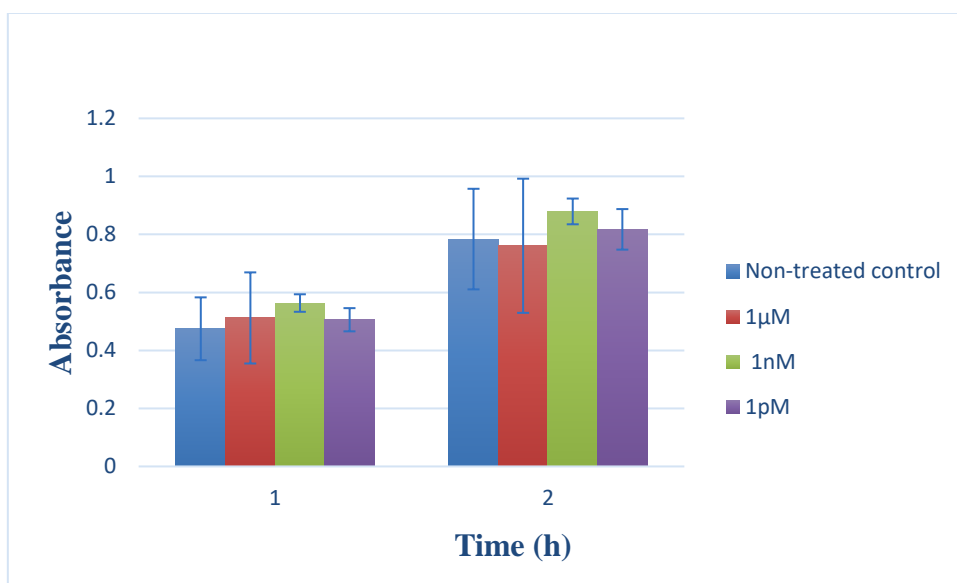
Figure 23. The proliferation efficiency of 3-Amino-4-chloro-N-(m-tolyl)benzenesulfonamide **4** on HEK-293 cells.
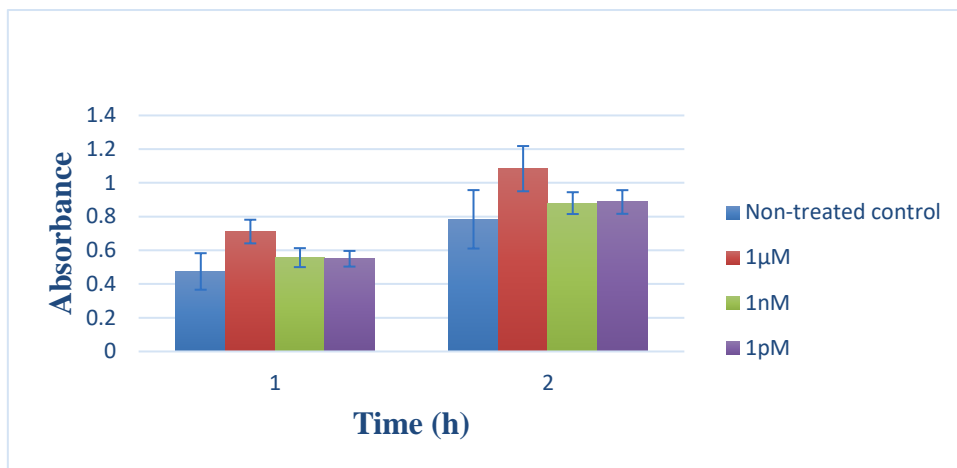
Figure 24. The proliferation efficiency of methyl piperazine-2-carboxylate **5** on HEK-293 cells.

