Analysis of DNA methyltransferase DNMT1 structure and function and development of noncovalent inhibitors using computational methods

Morović, Viktor

Master's thesis / Diplomski rad

2023

Degree Grantor / Ustanova koja je dodijelila akademski / stručni stupanj: **University of Rijeka / Sveučilište u Rijeci**

Permanent link / Trajna poveznica: https://urn.nsk.hr/urn:nbn:hr:193:785350

UNIVERSITY OF RIJEKA
DEPARTMENT OF BIOTECHNOLOGY
Graduate programme
Medicinal chemistry

Viktor Morović

Analysis of DNA methyltransferase DNMT1 structure and function and development of noncovalent inhibitors using computational methods

Master's Thesis

Rijeka, 2023.

UNIVERSITY OF RIJEKA
DEPARTMENT OF BIOTECHNOLOGY
Graduate programme
Medicinal chemistry

Viktor Morović

Analysis of DNA methyltransferase DNMT1 structure and function and development of noncovalent inhibitors using computational methods

Master's Thesis

Rijeka, 2023.

Mentor: Doc. dr. sc. Željko M. Svedružić
Co-mentor: Doc. dr. sc. Jurica Novak

SVEUČILIŠTE U RIJECI
ODJEL ZA BIOTEHNOLOGIJU
Diplomski sveučilišni studij
Medicinska kemija

Viktor Morović

Analiza strukture i funkcije DNA metiltransferaze DNMT1 i razvoj nekovalentnih inhibitora koristeći računalne metode

Diplomski rad

Rijeka, 2023.

Mentor: Doc. dr. sc. Željko M. Svedružić
Komentor: Doc. dr. sc. Jurica Novak

The thesis was defended on 14.7.2023. in front of committee members:

1. izv. prof. dr. sc. Rozi Andretić Waldowski
2. izv. prof. dr. sc. Nela Malatesti
3. Doc. dr. sc. Jurica Novak

The thesis has 36 pages, 19 images, 3 tables and 27 references.

Sažetak

DNA metiltransferaza 1 (DNMT1) je enzim koji ima ključnu ulogu u epigenetskoj regulaciji genske ekspresije kod sisavaca. Disregulacija procesa metilacije može dovesti do raznih patoloških stanja, što je posljednjih godina dovelo do opsežnog proučavanja strukture I funkcije DNMT1. DNMT1 postoji u više oblika ovisno o stupnju metilacije DNA, što ograničava mogućnost u otkrivanju i modifikaciji novih lijekova. Korištenjem *in silico* metoda nastojali smo identificirati spojeve slične lijekovima koji se mogu vezati I inhibirati funkciju DNMT1.

Analiza simulacija molekularne dinamike (MD pokazale su jače vezanje SAH na sva tri proteina DNMT1 (4WXX, 3PTA i 4DA4), u usporedbi s testiranim spojevima. Dva spoja pomaknula su se s veznog mjesta SAH tijekom početnih simulacija, zbog čega su isključena iz daljnjeg istraživanja.

Ispitani spojevi pridržavaju se Lipinskog pravila, ali stvaraju manje vodikovih veza s DNMT1, u usporedbi sa SAH. Buduća bi istraživanja trebala iskoristiti potencijal π-π interakcija s aminokiselinskim ostacima unutar SAH veznog džepa.

Ključne riječi:

DNA metiltransferaza DNMT1, inhibitori DNMT1, dizajn lijekova, molekularna dinamika

Abstract

DNA methyltransferase 1 (DNMT1) is an enzyme that plays an essential role in epigenetic regulation of mammalian genetic expression. Dysregulation of methylation processes can lead to various pathological conditions, which in recent years ,lead to extensive study of structure and function of DNMT1. DNMT1 has multiple forms dependant of the stage of DNA methylation, which posses limitations in both discovery and modification of new drugs. With use of *in silico* methods we sought to identify drug like compounds capable of binding to, and inhibiting the function of DNMT1.

Analysis of molecular dynamics (MD) simulations showed stronger binding of SAH to all three DNMT1 proteins (4WXX, 3PTA and 4DA4), compared to tested compounds. Two compounds moved from the SAH binding site during initial simulations, for which they were excluded from further research.

Tested compounds adhere to the Lipinski rule but lack the formation of hydrogen bonds with DNMT1, compared to SAH. Future research should use the potential of π- π stacking interaction with residues inside SAH binding pocket.

Keywords:

DNA methylation, DNA methyltranferase DNMT1, DNMT1 inhibitors, drug design, molecular dynamics

Table of contents

1. Introduction

1.1 Epigenetics

Epigenetics is the study of heritable, reversible forms of gene regulation that do not involve DNA sequence alteration. Epigenetic modification include methylation, acetylation, ubiquitination as well as phosphorylation. These processes can alter gene expression by upregulating, downregulating or silencing genes completely. Dysregulation of epigenetic processes can lead to cancer development, cardiovascular disease as well as neurological and metabolic disorders (1).

Identification of drugs that can influence epigenetic modifications are of great clinical interest for treatment of these disorders. Major classes of epigenetic drugs currently on the market are inhibitors of DNA methylation, bromodomain inhibitors, histone deacetylase inhibitors, protein methyltransferase inhibitors, and histone methylation inhibitors (1).

1.2 DNA methylation

DNA methylation in mammalian cells is one of major epigenetic modification that influence gene expression as it is responsible for transcriptional silencing of retrotransposones, genomic imprinting as well as X chromosome inactivation in females (2). DNA methyltransfetases (DNMTs) are a family of enzymes that catalyze DNA methylation reactions by transfering methyl group on C5 position of cytosine (C) to form 5-methylcytosine (5mC) (Figure 1). S-adenosylmethionine (SAM) is a methyl donor in this reaction, and the transfer of methyl group results in conversion of SAM to S-adenosyl homocysteine (SAH). Demethylation of SAM is an irreversible reaction, and the cofactor (SAH) is a potent inhibitor of all methyltransferases. 5mC can spontaneously transiotion to thymine (T) due to deamination, making it inherently mutagenic.

Mammalian DNA methylation occurs within the context of cytosine phosphate guanine (CpG), where it accounts for 70-80% of CpG sites (2). After

fertilization, as well as germline cell specification, mammalian DNA is experiencing reprogramming of CpG methylation patterns (3).
Organism that have the ability for CpG methylation have smaller CpG content, for example, mammals have ~80% fewer CpG dinucleotides than is expected looking from its DNA sequence (3).

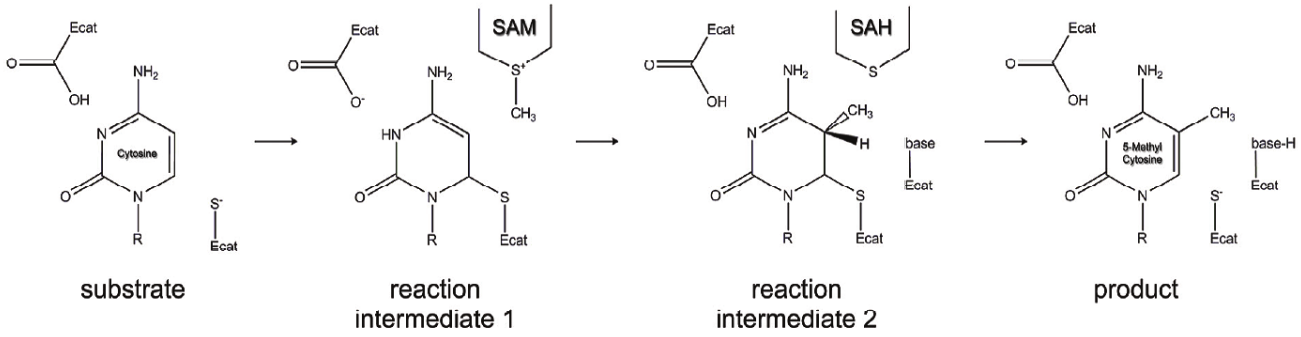


Figure 1. Reaction mechanism of DNMTs. SAM is the methyl donor and SAH is the cofactor product. Ecat are catalytic residues and base-H represent basic residues. (Source: Dhe-Paganon S, Syeda F, Park LC. DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. Int J Biochem Mol Biol 2: 58-66. *Int J Biochem Mol Biol* 2011; **2**: 58–66.)

When CpG sites are methylated, gene silencing is reinforced by a variety of mechanisms, such as binding to DNA promoter regions, inhibition of RNA polymerase and recruitment of transcriptional repressor complexes (5).
DNA methylation has three phases; establishment (*de novo* DNA methylation), maintenance and demethylation. For *de novo* DNA methylation are responsible DNA methyltransferase 3A (DNMT3A) and DNA methyltransferase 3B (DNMT3B) enzymes, whereas DNA methyltransferase 1 (DNMT1) has a role of propagating DNA methylation during replication by adding methyl groups to the unmethylated daughter strand (Figure 2) (1-3).
Besides DNMT1s ability to directly methylate DNA, it can also directly binds to methyl-CpG binding proteins that can dock to constrictive histone enzymes, which play a role in gene silencing (5).
There is evidence that DNMT3A and DNMT3B also have an important role in the maintenance phase of DNA methylation, whereas DNMT1 has the ability to perform *de novo* methylation (2).

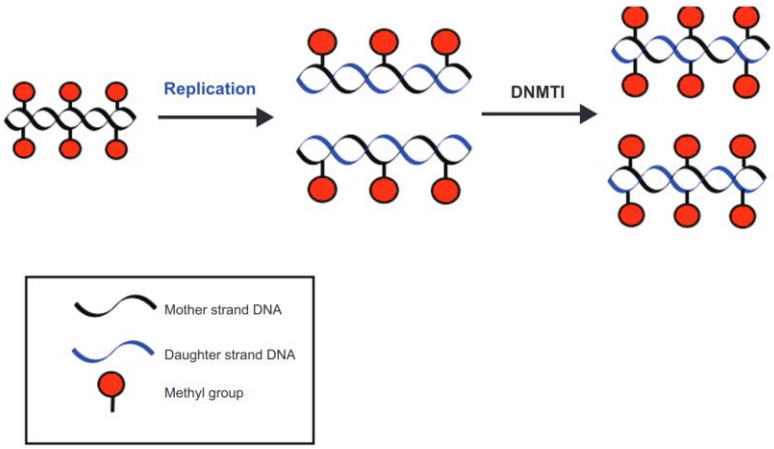


Figure 2. Maintenance of DNA methylation. After DNA replication, DNMT1 copies methylation pattern from mother strand to daughter strand on the specific CpG residues. (Source: Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of Epigenetic Drugs in Disease: An Overview. Genet Epigenetics 2014; 6: GEG.S12270)

Dysregulation of epigenetic processes may lead to emergence of many types of diseases, such as cardiovascular and neurological diseases, as well as metabolic disorders and cancer. Deletion of DNMT1 during embryonic stages of development will cause global demethylation leading to embryonic lethality. Upregulation of DNMT1 can cause methylation of CpG islands in promoter site of cell cycle inhibitors as well as pro-apoptotic genes, leading to silencing of tumor suppressor genes. When comparing cancer cells with normal tissue, almost all cancer cells have high expression level of DNMT1 (1).
Epigenetic changes during development of cancer are reversible, which makes them a good target for research and development of drug candidates (1).

1.3 Structure of DNMT1

DNMT1 contains 1616 amino acid residues and is the most abundant enzyme in the DNMT family in mammals. DNMT1 is made up of two subunits, the *N*-terminal regulatory domain and the *C*-terminal catalytic methyltransferase domain. Regulatory subunit includes residues 1-1139, while *C*-terminal methyltransferase (Mtase) domain corresponds to residues 1140-1616 (Figure 3). Mtase domain and regulatory region are linked by highly conserved thirteen glycine-lysine (GK) repeats (4).

N-terminal regulatory region starts with DNA methyltransferase associated protein 1 (DMAP1) binding site (residues 15-115), the Replication Foci Targeting Sequence (RFTS) domain (residues 351Lys - Arg600), followed by the CXXC zinc binding domain (residues Lys645 - Cys690) which is connected to Bromo-Adjacent Homology 1 (BAH1) domain (residues Thr755 - Phe875) and BAH2 domain (residues Glu965 - Pro1100) via an autoinhibitory linker (Figure 4). BAH 1 and BAH 2 are separated by one *α*-helix, as can be seen in Figure 4 colored in light sea green.