The effect of studied compounds on the cell proliferation can be clearly seen also by examining the images of the cell cultures. As compared to the negative control, i.e. no inhibitor added (Figure 25, a) to the HEK-293 cell culture, the cell cultures with 1 μM (Figure 25,b) and 0.75 nM (Figure 25,c) added have after 24 hours substantial increase in the number of cells as well as higher regularity.
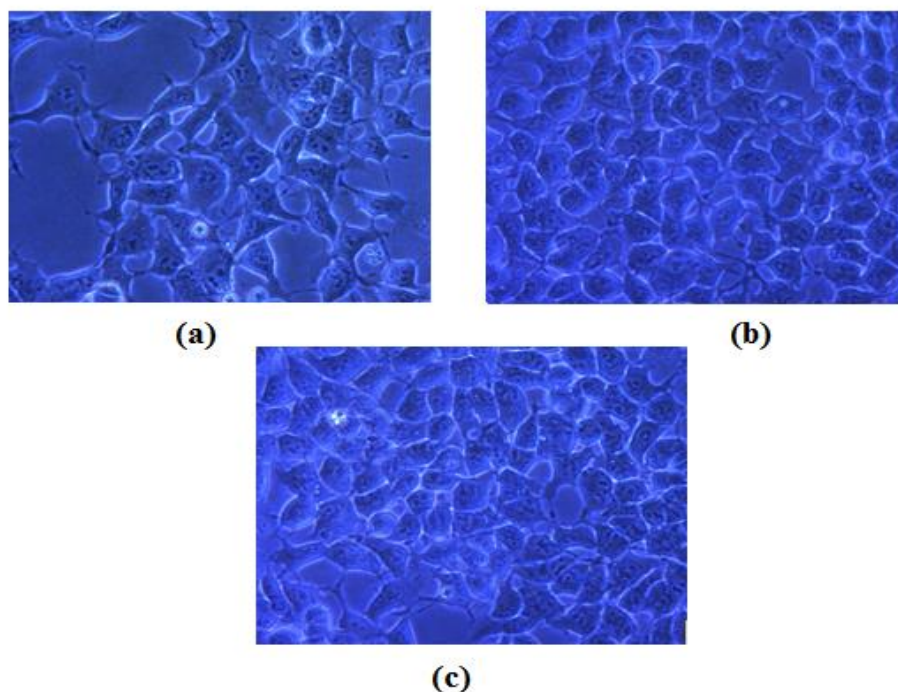
Figure 25. Images of the HEK-293 cell cultures at 24 hours after addition of compound **1** (in 0.05% DMSO): no compound added (a); 1 μM compound added (b); 0.75 nM compound added (c).

These results demonstrate that compounds **1** (methyl piperidine-3-carboxylate) and **2** (4-(9H-purin-6-yloxy)aniline) have significant effects on cell proliferation at low nanomolar concentrations and may thus have valuable medical applicability as supporters of the tissue regeneration.

5. Summary

Recent studies have shown that RNA m6A methylation plays significant and broad role in various physiological processes, such as development, fertility, carcinogenesis, stemness, early mortality, meiosis and circadian cycle, and links to obesity, cancer, and other human diseases [20]. The detailed investigation and development of chemical agents regulating this process is therefore of large biological and medical importance. Thus the goal of the present work was to design the inhibitory ligands for the enzymes regulating the RNA adenosine N6-methylation using methods of computational chemistry. The molecular docking using AutoDock 4.2 software was used to locate the potential ligand binding sites of the RNA methyltransferase complex Mettl3/Mettl14, RNA demethylase FTO and RNA "reader" protein YTHDF1. A virtual screening on the large compound databases ZINC, MolPort and DrugBank 4.0 was carried out to identify compounds structurally fitting to predicted binding sites. The best docking efficiencies were found for a series of piperidine and piperazine derivatives binding to the methyltransferase Mettl3/Mettl14 complex. Because of large potential biomedical importance of the RNA methylation inhibitors, the study was then concentrated on these compounds. The molecular dynamics simulations using the Desmond 4.2 code within Schrödinger LLC Maestro program was carried out for two compounds representing the most promising scaffolds. The results of the simulations enabled to detail the specific ligand-protein interactions for the further structure optimization of inhibitors. The biological activity for five predicted potential inhibitors of the Mettl3/Mettl14 complex was investigated using the tetrazolium salt WST-1 (Roche®) assay that quantifies the cell proliferation. Two compounds (methyl piperidine-3-carboxylate and 4-(9H-purin-6-yloxy)aniline) exhibited high proliferation enhancing activity on human embryonic kidney HEK-293 cells, even at nanomolar concentration level. The promotion of the cell proliferation by Mettl3/Mettl14 complex inhibitors was also demonstrated using morphological cell culture imaging studies. The best designed compounds are the first known RNA methyltransferase complex Mettl3/Mettl14 inhibitors and therefore of high interest for further biochemical, cellular and *in vivo* studies.

6. Kokkuvõte

Hiljutised uuringud on näidanud, et RNA adenosiini N6-metüleerimine omab olulist rolli paljudes füsioloogilistes protsessides nagu näiteks meioos, organismi areng, viljakus, kartsinogenees, tüvirakulisus, varajane suremus ja ööpäevased tsüklid ning seondub ülekaalulisuse, vähihaiguse ja paljude teiste haigustega. Seetõttu on RNA metüleerimist reguleerivate keemiliste agentide uurimine ja väljaarendamine suure bioloogilise ja praktilise meditsiinilise tähtsusega. Käesoleva töö eesmärgiks oli seega leida inhibeerivaid madalmolekulaarseid ligande RNA adenosiini N6-metüleerimist reguleerivatele ensüümidele, kasutades selleks arvutikeemia meetodeid. Molekulaarsildamise programmi AutoDock 4.2 abil selgitati välja potentsiaalsed ligandide sidumiskohad RNA metüültransferaasi kompleksile Mettl3/Mettl14, RNA demetülaasile FTO ning RNA „lugeja" valgule YTHDF1. Keemiliste ühendite suurtest andmebaasidest ZINC, MolPort ja DrugBank 4.0 leiti virtuaalse sõelumise teel ennustatud sidumiskohtadesse sobivad ligandid. Parimate sildamise efektiivsustega (DE > 0.5) olid mitmed Mettl3/Mettl14 kompleksile seonduvad piperidiini ja piperasiini derivaadid. RNA metüleerimise inhibiitorite suure biomeditsiinilise tähtsuse tõttu keskenduti edasises uuringus just neile ühenditele. Kahe parema erineva põhistruktuuriga ühendi jaoks viidi läbi molekulaardünaamika arvutused Schrödinger LLC Maestro tarkvaras oleva Desmondi programmi abil. Arvutuste tulemused võimaldasid täpsustada spetsiifilisi ligand-valk interaktsioone, mis on olulised inhibeerivate ligandide struktuuri edasiseks optimeerimiseks. Viie ennustatud Mettl3/Mettl14 kompleksi inhibiitori bioloogilist aktiivsust uuriti, kasutades tetrasooliumi soola WST-1 (Roche®) testi, mis võimaldab määrata rakkude proliferatsiooni taset. Kaks ühendit (metüül-piperidiin-3-karboksülaat ja 4-(9H-puriin-6-üüloksü)aniliin) omasid suurt proliferatsiooni taset tõstvat võimet inimese embrüonaalsetele neerurakkudele HEK-293, isegi nanomolaarse kontsentratsiooni tasemel. Rakkude proliferatsiooni kasv Mettl3/Mettl14 kompleksi inhibiitorite toimel oli selgelt detekteeritav ka rakukultuuride morfoloogilistes uuringutes. Parimad leitud ühendid on esimesed teadaolevad RNA metüültransferaasi Mettl3/Mettl14 kompleksi inhibiitorid ning pakuvad seega suurt huvi edasisteks biokeemilisteks, raku- ja loomkatseteks.

References

1. Felsenfeld, G. A Brief History of Epigenetics. *CSH Perspect Biol.* **2014***, 6,* 1-10.

2. Maity, A.; Das, B. N6-methyladenosine Modification in mRNA: Machinery, Function and Implications for Health and Diseases. *FEBS Journal*. **2016**, *283*, 1607-1630.