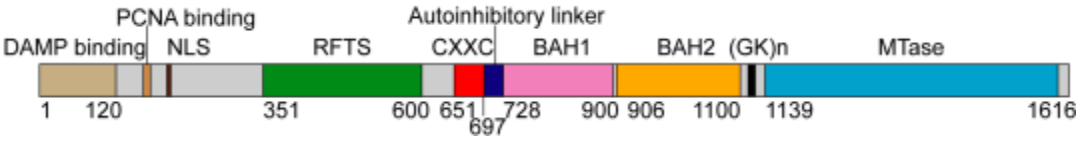


Figure 3. Domain architecture of human DNMT1 enzyme. Numbers represent residues that are domain borders. (Source: Ren W, Gao L, Song J. Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes* 2018; **9**: 620)

DNMT1 associated protein (DMAP) binding domain (residues 12-105) is able to partly suppress transcription by direct interaction with histone deacetylase 2 (HDAC2). In addition, the DMAP binding domain can form a complex with DMAP1 and HDAC2, and when this occurs, the complex functions as transcription co-repression (4).

Furthermore, the *N*-terminal region contains short motifs that aid DNMT1 in forming complex with proteins and DNA. One of these motifs (residue 162-171) binds proliferating cell nuclear antigen (PCNA) which enables DNMT1 to be recruited to replication site. DNMT1 also contains three nuclear localization signals (NLS) (residues 191–112, 259–378, and 630–757) whose function is still uncertain, but some evidence suggests that it may be important for nucleo-cytoplasmic shuttling. Interactions between *N*-terminal and *C*-terminal domains, such as CXXC region and catalytic domain, and RFTS region and catalytic domain, have been reported, which may be important for proper function of DNMT1(4).

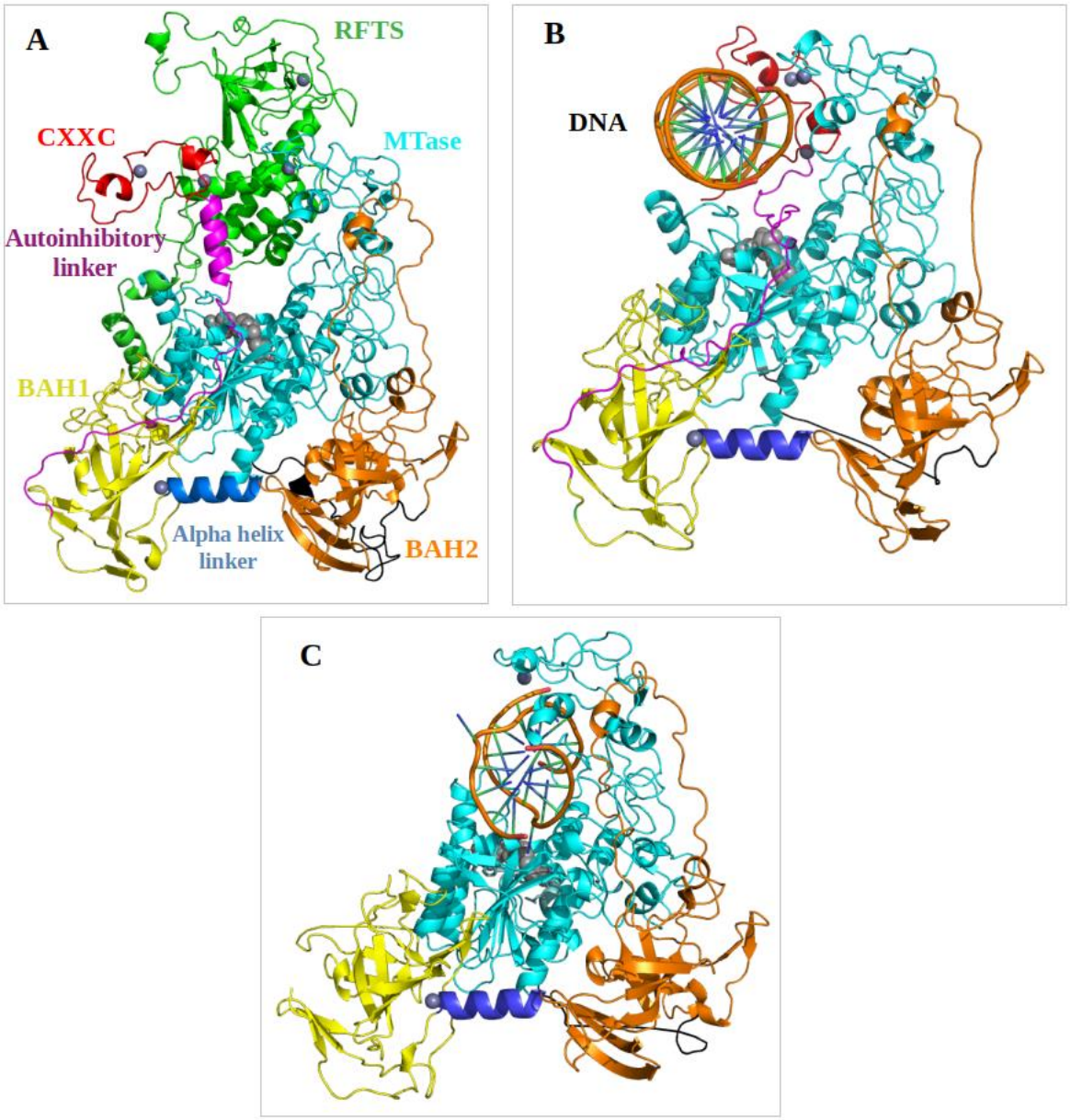


Figure 4. DNMT1 protein with its domains. A) 4WXX contains RFTS, Mtase, CXXC, Autoinhibitory linker, Alpha helix linker and 2 BAH domains, B) 3PTA has DNA molecule bound and is without RFTS domain, C) 4DA4 is without RFTS, CXXC and autoinhibitory linker. These figures represent three modes of conformation of DNMT1 protein; not bounded to DNA (A), bound to DNA but not active (B), bound to DNA and active (C).

When DNMT1 is not in complex with other proteins, it assumes an autoinhibitory configuration because its RFTS domain (residues 350-600) is partly situated in the catalytic MTase domain, preventing the process of methylation. Removal of RFTS domain from DNMT1 catalytic domain is dependent on E3 ubiquitin protein ligase UHRF1. When UHFR1 binds to DNA replication forks, it binds to RFTS domain of DNMT1 with its ubiquitin-like

(UBL) domain. This releases RFTS from DNMT1 catalytic domain, allowing the RFTS domain to bind to histone H3 tail and the MTase domain to methylate the daughter DNA strand (3). RFTS domain contains one zinc ion situated between 4 residues. On protein 4WXX these residues are Cys353, Cys356, Cys414 and His418. (Figure 5.A)

The CXXC zinc finger domain of mDNMT1 contains amino acid residues 621-698 and has a crescent-like fold. It got its name because it has eight Cys residue holding zinc cations. Cys residues are part of two short helical segments (Figure 5.B). CXXC domain is responsible for all contact between DNMT1 and DNA that is sequence specific, and it does so by targeting unmethylated CpG dinucleotides. CXXC domain targets major groove of DNA with a loop segment containing residues Ala684 - Glu687 (Figure 5.B) (6).

CXXC domain has two zinc ions, each connected to four Cys residues. On protein 4WXX, first zinc ion is connected to Cys653, Cys656, Cys659 and Cys691, and the second to Cys664, Cys667, Cys670 and Cys686. (Figure 5.B)

Linker connecting CXXC and BAH domain is called autoinhibitory linker because of its role in CXXC and RFTS domain mediated autoinhibition (7).

BAH1 and BAH2 domains are connected via an alpha helix linker and their function remains mostly unknown (Figure 4). Deletion of BAH1 and BAH2 domains results DNMT1 inactivity by preventing DNMT1 to form complex with DNA at replication foci, even in the presence of RFTS. This indicates that BAH domains have some crucial role in RFTS and UHRF1 dependent targeting (8).

BAH1 domain is connected to CXXC domain via BAH1-CXXC linker, which in vertebrates, is extremely conserved. The function of BAH1-CXXC linker is to prevent unmethylated CpG dinucleotides from entering in the active site of MTase (8). BAH1 domain contains one zinc ion situated between 4 residues. On protein 4WXX these residues are His793, Cys820, Cys893 and Cys896 (Figure 5.C). BAH1 and BAH2 are connected via alpha helix linker (Figure 4). BAH2 domain is connected to MTase domain via BAH2-TRD loop, which has a regulatory function by preventing TRD-DNA interaction (8).

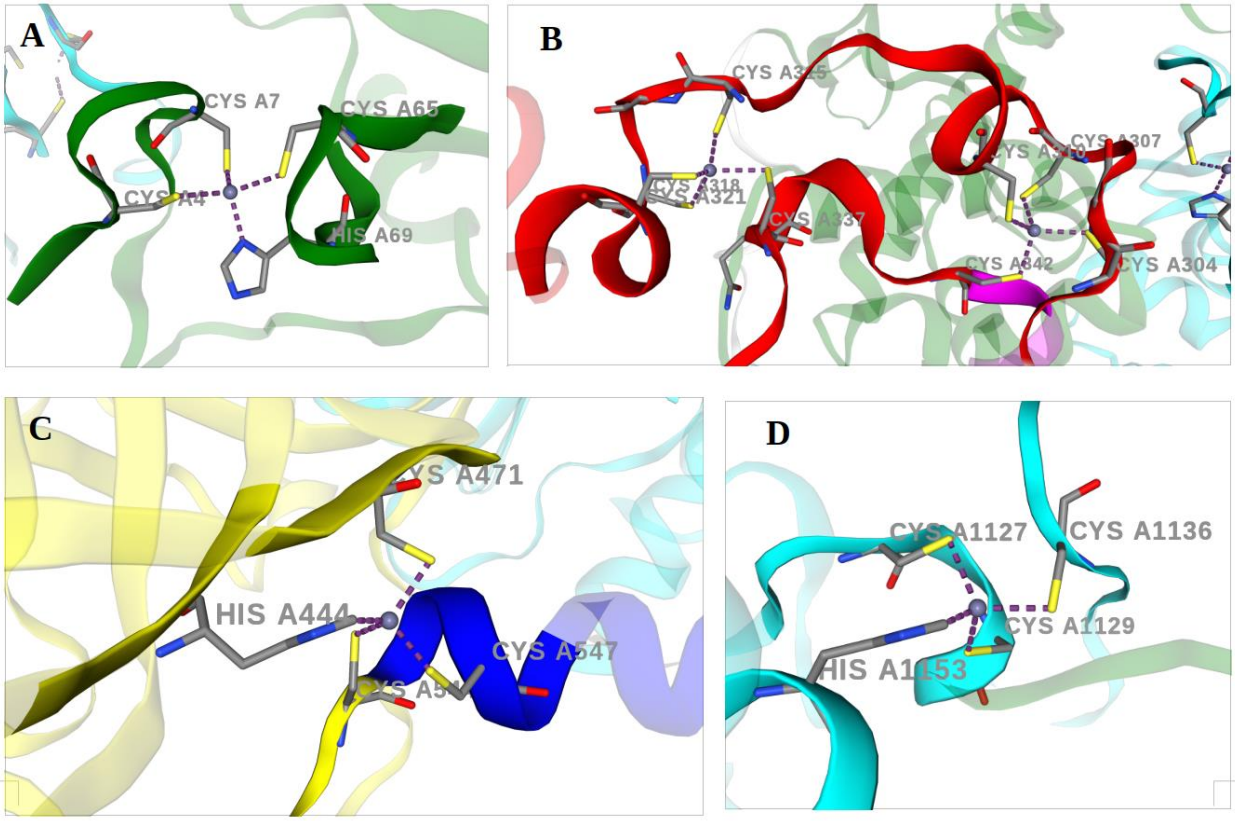


Figure 5. Zinc finger domains of DNMT1. A) RFTS zinc finger domain, B) CXXC zinc finger domains, C) BAH1 zinc finger domain and D) MTase zinc finger domain.

Mtase domain of DNMT1 contains residues Asn1140 - Trp1616 and includes 2 subdomains: catalytic core and target recognition domain (TRD). Architecture of DNMT1 catalytic core puts it in class I methyltransferase family, specifically, because of its secondary structure called Rossmann fold, which is responsible for binding *S*-adenosyl methionine (SAM). Catalytic core also includes motifs responsible for catalyzation of methylation reaction. The role of TRD subdomain is to recognize the binding site of DNA. TRD subdomain is placed inside the catalytic core, between the central beta sheet and the last alpha helix (2).

Catalytic domain contains the binding pocket where it connects *S*-adenosyl homocysteine (SAH or AdoHcy) via residues Gly1150, Ile1151, Glu1168, Met1169, Asp1190, Cys1191, Ser1146, Gly1149, Phe1145 and Val1580 (Figure 6). Unmethylated DNA binds to MTase domain in proximity to SAM.