3. Chi, K.R.The RNA Code Comes into Focus. *Nature*. **2017**, *542*, 503-506.

4. Song, J.; Yi, C. Chemical Modifications to RNA: A New Layer of Gene Expression Regulation. *ACS Chem. Biol.* **2017**, *12*, 316-325.

5. Mason, C.; Dunnill, P. A Brief Definition of Regenerative Medicine. *Regen. Med*. **2008**, *1*, 1-5.

6. Munsie, M.; Hyun, I.; Sugarman, J. Ethical Issues in Human Organoid and Gastruloid Research. *Dev.* **2017**, *144,* 942-945.

7. Reya, T.; Morrison, S.J.; Clarke, M.F.; Weissman, I.L. Stem Cells, Cancer, and Cancer Stem Cells. *Nature*. **2001**, *414*, 105-111.

8. Dawson, M.A. The Cancer Epigenome: Concepts, Challenges, and Therapeutic Opportunities. *Science*. **2017**, *335*, 1147-1152.

9. Takahashi, K.; Yamanaka, S. Induction of Pluripotent Stem Cells from Mouse Embryonic and Adult Fibroblast Cultures by Defined Factors. *Cell*. **2006**, *126*, 663-676.

10. Yu, J.; Vodyanik, M.A.; Smuga – Otto , K.; Antosiewicz – Bourget, J.; Frane, J.L.; Tian, S.; Nie, J.; Jonsdottir, G.A.; Ruotti, V.; Stewart, R.; Slukvin, I.I.; Thomson, J.A. Induced Pluripotent Stem Cell Lines Derived from Human Somatic Cells. *Science*. **2007**, *318*, 1917-1920.

11. Yamanaka, S.; Blau, H.M. Nuclear Reprogramming to a Pluripotent State by Three Approaches. *Nature*. **2010**, *465*, 704-712.

12. Mahla, R.S. Stem Cells Applications in Regenerative Medicine and Disease Therapeutics. *Int. J. Cell Biol*. **2016**, *2016*, 1-24.

13. Srivastava, D.; DeWitt, N. In Vivo Cellular Reprogramming: The Next Generation. *Cell.* **2016***, 166,* 1386-1396.

14. Villota – Salazar, N.A.; Mendoza – Mendoza, A.; González – Prieto, J.M. Epigenetics: from the Past to the Present. *Front. Life Sci*. **2016**, *9*, 347-370.

15. Cedar, H.; Bergman, Y. Linking DNA methylation and histone modification: patterns and paradigms. *Nat. Rev. Genet*. **2009**, *10*, 295-304.

16. Rose, N.R.; Klose, R.J. Understanding the relationship between DNA methylation and histone lysine methylation. *BBA*. **2014**, *1839*, 1362-1372.

17. Mauer, J.; Luo, X.; Blanjoie, A.; Jiao, X.; Grozhik, A.V.; Patil, D.P.; Linder, B.; Pickering, B.F.; Vasseur, J.-J.; Chen, Q.; Gross, S.S.; Elemento, O.; Debart, F.; Kiledjian, M.; Jaffrey, S.R.Reversible Methylation of m^{6}A$_m$ in the 5′ Cap Controls mRNA Stability. *Nature*. **2017,** *541*, 371-375.

18. Xiang, Y.*;* Laurent, B.; Hsu, C.-H.; Nachtergaele, S.; Lu, Z.; Sheng, W.; Xu, C.; Chen, H.; Ouyang, J.; Wang, S.; Ling, D.; Hsu, P.-H.; Zou, L.; Jambhekar, A.; He, C.; Shi, Y. RNA m^{6}A methylation regulates the ultraviolet-induced DNA damage response. *Nature*. **2017**, *543*, 573-576.