Both BAH domains interact with MTase domain via hydrophobic interactions, and BAH1 domain is additionally connected to catalytic domain with antiparallel

beta sheets pairing (2). MTase domain contains one zinc ion situated between 4 residues. On protein 4WXX these residues are Cys1476, Cys1478, Cys1485 and His1502 (Figure 5.D).

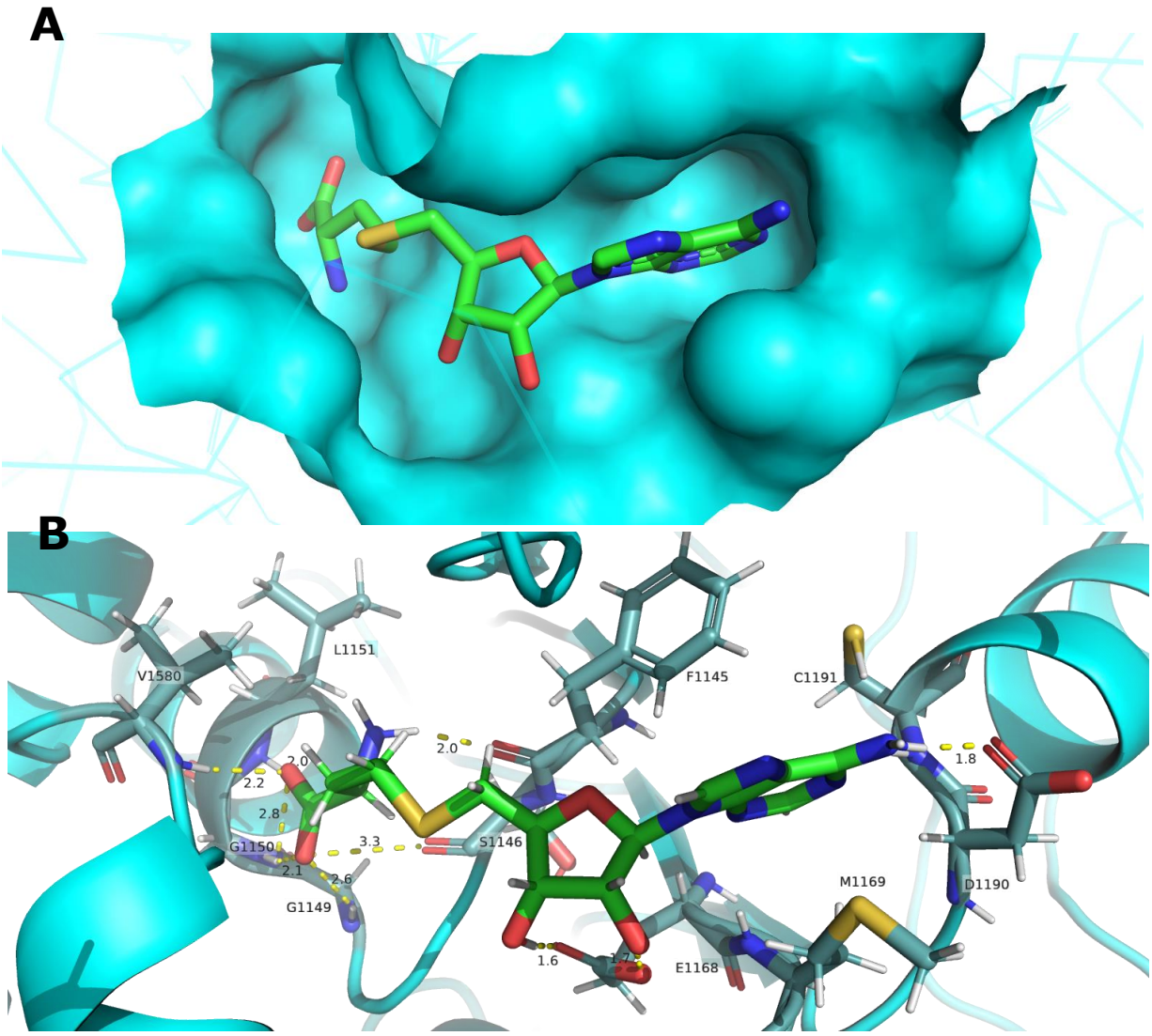


Figure 6. SAH in DNMT1 binding pocket. A) Protein 4WXX is shown with representation of solid surface. B) Protein is depicted with representation cartoon and licorice. SAH is depicted with representation licorice, where atoms are colored green (C), blue (N), red (O), yellow (S) and white (H). Images were created using UCSF Chimera.

1.4 Structure based drug design (SBDD)

Structure based drug design SBDD is a strategy in drug development that uses information of three-dimensional (3D) protein structure to develop new lead

compounds. Its use demands knowledge of medicinal chemistry, informatics, bioinformatics, computational chemistry, biology and biochemistry (9).

Prerequisite for successful use of SBDD is a detailed information about 3D protein structure, which includes position of every amino acid in 3D space, presence of ions important for structure maintenance or catalytic activity, hydrophobic and hydrophilic properties of ligand-enzyme binding pocket, etc. Advances in protein generation and purification techniques, as well as ever cheaper and more powerful computers have made SBDD more efficient and cost effective (9).

3D structure of proteins can originate from experimental data such as NMR spectroscopy, X-ray crystallography, or from computational homology modeling and this information can be found in Protein Data Bank (PDB) (9,10).

Protein structures obtained from PDB have a 4-character identification code, consisting of numbers (0-9) or letters from the Latin alphabet (A-Z) (10). PDB files usually lack important information for SBDD, such as protonation, formal atomic charge, hydrogen atoms etc., and this information must be generated *in silico.*

Next step in SBDD after acquiring necessary data for target structure is virtual screening (VS). VS is a computer assisted process in which new hits are discovered by screening molecules available in commercially available databases. Docking programs are used to examine different conformations of ligands in the proteins binding site, thereby predicting the structure of protein-ligand complex. Most docking programs assume a rigid protein structure, as a way to minimize computational resources needed, but it has been shown that inclusion of protein flexibility leads to better performance with regard to pose prediction (11).

Databases contain tens, if not hundreds of thousands of compounds, and to assess every possible target on a large protein with every compound in the database would require an immense amount of energy as well as time (12).

For this reason, researchers must decide on an approximate docking site, as well as decide the optimal box size. Screening libraries are usually composed of chemically diverse molecules, so to expedite the search for hits several molecular filters are uses. These filters are commonly a variation of Lipinski’s

rule of five, that can include original rules for druglikeness (number of hydrogen bond donors/acceptors, molecular mass less than 500 Daltons, log*P* not greater than 5 (where P is a measure of lipophilicity), as well as variations including number of rotatable bonds, limit of polar surface etc. Docking programs use scoring functions that approximate the binding free energy of a complex in order to rank compounds (13).

1.5 Molecular dynamics

Molecular dynamics (MD) simulation is a computer simulation method used to study the motions of atoms in a complex macromolecular system (e.g. protein) (14). Classical All Atom MD (AAMD) simulations calculate the motion of atoms in a complex system by numerically solving the Newtonian equations of motion

$$m_\alpha \ddot{\vec{r}}_\alpha = -\frac{\partial}{\partial \vec{r}_\alpha} U_{\text{total}}(\vec{r}_1, \vec{r}_2, \ldots, \vec{r}_N), \quad \alpha = 1, 2 \ldots N,$$

where m_a is the mass of atom a, r_a is its vector position, and U_{total} is the total potential energy dependent on the position of all atoms (15). AAMD simulations are calculated on time scale measured in nanoseconds and microseconds and are useful for studying conformational changes of large biological complexes, such as protein-ligand interactions, as well as protein-protein and protein-DNA/RNA interactions (16).

Most commonly performed analysis for calculated motion of MD simulation is the analysis of structural fluctuations of the macromolecular system, and two most common measures are the Root Mean Square Deviation (*RMSD*) and the Root Mean Square Fluctuations (*RMSF*) (14).

Most commonly performed analysis for calculated motion of MD simulation is the analysis of structural fluctuations of the macromolecular system, and two most common measures are the Root Mean Square Deviation (*RMSD*) and the Root Mean Square Fluctuations (*RMSF*) (14).

RMSD is a quantitative measure of the average distance between atoms at one point in the simulation and the reference structure, which is usually the first frame of the simulation. RMSD is defined with the equation:

$$RMSD(t) = \sqrt{\frac{1}{N}\sum_i^N \left\| x_i(t) - x_i^{ref} \right\|^2}$$

where N is the number of atoms in a system, $x_i(t)$ are the coordinates of N atoms at any time t of the simulation and x_i^{ref} are coordinates of a reference structure (14). Since RMSD is a measure of distance, it is expressed in length units, most commonly in Ångström (Å).

RMSF is a measure of the deviation between the position of an atom or group of atoms (e.g. amino acids), relative to the position of a reference particle, in time dependant manner. RMSF is defined with the equation:

$$RMSF_i = \sqrt{\frac{1}{T}\sum_{t=1}^{T}\langle (r'_i(t)) - r_i(t_{ref}))^2 \rangle}$$

where T is time in which RMSF is measured (or frames of simulation), r_i is the position of an atom (or group of atoms), r'_i represents the coordinates of atom (or group of atoms) after superimposition on the reference atom and t_{ref} is the reference time.

RMSD can be used to determine the stability of the protein-ligand complex. Comparing RMSD values of tested ligands against a reference inhibitor can give us insight into which compound has a greater probability of stabilizing the protein. When comparing RMSD values of two complexes, the lower RMSD means better stability, therefore better inhibitor.

2. Aim of the study

Purpose of this study was to test potential inhibitors of DNMT1, which are available from Elite and Maybridge databases, using *in silico* methods. For a tested compound to become eligible for further study, it should make a stable connection to three DNMT1 structures: 4WXX, 3PTA and 4DA4.

3. Materials and methods

The crystal structures of DNMT1 were obtained from Protein Data Base (PDB). Four-digit alphanumerical codes 4WXX and 3PTA are PDB codes for human DNMT1 proteins and 4DA4 is PDB code for murine DNMT1 protein (6,7,17). Visualization of protein, ligands and DNA were done using programs VMD, PyMOL and UCSF Chimera (18-20). Docking, MD and subsequent analysis was performed on chain A for proteins 4WXX and 3PTA. For protein 4DA4 chains A, C and D were used, and chains B, E and F were excluded from experiments. Since DNMT1 is a dimer, chains in proteins were excluded to avoid complexity, therefor minimizing time and resources for conducting experiments. Chain removal was possible because the binding site, as well as catalytic residues do not fall between the two chains.

The structures of ligands (Table 2) were obtained from Elite and Maybridge databases. Compounds were visualized using ChemAxon Marvin (21). We used MarvinSketch to perform a pK_a based protonation state check at pH=7,2 (21). Absorption, distribution, metabolism, and excretion (ADME) parameters for tested compounds were computed using a web tool SwissADME.

Molecular docking of Maybridge and Elite compounds on proteins was preformed using Autodock Vina on UCSF Chimera interface (22). To prepare protein for docking, first we deleted chains as mentioned, and water molecules in file downloaded from PDB. We used the Modeller interface in UCSF Chimera to add the missing residues.

We used Dock Prep tool with default settings to add hydrogens, assign charges to residues using AMBER ff99SB force field and saved the file in .pdbqt format.

Using AutoDock Vina in UCSF Chimera, we selected the protein and ligand of interest, and determined the size of a grid box, i.e. volume in which the program will search the optimal pose of the ligand. Grid box for ligands was set with dimension 30 Å × 25 Å × 30 Å surrounding the SAH binding site. Center of the grid box was set to O3 atom of the SAH ligand. Number of binding modes was set to 10, exhaustiveness of search to 8 and maximum energy difference to 3 kcal/mol. Modes with lowest binding energy were selected for further research.

CHARMM-GUI was used to prepare protein-ligand complexes for molecular dynamics and output analysis (23). Ligand parameters for the tested complexes were created in the Ligand Reader & Modeler module. Solution builder module was used to create an aqueous solution. Generated water box had 0.15 M potassium chloride (KCl), the ion placing method was Monte-Carlo, temperature was set to 303.15 K and pressure to 1 bar.

GROMACS suite version 2020.4 was used for minimization, equilibration and production (24). All protein-ligand complexes were propagated for 50 million steps with time step 0.002 ps, generating a simulation of 100 ns. When simulation was not stable, then we set 1.5 million steps to let it stabilize, and then continued to 50 million steps. All GROMACS calculations were performed on Bullx DLC720 supercomputer Bura using 2 nodes for each calculation.

VMD was used to visualize results from GROMACS, to export results in .pdb and .dcd formats as well as to calculate the number of hydrogen bonds formed between ligands and protein. Parameters for hydrogen bonds were set for donor-acceptor distance at 3.5 Å and angle cutoff at 20°.