19. Li, X.; Xiong, X.; Yi, C. Epitranscriptome Sequencing Technologies: Decoding RNA Modifications. *Nature Methods*. **2017**, *14*, 23-31.

20. Wu, R.; Jiang, D.; Wang, Y.; Wang, X. N6-Methyladenosine (m6A) Methylation in mRNA with A Dynamic and Reversible Epigenetic Modification. *Mol. Biotechnol*. **2016**, *58*, 450-459.

21. Chen, K.; Zhao, B.S.; He, C. Nucleic Acid Modifications in Regulation of Gene Expression. *Cell Chem. Bio*. **2016**, *23*, 75-85.

22. Bakin, A.; Lane, B.G.; Ofengand, J. Clustering of pseudouridine residues around the peptidyltransferase center of yeast cytoplasmic and mitochondrial ribosomes. *Biochemistry*. **1994**, *33*, 13475-13483.

23. Machnicka, M.A.; Milanowska, K.; Osman Oglou, O.; Purta, E.; Kurkowska, M.; Olchowik, A.; Januszewski, W.; Kalinowski, S.; Dunin-Horkawicz, S.; Rother, K.M.; Helm, M.; Bujnicki, J.M.; Grosjean, H. MODOMICS: a Database of RNA Modification Pathways– 2013 update. *Nucleic Acids Res*. **2014**, *41*, D262-D267.

24. Park, C.H.; Hong, K. Epitranscriptome: m6A and its function in stem cell biology. *Genes Genom*. **2017**, *39*, 371-378.

25. Niu, Y.; Zhao X.; Wu, Y.-Z.; Li, M-M.; Wang, X.-J.; Yang, Y.-G. N6-Methyl-Adenosine (m6A) in RNA: An Old Modification with A Novel Epigenetic Function. *GPB*. **2013**, **11**, 8-17.

26. Sledz, P.; Jinek, M. Structural Insights into the Molecular mechanism of the m(6)A Writer Complex. *Elife*. **2016**, 5, e18434.

27. Scavetta , R.D.; Thomas, C.B.; Walsh, M.A.; Szegedi, S.; Joachimiak, A.; Gumport R.I.; Churchill, M.E.A. Structure of *Rsr*I Methyltransferase, a Member of the *N*6-adenine β Class of DNA Methyltransferases. *Nucleic Acids Res*. **2000**, *28*, 3950-3961.

28. Xiang, W.; Jing, F.; Yuan, X.; Guan, Z.; Zhang, D.; Liu, Z.; Gong, Z.; Wang, Q.; Huang, J.; Tang, C.; Zou, T.; Yin, P. Structural basis of N-6-Adenosine Methylation by the METTL3-METTL14 Complex. *Nature*. **2016**, 534, 575-578.

29. Wang, P.; Doxtader, K.A.; Nam, Y. Structural Basis for Cooperative Function of Mettl3 and Mettl14 Methyltransferases. *Mol. Cell*. **2016**, *63*, 306-317.

30. Zhou, K.I.; Pan, T. Structures of the m(6)A Methyltransferase Complex: Two Subunits with Distinct but Coordinated Roles. *Mol. Cell*. 2016, 63, 183-185.

31. Jia, G.; Fu, Y.; Zhao X.; Dai, Q.; Zheng, G.; Yang, Y.; Yi, C.; Lindahl, T.; Pan, T.; Yang Y.G.; He, C. N6-Methyladenosine in Nuclear RNA is a Major Substrate of the Obesity-Associated FTO. *Nat. Chem. Biol*. **2017**, *12*, 885–887.

32. Han, Z.; Niu, T.; Chang, J., Lei, X., Zhao, M., Wang, Q.; Cheng, W.; Wang, J., Feng, Y., Chai, J. Crystal Structure of the FTO Protein Reveals Basis for Its Substrate Specificity. *Nature*. **2010**, *464*, 1205-1209.

33. Aik, W.S.; Demetriades, M.; Hamdan, M.K.K.; Bagg, E.A.L.; Yeoh, K.K.; Lejeune, C.; Zhang, Z.; McDonough, M.A.; Schofield, C.J.; Structural basis for Inhibition of the Fat Mass and Obesity Associated Protein (FTO. *J.Med.Chem*. **2013**, *56*, 3680-3688.

34. Toh, J.D.W.; Sun, L.; Lau, L.Z.M.; Tan, J.; Low, J.J.A.; Qiang, W.; Tang, C.W.Q.; Yi, J.; Cheong, E.J.Y.; Tan, M.J.H.; Chen, Y.; Hong, W.; Gao, Y.G.; Woon, E.C.Y. A Strategy Based on Nucleotide Specificity Leads to Subfamily-Selective and Cell-Active Inhibitor of N6-Methyladenosine Demethylase FTO. *Chem. Sci*. **2015**, *6*, 112-122.

35. Huang, Y.; Yan, J.; Li, Q.; Li, J.; Gong, S.; Zhou, H.; Gan, J.; Jiang, H.; Jia, G.F.; Luo, C.; Yang, C.-G. Meclofenamic Acid Selectively inhibits FTO Demethylation of m6A over ALKBH5. *Nucleic Acids Res*. **2014**, *43*, 373-384.

36. Feng, C., Liu, Y., Wang, G., Deng, Z., Zhang, Q., Wu, W., Tong, Y., Cheng, C., Chen, Z., Crystal Structures of the Human RNA Demethylase Alkbh5 Reveal Basis for Substrate Recognition. *J. Biol. Chem*. **2014**, *289*, 11571-11583.

37. Xu, C.; Liu, K.; Ahmed, H.; Loppnau, P.; Schapira, M.; Min, J. Structural Basis for the Discriminative Recognition of N6-Methyladenosine RNA by the Human YT521-B Homology Domain Family of Proteins. *J. Biol. Chem*. **2015**, *290*, 24902-24913.

38. Li, F.; Zhao, D.; Wu, J.; Shi, Y. Structure of the YTH Domain of Human YTHDF2 in Complex with an m(6)A Mononucleotide Reveals an Aromatic Cage for m(6)A Recognition. *Cell Res*. **2014**, *24*, 1490-1492.

39. Xu, C. Wang, X.; Liu, K.; Roundtree, I.A.; Tempel, W.; Li, Y.; Lu, Z.; He, C.; Min, J. Structural Basis for Selective Binding of m(6)A RNA by the YTHDC1 YTH Domain. *Nat.Chem.Biol*. **2014**, *10*, 927-929.

40. Chen, B.; Ye, F.; Yu, L.; Jia, G.; Huang, X.; Zhang, X.; Peng, S.; Chen, K.; Wang,

M.; Gong, S.; Zhang, R.; Yin, J.; Li, H.; Yang, Y.; Liu, H.; Zhang, J.; Zhang, H.; Zhang, A.; Jiang, H.; Luo, C.; Yang, C.-G. Development of Cell-Active N6-methyladenosine RNA Demethylase FTO Inhibitor. *J. Am. Chem. Soc*. **2012**, *134*, 17963-17971.

41. Huang, Y.; Yan, J.; Li, Q.; Li, J.; Gong, S.; Zhou, H.; Gan, J.; Jiang, H.; Jia, G.-F.; Luo, C.; Yang, C.-G. Meclofenamic Acid Selectively Inhibits FTO Demethylation of m(6)A over ALKBH5. *Nucleic Acids Res*. **2015**, *43*, 373-384.

42. Xu, C.; Liu, K.; Tempel, W.; Demetriades, M.; Aik, W.; Schofield, C.J.; Min, J. Structures of Human ALKBH5 Demethylase Reveal a Unique Binding mode for Specific Single-Stranded N6-methyladenosine RNA Demethylation. *J. Biol. Chem*. **2014**, *289*, 17299-17311.