RMSD and RMSF were calculated in RStudio using Bio3D package and in VMD to check calculations (25). Charts were created using LibreOfficeCalc from data generated in RStudio. RMSD was calculated for all atoms in proteins.

Methods for docking and MD simulations with SAH complexes are described in (26). Data for analysis were downloaded from Bura server.

LigPlot+ was used to analyze protein-ligand interactions at t=100 ns (27). With LigPlot+ we were also able to closely inspect distance of hydrogen bonds formed in protein-ligand complexes, as well as which atoms were involved and which amino acid residues.

4. Results

4.1 Docking the ligands to DNMT1

Docking a drug candidate to a protein produced 10 poses of ligand in protein-ligand complexes, where the protein remained in a fixed position and the ligand assumed different conformations. For each conformation/pose of the ligand, the program calculated a binding score, based on which we selected the initial geometry of protein-ligand complex for further simulation. We performed docking for molecules Elite1, Elite6, and Maybridge9 on the protein 4WXX (Figure 7). Elite6 docking was also performed on proteins 4DA4 and 3PTA (Figure 8 and Table 1).

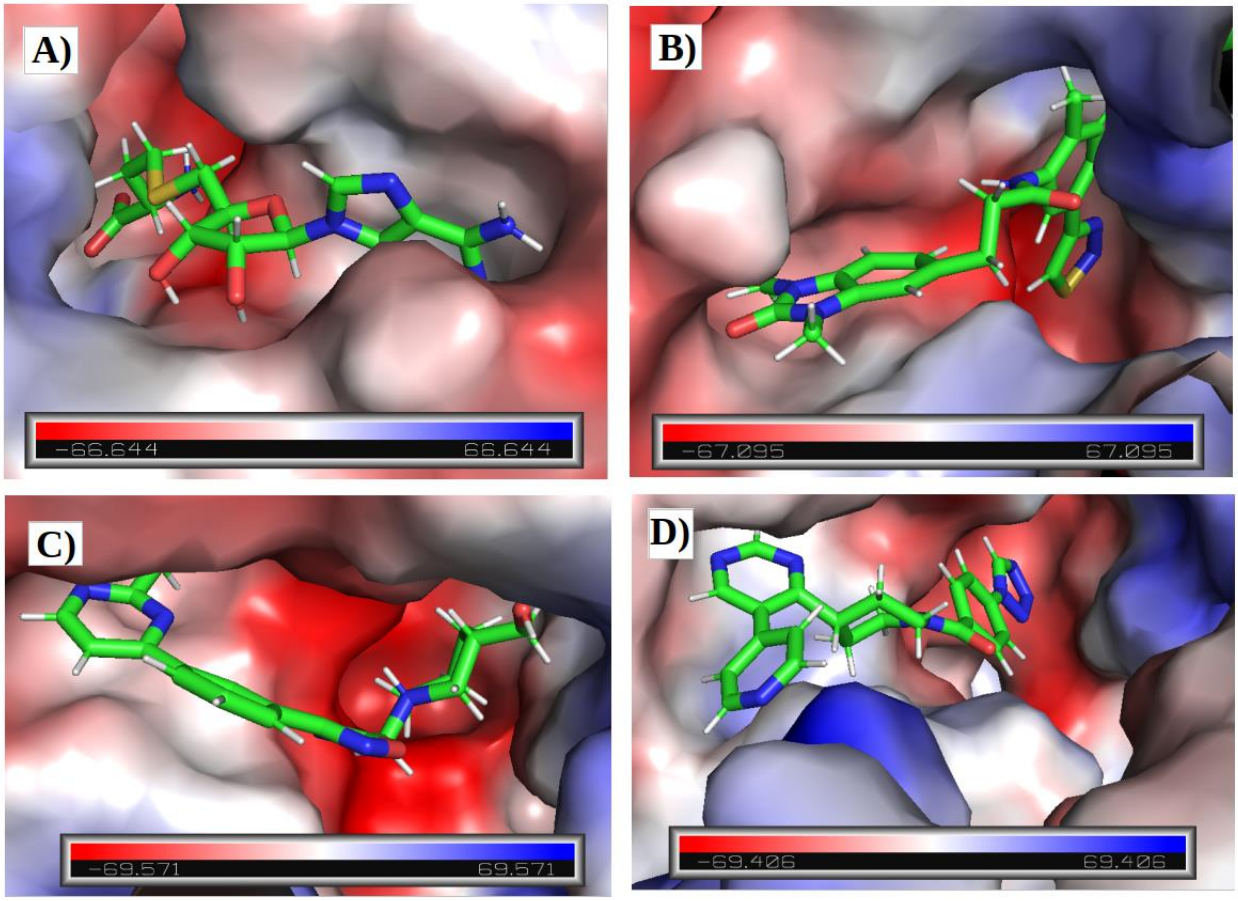


Figure 7. Ligands in the 4WXX binding pocket. A) SAH, B) Elite6, C) Elite1 and D) Maybridge9. The images were created using PyMol. The protein is colored with a vacuum electrostatics representation, which calculates the protein contact potential. The red areas are negative, blue positive and white neutral charge.

4WXX-Elite6		3PTA-Elite6		4DA4-Elite6	
Pose	Score	Pose	Score	Pose	Score
1	-9.6	1	-7.7	1	-10.1
2	-8.8	2	-7.5	2	-9.3
3	-8.7	3	-7.2	3	-9.2
4	-8.7	4	-7	4	-9
5	-8.6	5	-6.8	5	-8.9
6	-8.4	6	-6.8	6	-8.8
7	-8.1	7	-6.8	7	-8.5
8	-8.1	8	-6.7	8	-8.5
9	-8	9	-6.6	9	-8.1
10	-7.9	10	-6.5	10	-8

Table 1. AutoDock Vina results of the binding affinity (in kcal/mol) for different conformations of Elite6 in the DNMT1-Elite6 complex. Lower score indicate more favorable binding.

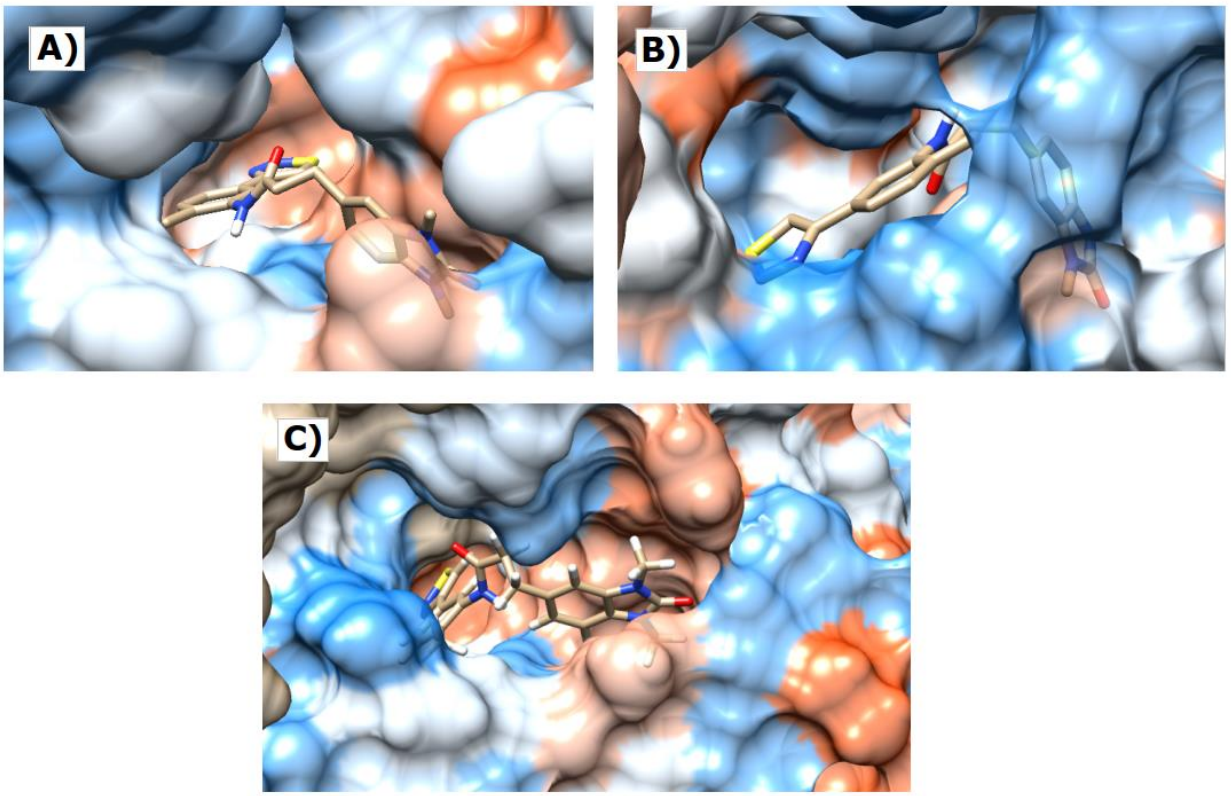


Figure 8. Elite6 docked in the binding pocket of DNMT1 with highest binding scores: A) 4WXX, B) 3PTA and C) 4DA4. Proteins are shown with hydrophobicity surface, where blue represents the hydrophilic surface, and orange represents the hydrophobic surface.

4.2 MD simulations of docked complexes

MD simulations for all three protein-SAH and protein-Elite6 complexes were performed for a time period of 100 ns, while for 4WXX-Elite1 and 4WXX-Maybridge9 the simulations were stopped at t=28 ns and t=84 ns, respectively (Figure 9) (26). RMSD analysis for protein-SAH and protein-Elite6 have similar values, and RMSF analysis showed a decrease in values for protein-Elite6 complexes. We found that protein-SAH complexes had formed more hydrogen bonds, as well as more hydrophobic interactions, compared to protein-Elite6 complexes.

After docking, we performed AAMD simulation with 4WXX-ligand complexes in an aqueous environment to study interactions between the 4WXX form of DNMT1 and the drug candidates Elite1, Elite6 and Maybridge9 and compared the results from the 4WXX-SAH simulation. Based on analysis of these complexes, we excluded Elite1 and Maybridge9 from further research. We proceeded with AAMD simulations for Elite6 on proteins 4DA4 and 3PTA and compared the results with the 4DA4-SAH and 3PTA-SAH complexes from (26).

4WXX-Maybridge9 has significantly higher RMSD values compared to other complexes, with the highest value of 20.348 Å at t=45 ns. Maybridge9 was excluded from further analysis due to its removal from the binding pocket, separation from the protein at t = 36 ns and the formation of a hydrogen bond at t = 46 ns with the protein, outside the binding pocket (Figure 10). For this reason, the simulation was terminated at t=84 ns. 4WXX-Elite1 simulation was abrupted at t=28 ns for similar reasons.

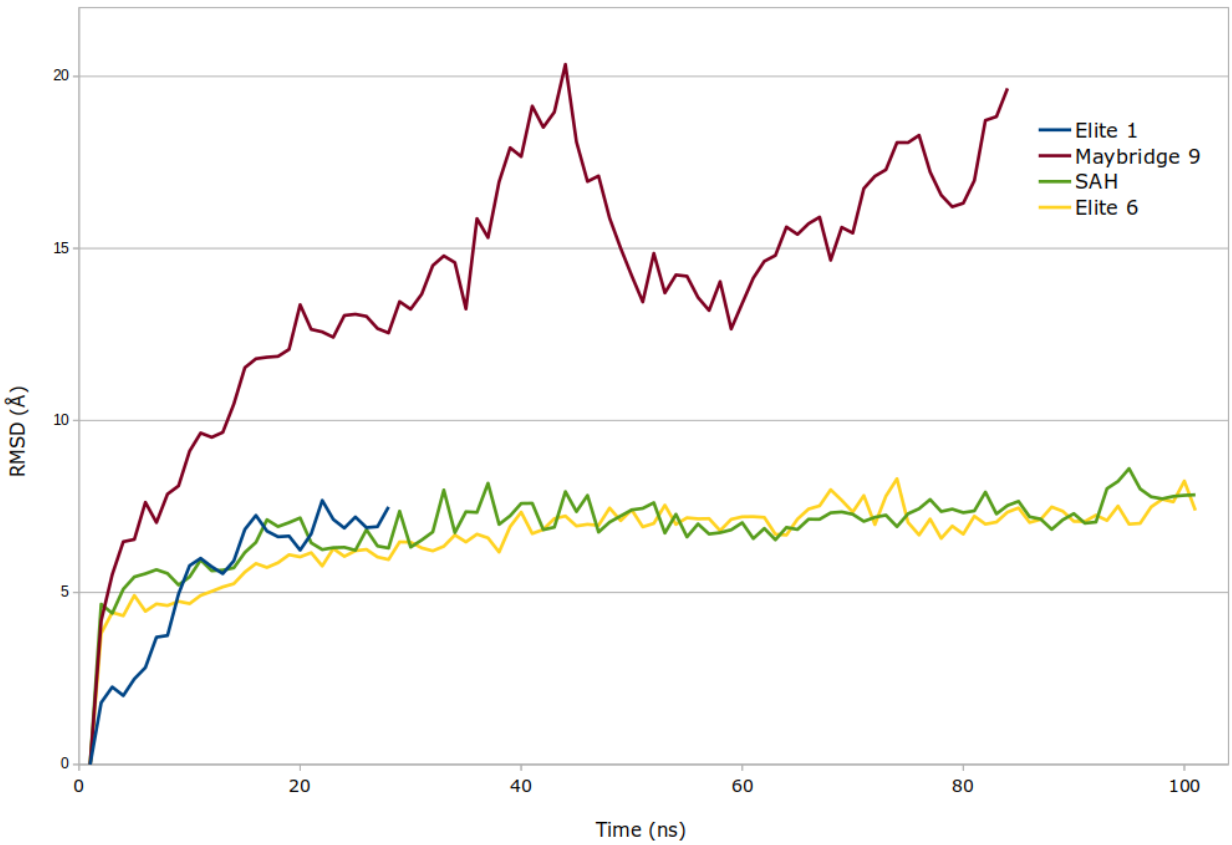


Figure 9. RMSD for tested compounds on 4WXX. RMSD values for Elite1 (blue), Maybridge9 (brown), SAH (green), and Elite6 (yellow). The RMSD values were calculated using RStudio and the plot was generated in LibreOffice Calc.