43. Aik, W.; Scotti, J.S.; Choi, H.; Gong, L.; Demetriades, M.; Schofield, C.J.; Mcdonough, M.A. Structure of Human RNA N-Methyladenine Demethylase ALKBH5 Provides Insights into its Mechanisms of Nucleic Acid Recognition and Demethylation. *Nucleic Acids Res*. **2014**, *42*, 4741-4754.

44. He, W.; Zhou, B.; Liu, W.; Zhang, M.; Shen, Z.; Han, Z.; Jiang, Q.; Yang, Q.; Song, C.; Wang, R.; Niu, T.; Han, S.; Zhang, L.; Wu, J.; Guo, F.; Zhao, R.; Yu, W.; Chai, J.; Chang, J. Identification of a Novel Small-Molecule Binding Site of the Fat Mass and Obesity Associated Protein (FTO). *J. Med. Chem*. **2015**, *58*, 7341-7348.

45. Sastry, G. M.; Adzhigirey, M.; Day, T.; Annabhimoju, R.; Sherman, W. Protein and Ligand Preparation: Parameters, Protocols, and influence on Virtual Screening Enrichments. *J. Comput. Aid. Mol. Des.* **2013***, 27,* 221-234.

46. 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*, 455-461.

47. Morris, G. M.; Huey, R.; Lindstrom, W., Sanner, M. F.; Belew, R. K., Goodsell, D. S.; Olson, A. J. Autodock4 and AutoDockTools4: Automated Docking with Selective Receptor Flexiblity. *J. Comput. Chem.* **2009**, *16*, 2785-91.

48. Stephens, P.J.; Devlin, F.J.; Chabalowski, C.F.; Frisch, M.J. Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields. *J. Phys. Chem*. **1994**, *98*, 11623-11627.

49. Bowers, K.J.; Chow, D.E.; Xu, H.; Dror, R.O.; Eastwood, M.P.; Gregersen, B.A.; Klepeis, J.L.; Kolossvary, I.; Moraes, M.A.; Sacerdoti, F.D.; Salmon, J.K.; Shan, Y.; Shaw, D.E. Scalable Algorithms for Molecular Dynamics Simulations on Commodity Clusters. In: Proceedings of the ACM/IEEE SC **2006** conference, 43-43.

50. Banks, J.L.; Beard, H.S.; Cao, Y.; Cho, A.E.; Damm, W.; Farid, R.; Felts, A.K.; Halgren, T.A.; Mainz, D.T.; Maple, J.R.; Murphy, R.; Philipp, D.M.; Repasky, M.P.; Zhang, L.Y.; Berne, B.J.; Friesner, R.A.; Gallicchio, E.; Levy. R.M. Integrated Modeling Program, Applied Chemical Theory (IMPACT). *J. Comput. Chem*. **2005**, *26*, 1752-1780.

51. Toukmaji, A.Y.; Board Jr., J.A. Ewald Summation Techniques in Perspective: a Survey. *Comput. Phys. Commun*. **1996**, *95*, 73-92.

52. Zielkiewicz, J. Structural Properties of Water: Comparison of the SPC, SPCE, TIP4P, and TIP5P Models of Water. *J. Chem. Phys*. **2006**, *124*, 109901.

53. Martyna, G.J.; Klein, M.L. Nosé–Hoover Chains: The Canonical Ensemble Via Continuous Dynamics. *J. Chem. Phys*. **1992**, *97*, 2635-2643.

54. Li, F.; Kennedy, S.; Hajian, T.; Gibson, E.; Seitova, A.; Xu, C.; Arrowsmith, C.A.; Vedadi, M. A Radioactivity Based Assay for Screening Human m6A-RNA Methyltransferase, METTL3-METTL14 Complex, and Demethylase ALKBH5. *J. Biomol. Screen*. **2016**, *21*, 290-297.

55. Irwin, J.J.; Shoichet, B.K. ZINC – A Free Database of Commercially Available Compounds for Virtual Screening. J. Chem. Inf. Model. 2005, 45, 177-182.