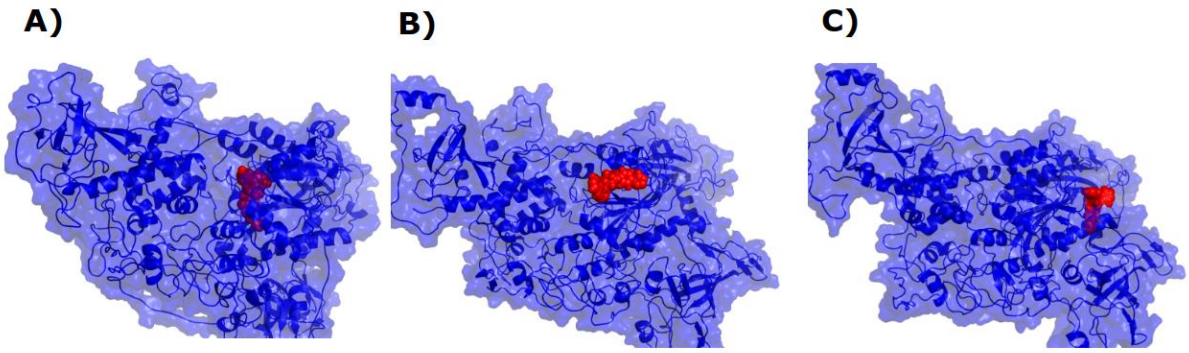


Figure 10. Maybridge9 has moved out of the DNMT1 binding pocket during the simulation. A) Maybridge9 at t=0 ns is in the binding pocket, B) Maybridge9 at t=35 ns does not have any bonds with DNMT1, C) Maybridge9 at t=63 ns is binding to DNMT1 but not in the binding pocket. The images were created using PyMol. 4WXX is colored in blue and Maybridge9 in red.

For Elite6 and SAH in complex with all three DNMT1 proteins, we performed RMSD and RMSF analysis. The number of hydrogen bonds and hydrophobic interactions was then calculated for protein-SAH and protein-Elite6 complexes. RMSD analysis for protein-SAH and protein-Elite6 complexes were calculated to assess their stability during 100 ns MD simulation. 4WXX-SAH converged after ~30 ns, having average RMSD fluctuation ~7 Å (Figure 11.A). 3PTA-SAH did not converged during 100 ns MD simulation and had the average RMSD fluctuation ~4 Å (Figure 11.B). 4DA4-SAH complex converged after ~15 ns with average RMSD value of ~3.5 Å (Figure 11.C)

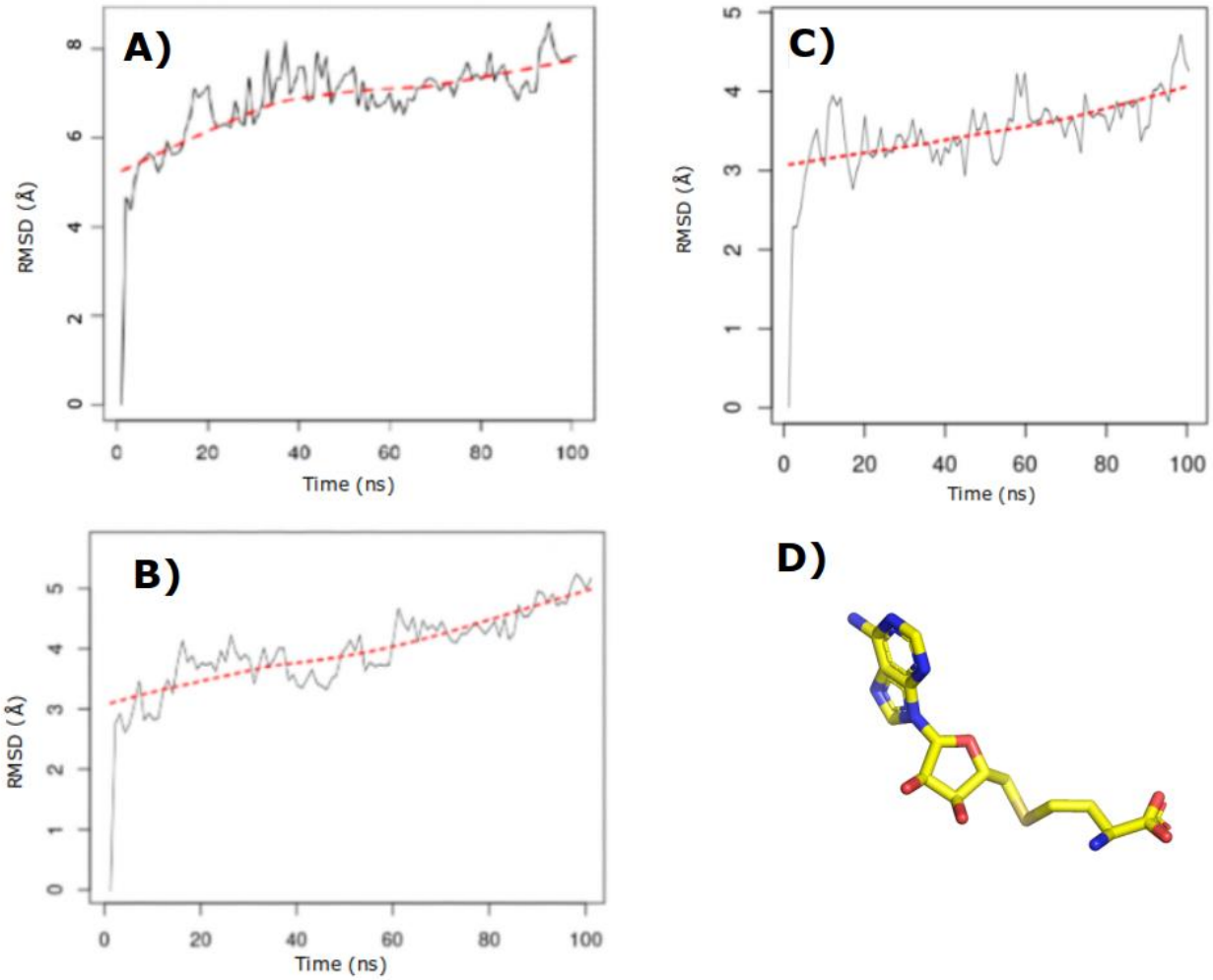


Figure 11. RMSD values and its trend line for the SAH complexes. A) 4WXX-SAH, B) 3PTA-SAH, C) 4DA4-SAH and D) SAH. RMSD values were calculated, and graphs generated using the bio3d package in RStudio. SAH image was created using PyMol.

4WXX-Elite6 complex started to converge at ~50 ns with average RMSD value of 7 Å (Figure 12.A). 3PTA-Elite6 reached equilibrium ~15 ns after the start of the MD simulation with average RMSD value of ~4 Å (Figure 12.B). 4DA4-Elite6 had the lowest average RMSD value of ~3 Å and started to converge after ~15 ns (Figure 12.C).

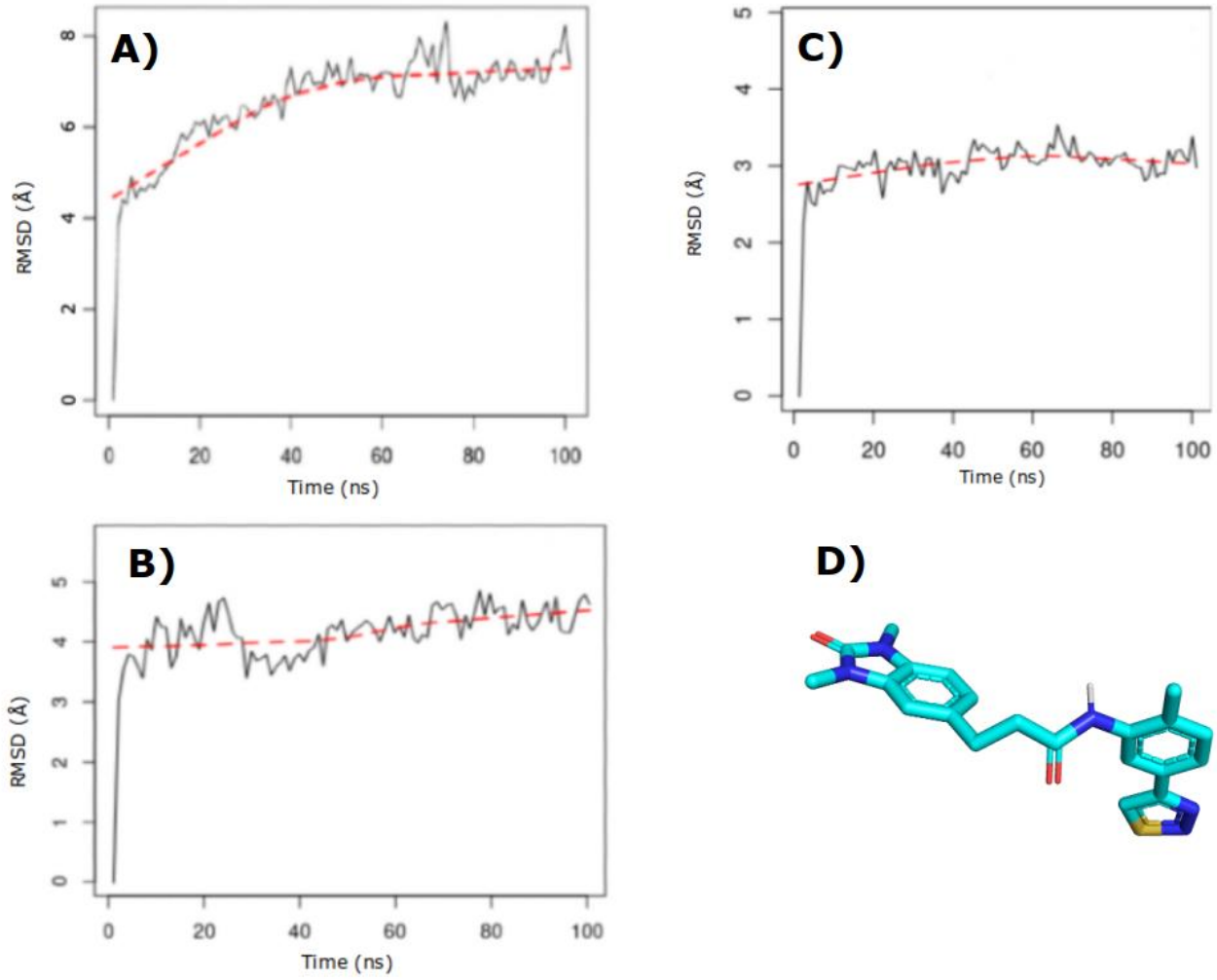


Figure 12. RMSD values and its trend line for the Elite6 complexes. A) 4WXX-Elite6, B) 3PTA-Elite6, C) 4DA4- Elite6 and D) Elite6. RMSD values were calculated, and graphs generated using the bio3d package in RStudio. SAH image was created using PyMol.

4.3 RMSF analysis of SAH and Elite6

By analyzing the binding of SAH to 4WXX, 3PTA and 4DA4, we were able to compare the binding of other drug candidates. The binding site for SAH on 4WXX and 3PTA includes residues Gly1145 - Leu1247 and Val1580, and on

4DA4 Gly1148 - Leu1250 and Val1583. SAH forms stable complexes with all three forms of DNMT1. The binding pocket for SAH is hydrophobic, and since SAH is a relatively hydrophilic molecule, drug candidates should also be hydrophilic to have a better chance of becoming drugs.

We calculated RMSF for protein-SAH and protein-Elite6 complexes in order to compare flexibility of DNMT1 protein when in complex with SAH and Elite6.

The average RMSF value for 4WXX-SAH complex was 2.51 Å and maximum 12.58 Å. There are five regions with higher RMSF than 8.0 Å, centered around Glu384, Pro639, Asp1038, Ala1381 and Trp1519 (Figure 13). RFTS domain had an average RMSF value of 3.24 Å. The average RMSF value for 4WXX-Elite6 complex was 2.35 Å with a maximum of 10.82 Å. Four regions had RMSF higher than 8.0 Å, with Pro351, Glu384, Ile627 and Arg920. RFTS domain had an average RMSF value of 3.52 Å (Figure 13).

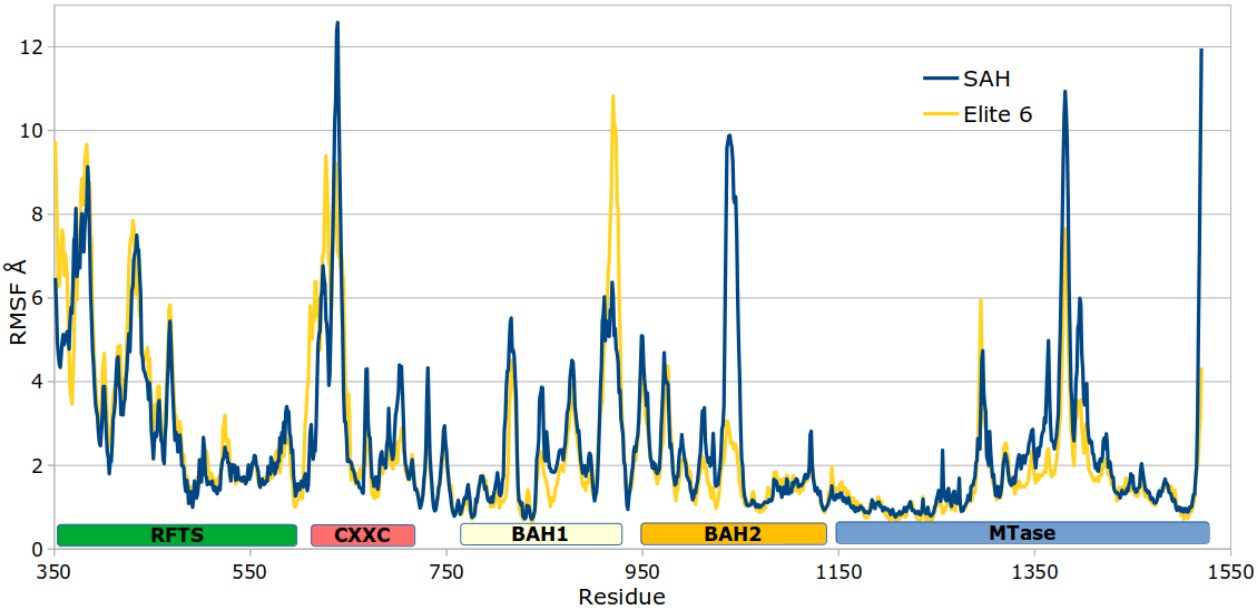


Figure 13. RMSF graphs for 4WXX-SAH and 4WXX-Elite6 complexes. Under the graphs are marked DNMT1 domains. Figures were created using LibreOfficeCalc.

The average RMSF value for 3PTA-SAH complex was 1.84 Å with maximum of 8.81 Å and 2 peaks with higher values than 8 Å (Gly665 and Val939). The average RMSF value for 3PTA-Elite6 was 1.77 Å with maximum of 9.41 Å. There was one region with higher RMSF of 8 Å, Val674 (Figure 14).

The average RMSF for 4DA4-SAH complex was 1.52 Å with maximum of 7.42 Å. The average RMSF value for 4DA4-Elite6 complex was 1.24 Å with a maximum of 4.57 Å (Figure 15).

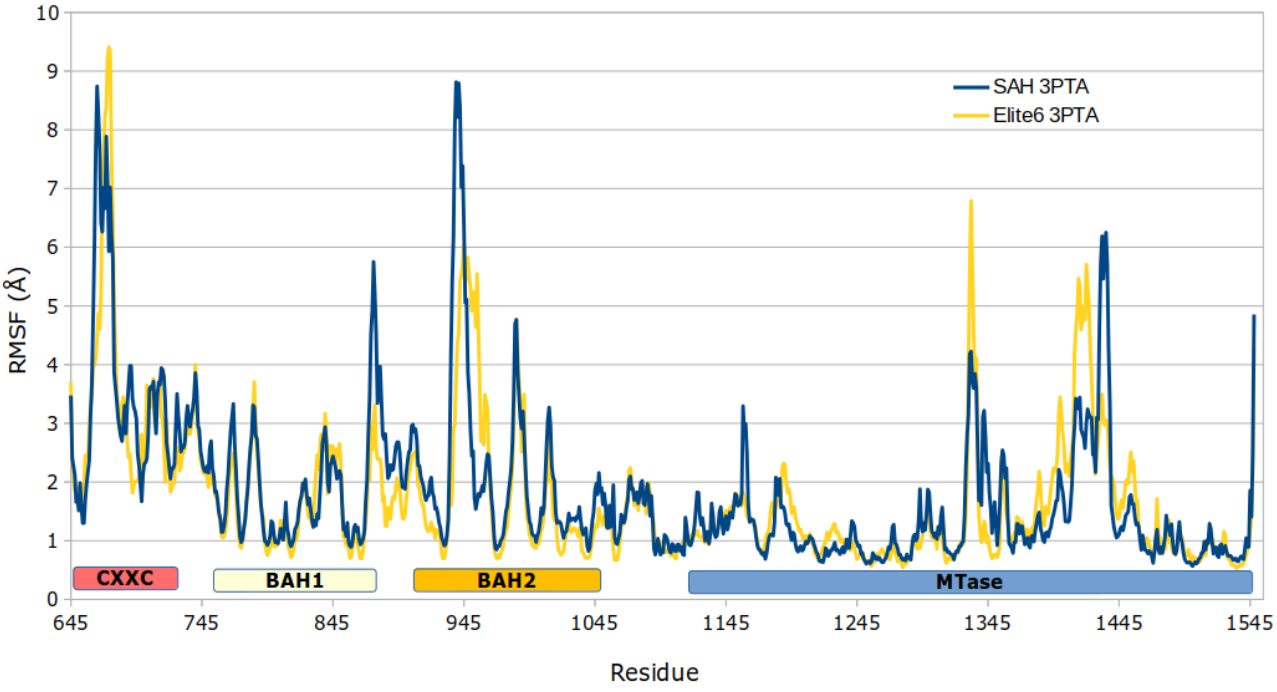


Figure 14. RMSF graphs for 3PTA-SAH and 3PTA-Elite6 complexes. Under the graphs are marked DNMT1 domains. Figures were created using LibreOfficeCalc.

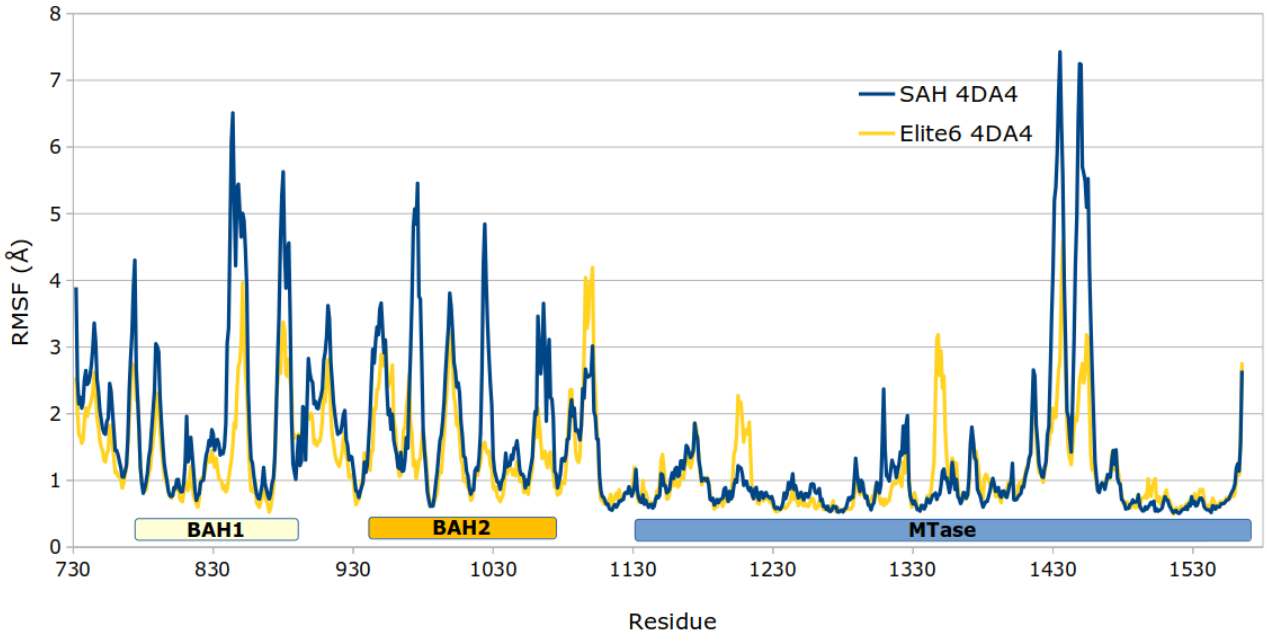


Figure 15. RMSF graphs for 4DA4-SAH and 4DA4-Elite6 complexes. Under the graphs are marked DNMT1 domains. Figures were created using LibreOfficeCalc.

4.4. Hydrogen bonds and hydrophobic interactions

Using VMD, we identified the number of amino acid residues that participate in formation of hydrogen bonds in protein-SAH and protein-Elite6 complexes during 100 ns MD simulation. All protein-SAH complexes formed more hydrogen bonds compared to protein-Elite6 complexes. 4WXX-SAH complex formed up to 9 hydrogen bonds, with average number of bonds during simulation was 3.83. 4WXX-Elite6 complex formed up to 5 hydrogen bonds during the MD simulation with the average number of bonds was 2.06 (Figure 16.A).

SAH formed up to 8 hydrogen bonds with proteins 3PTA and 4DA4, with the average 3.59 and 4.56, respectively (Figure 16.B).

Elite6 formed up to 3 hydrogen bonds with proteins 3PTA and 4DA4. Average number of hydrogen bonds during the simulation with 3PTA was 0.83 and with 4DA4 was 0.54 (Figure 16.C).

Analyzing the last frame of MD simulations of the 4WXX ligand (t = 100 ns) we found that 13 residues of the 4WXX protein are involved in hydrophobic contacts with the SAH ligand. Four residues that are involved in hydrophobic contacts also form hydrogen bonds with SAH. These residues are Leu1151 (d=2.67 Å), Val1580 (d=3.17 Å), Gly1150 (d=3.00 Å) and Glu1168 (d=2.82 Å) (Figure 17.A).

For 4WXX-Elite6, in the last frame we found 12 residues involved in hydrophobic interaction, of which three residues, beside hydrophobic interactions, are also involved in the hydrogen bond with Elite6 (Figure 17.B). The residues involved in hydrogen bonds and hydrophobic interactions are Glu1168 (d=2.57 Å), Val1580 (d=3.31 Å) and Gly1150 (d=3.22 Å).

3PTA had twelve residues involved in hydrophobic interactions with SAH and 11 with Elite6 (t=100 ns) (Figure 18.A and 18.B). For 3PTA-SAH, LigPlot+ calculated 3 hydrogen bonds at t = 100 ns, involving the residues Met1169 (d=3.34 Å), Cys 148 (d=2.91 Å) and Glu1168 (d=2.72 Å). For the 3PTA-Elite6 complex, LigPlot+ calculated 2 hydrogen bonds, including Phe648 (d=2.85 Å) and Gln1127 (d=2.89 Å).