56. Law, V.; Knox, C.; Djoumbou, Y.; Jewison, T.; Guo, A.C.; Liu, Y.; Maciejewski, A.; Arndt, D.; Wilson, M.; Neveu, V.; Tang, A.; Gabriel, G.; Ly, C.; Adamjee, S.; Dame, Z.T.; Han, B.; Zhou, Y.; Wishart, D.S. DrugBank 4.0: Shedding New Light on Drug Metabolism. *Nucleic Acids Res*. **2014**, *42*, D1091-D1097.

57. Bento, A.P.; Gaulton, A.; Hersey, A.; Bellis, L.J.; Chambers, J.; Davies, M.; Krüger, F.A.; Light, Y.; Mak, L.; McGlinchey, S.; Nowotka, M.; Papadatos, G.; Santos, R.; Overington, J.P. The ChEMBL Bioactivity Database: an Update. *Nucleic Acids Res*. **2014**, *42*, 1083-1090.

58. Lichinchi, G.; Zhao, B.S.; Wu, Y.; Lu, Z.; Qin, Y.; He, C.; Rana, T.M. Dynamics of Human and Viral RNA Methylation During Zika Virus Infection. *Cell Host Microbe*. **2016**, *20,* 666-673.

59. Dearden, C.E.; Else, M.; Catovsky, D. Long-Term Results for Pentostatin and Cladribine Treatment of Hairy Cell Leukemia. *Leuk. Lymphoma*. **2011**, *52*, 21-24.

60. Genheden, S.; Ryde, U. The MM/PBSA and MM/GBSA. *Expert Opin. Drug Discov.* **2015**, *10*, 449-461.

61. Batista, P.J.; Molinie, B.; Wang, J.K.; Qu, K.; Zhang, J.J.; Li, L.J.; Bouley, D.M.; Lujan, E.; Haddad, B.; Daneshvar, K.; Carter, A.C.; Flynn, R.A.; Zhou, C.; Lim, K.S.; Dedon, P.; Wernig, M. ; Mullen, A.C.; Xing, Y.; Giallourakis, C.C.; Chang, H.Y. M6A RNA

Modification Controls Cell Fate Transition in Mammalian Embryonic Stem Cells. *Cell Stem Cell*. **2014**, *15*, 707-719.

Tartu Ülikooli loodus-ja täppisteaduste valdkonna õppeprodekaanile
Taotlus lõputöö avaldamisele piirangute kehtestamiseks ja lõputöö kaitsmise kinniseks kuulutamiseks
Nimi   Simona Selberg
Sünniaeg   29.07.1993
Õppekava   Keemia
Juhendaja   Prof. Mati Karelson
Lõputöö pealkiri Molecular design of inhibitors for RNA methylation regulating enzymes

[x] Palun **mitte avaldada** minu lõputööd kuni **31.05.2020.**

    [x] teistele isikutele kuuluvate varaliste autoriõiguste tõttu.

    [ ] kuna töö sisaldab isikuandmeid, mille avaldamiseks pole andmesubjekti nõusolekut.

    [ ] riigisaladuse tõttu.

    [ ] ärisaladuse tõttu.

    [ ] lõputöö avaldatakse tulevikus teadusartiklina.

    [ ] muul põhjusel.

    Selgitus (*põhjendus, miks taotletakse piiranguid ja miks just selliseks perioodiks*):

    Töö on teostatud lepingu raames Chemestmed OÜ-ga, kellele kuuluvad lepingujärgselt intellektuaalomandi õigused.

[ ] Palun kuulutada minu **lõputöö kaitsmine kinniseks.**
    Selgitus (*põhjendus, miks taotletakse lõputöö kaitsmise kinniseks kuulutamist*):

Kuupäev ja üliõpilase
allkiri: _____

Kuupäev ja juhendaja
allkiri: _____

Taotlusele on lisatud järgmised lisad (*täita juhul, kui taotlusel on lisad, nt äriühingu poolne kinnitus, et töö sisaldab ärisaladust*):
1.