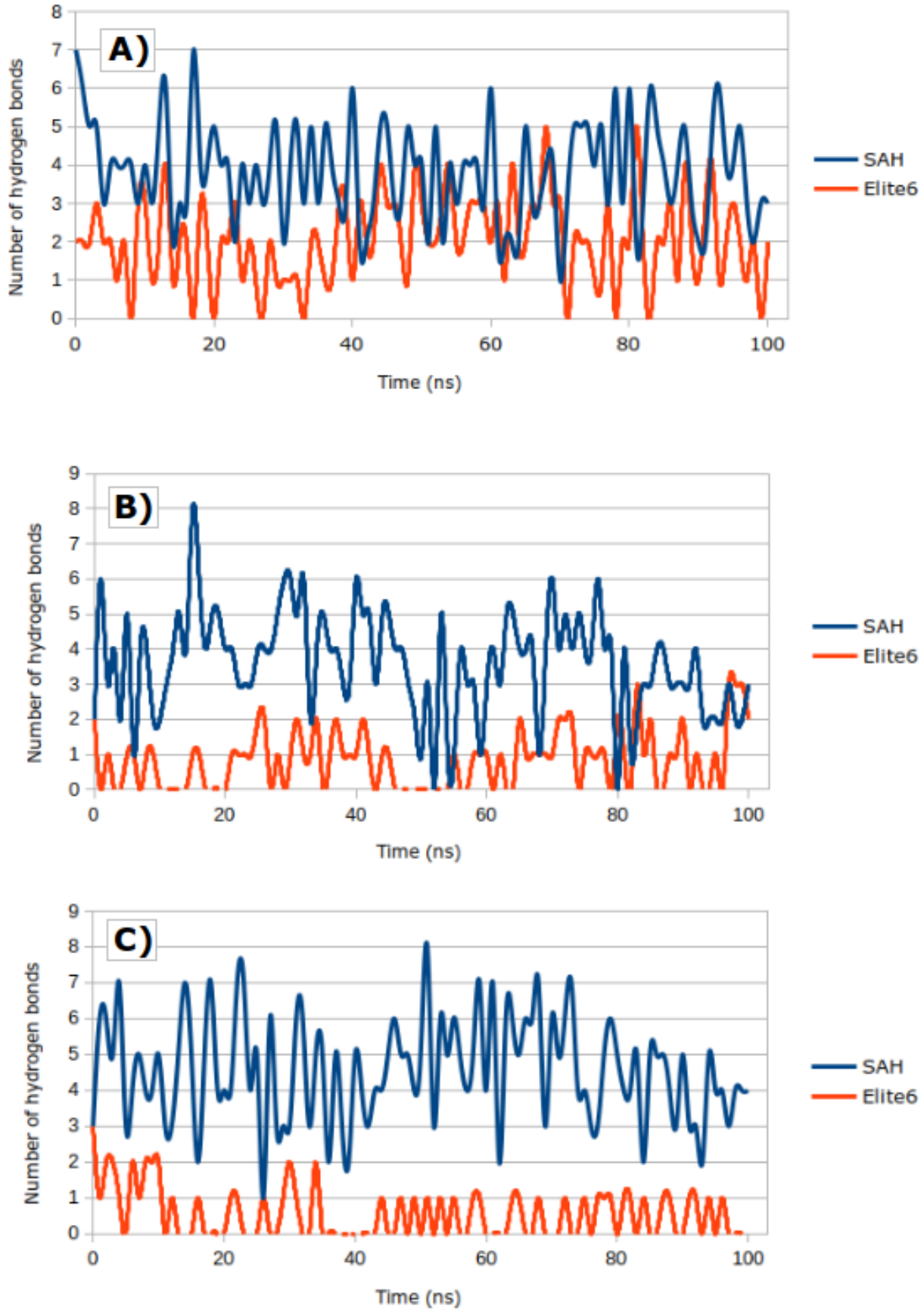


Figure 16. Comparison of hydrogen bonds. DNMT1: ligand complexes: A) 4WXX, B) 3PTA and B) 4DA4.

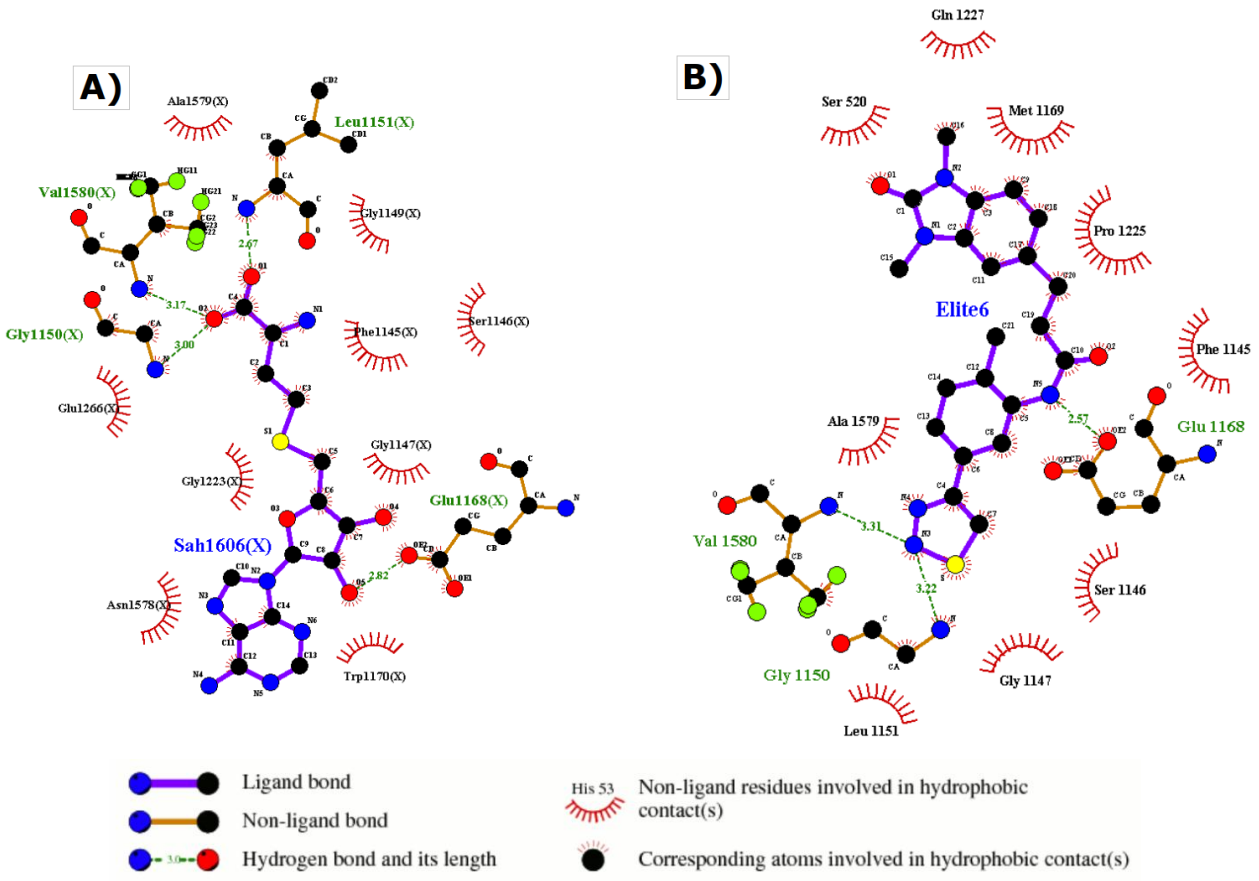


Figure 17. 2D representation of hydrophobic interactions. A)4WXX-SAH and B)4WXX-Elite6. Interactions were captured in the last frame of the MD simulation (t=100 ns). The lengths of the hydrogen bonds are expressed in Å.

The last frame of the 4DA4 MD simulation (t=100 ns) showed 18 residues involved in the hydrophobic interaction, of which seven residues form eight hydrogen bonds with the SAH ligand (Figure 18.C). Residues involved in hydrogen bonds are Gly153 (d=2.77 Å), Gly1152 (d=3.04 Å), Val1582 (d=3.16 Å), Leu1154 (d=2.83 Å), Cys1194 (d=3.22 Å), Asp1193 (d=2.69 Å) and also two hydrogen bonds involving Glu1171 (d=2,67 Å and d=2.56 Å).

Analyzing the last frame of the MD simulation 4DA4-Elite6, we counted fourteen hydrophobic interactions, of which one amino acid residue (Leu1115) is also involved in the hydrogen bond with Elite6 (Figure 18.D).

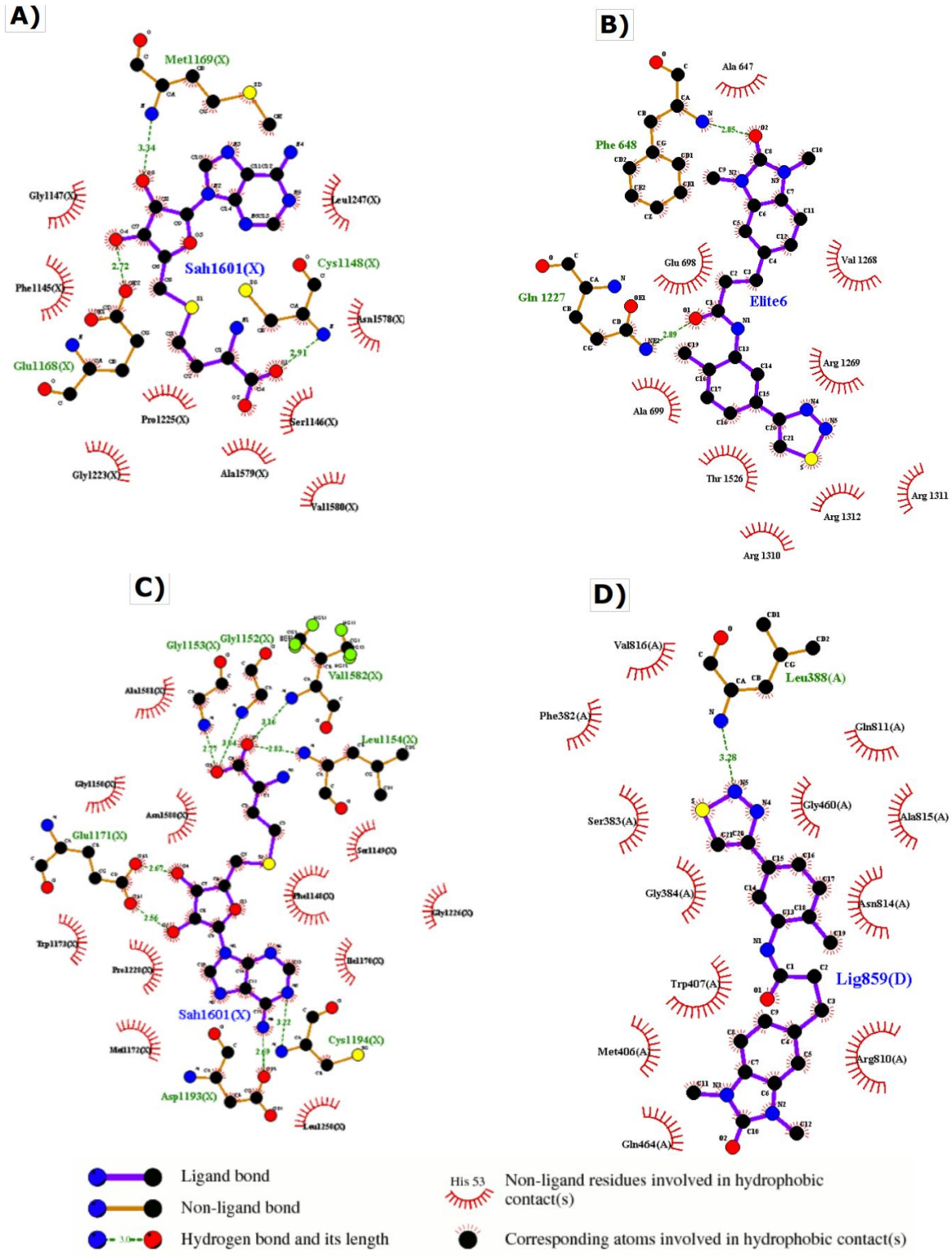


Figure 18. 2D representation of hydrophobic interactions. A) 3PTA-SAH, B) 3PTA-Elite6, C) 4DA4-SAH and D) 4DA4-Eliet6. Interactions were captured in the last frame of the MD simulation (t=100 ns). The lengths of the hydrogen bonds are expressed in Å. In the process of docking, residues shown on 4DA4-Elite6 are numbered 1 to 859, where Lig859(D) on D) represents the ligand Elite6, and Leu388(D) represents Leu-1115 in the original amino acid sequence.

After analyzing the 2D graphs generated in LigPlot+ we found π-π stacking interaction between Elite6 and 4DA4 as shown in Figure 18. Elite6 forms parallel-displaced π-π stacking interaction with the Phe648 residue. After closer inspection, we noticed another possible π-π stacking interaction between Elite6 and Trp1170 as shown in Figure 19. This interaction was not recognized by LigPlot+, only visually using VMD. Trp1170 is further from Elite6 than Phe648.

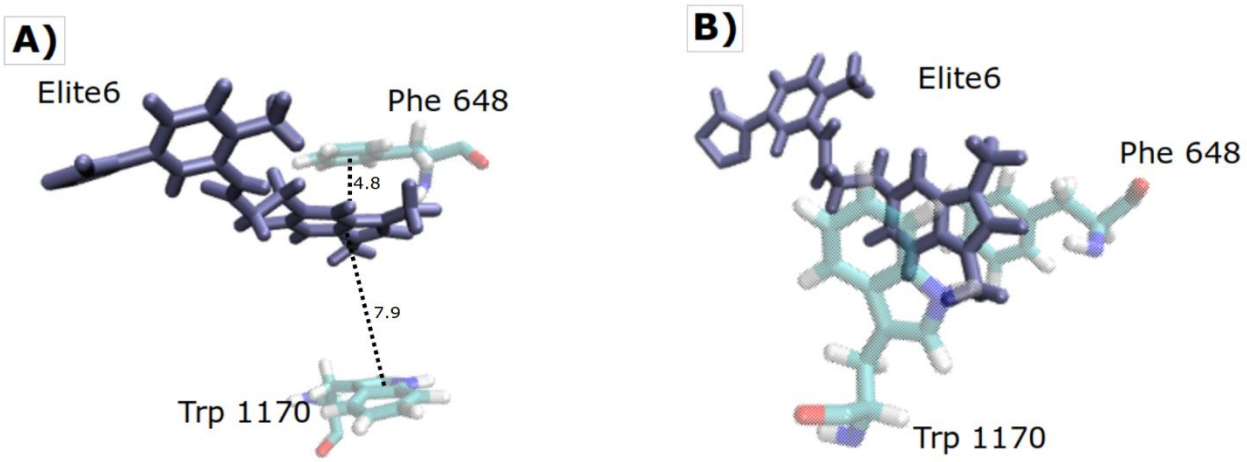


Figure 19. π- π stacking interaction between Elite6, Phe648 and Trp1170 (3PTA). A) Side view and B) Top view. The interaction was captured at t = 99 ns. Distance between centroid of residues and centroid of Elite6 is expressed in Å. Figure was created using VMD.

The chemical structures and IUPAC names of the tested ligands are shown in Table 2. Table 3 shows the ADME properties for the selected compounds. All tested compounds were calculated to be in the suitable physiochemical space for oral bioavailability and all three follow Lipinski's rule of 5. SAH has the highest number of hydrogen bond acceptors and donors (9 and 5, respectively), followed by Elite1 (7 and 1), Maybridge9 (7 and 0) and Elite6 (4 and 1). SAH also has the highest number of rotatable bonds (7), followed by Elite6 (6) and Elite1 and Maybridge9 both have 5.

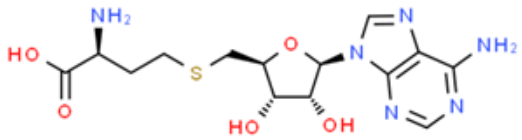
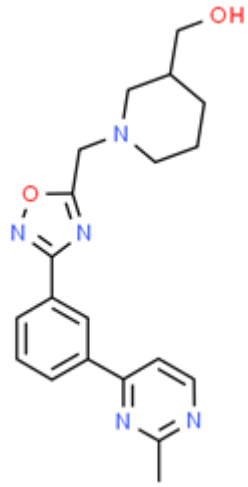
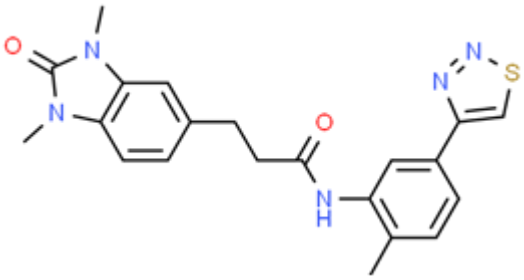
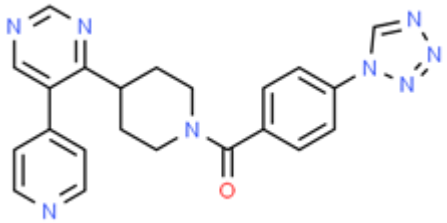
2D structure	Name used in study	IUPAC name
	SAH (*S*-adenosyl-L-homocysteine	(2*S*)-2-amino-4-[[(2*S*,3*S*,4*R*,5*R*)-5-(6-aminopurin-9-yl)-3,4-dihydroxyoxolan-2-yl]methylsulfanyl]butanoic acid
	Elite1	[1-({3-[3-(2-methyl-4-pyrimidinyl)phenyl]-1,2,4-oxadiazol-5-yl}methyl)-3-piperidinyl]methanol
	Elite6	3-(1,3-dimethyl-2-oxo-2,3-dihydro-1H-benzimidazol-5-yl)-N-[2-methyl-5-(1,2,3-thiadiazol-4-yl)phenyl]propanamide
	Maybridge9	{4-[5-(4-pyridinyl)-4-pyrimidinyl]-1-piperidinyl}[4-(1H-tetrazol-1-yl)phenyl]methanone

Table 2: Structures and IUPAC names of tested compounds. 2D structures were created using ChemSpider website, which also provided their IUPAC names.

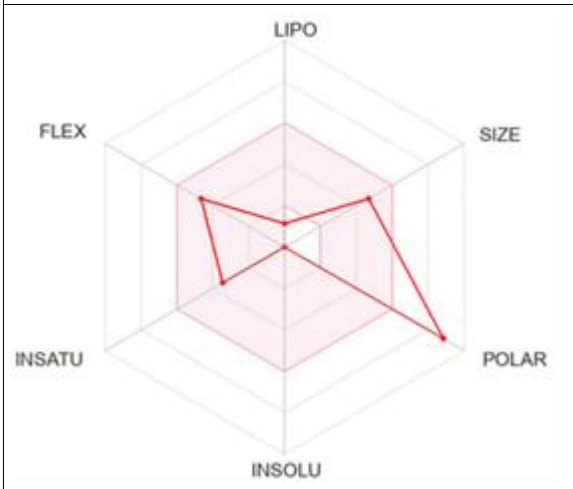
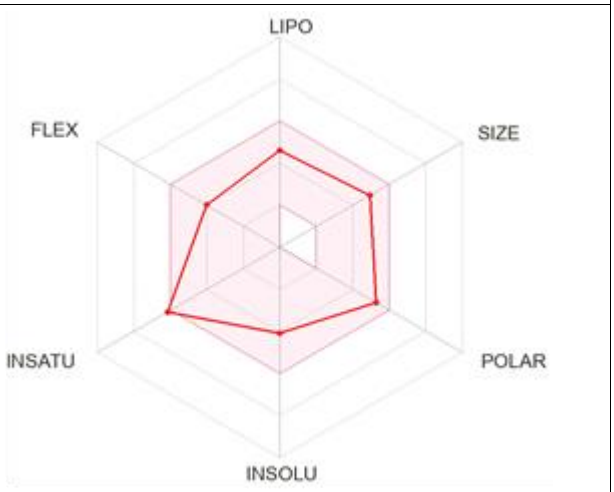
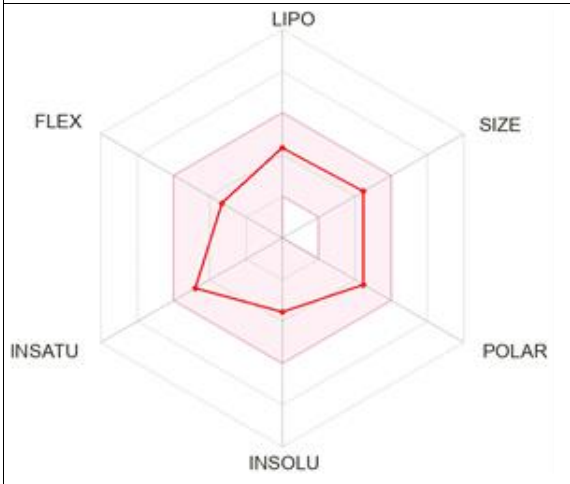
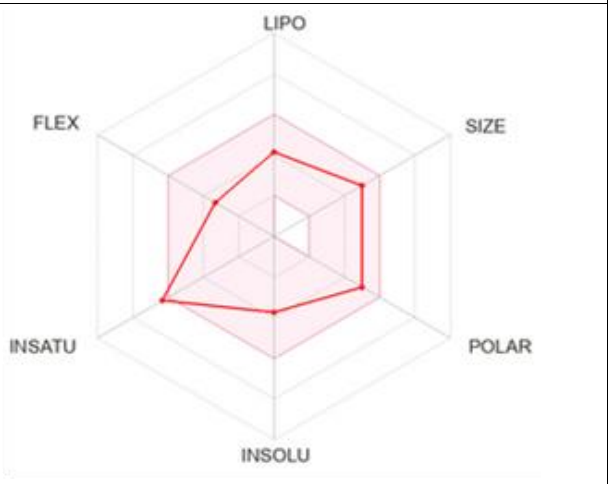
SAH	Elite6
	
Elite1	Maybridge9
	

Table 3. Bioavailability radar for tested compounds. The table presents bioavailability radar for SAH, Elite1, Elite6 and Maybridge9. The colored zone represents suitable physiochemical space for oral bioavailability. LIPO stands for lipophilicity, POLAR for polarity, INSOLU for solubility, INSATU for saturation and FLEX for flexibility.

5. Discussion

SAH is a DNMT1 inhibitor, and many researchers have tried to find a better inhibitor by modifying the SAH ligand, without promising results. We used the DNMT1-SAH binding pocket as a template and searched the ELITE and Maybridge databases to find and test *in silico* the best drug candidates. We found three distinct DNMT1 crystal structures with PDB codes 4WXX, 3PTA, and 4DA4. To become a good drug candidate, a molecule must form stable interactions with all three forms of DNMT1.

RMSD values were calculated to assess the stability of tested complexes. RMSD for all protein-Elite6 complexes show converging during 100 ns MD simulations, which implies that the simulations have stabilized (Figure 12). Compared to 3PTA-Elite6 and 4DA4-Elite6 complexes, 4WXX-Elite6 complex showed higher RMSD values and required longer time to reach equilibrium. This difference may be due to the higher residue content of 4WXX in comparison to 3PTA and 4DA4. The equilibrium of the 4WXX-SAH complex was reached after 50 ns, implying stability of the complex, but 3PTA-SAH and 4DA4-SAH did not converge during 100 ns MD simulations (Figure 11). The likely reasons for not stabilizing complexes during MD simulations are either insufficient docking of docked ligands, or more time is required for simulations to stabilize.

4WXX-Maybridge9 simulation was terminated at t=84 ns after Maybridge9 has moved out of the binding pocket. Possible reasons for its exit out of the binding pocket are that it has not made enough hydrogen bonds and/or hydrophobic interactions (Figure 10). Other reason could be that it was not properly docked. 4WXX-Elite1 simulation was stopped for similar reasons, and it moved out of the binding pocket for similar reasons as 4WXX-Maybridge9.

4WXX-Elite6 had an average increase in fluctuation for the RFTS domain by 0.28 Å compared to the 4WXX-SAH complex and overall, a 0.26 Å decrease in fluctuations of other domains. Greater fluctuation of RFTS domain implies that that domain is more active and since RFTS domain has an autoinhibitory role, this could go as an argument for Elite6 to be a better inhibitor than SAH. Of

particular importance is the decrease in RMSF values for MTase domain, which is responsible for catalyzation of methylation reaction. A decrease in activity in this domain would indicate that Elite6 is a stronger inhibitor than SAH.

RMSF analysis of 4WXX-ligands showed that 4WXX-Elite6 had a reduction in residue fluctuation in the BAH2 domain (residues His1036 - Glu1048), CXXC (residues Lys668 - Cys670), and MTase (residues Pro1375 - Leu1384) and also an increase in peak fluctuation around BAH1 zinc finger compared to 4WXX-SAH (Figure 13). 3PTA-Elite6 had higher RMSF values for the MTase domain (residues Phe1327 - Phe1330 and Phe1407 - Ser1418) and lower values for BAH1 (residues Asp870 - Glu878) and the MTase domain (residues Arg1426 - Gly1433) compared to 3PTA-SAH (Figure 14).

4DA4-Elite6 had a decrease in RMSF values across the whole protein except for residues Phe1096 - Asn1101 (BAH2 domain), residues Asn1205 - Arg1207, and Val1345 - Ser1352 (MTase domain) compared to 4DA4-SAH (Figure 15). These increases in RMSF values are between one and two Å, while 4DA4-SAH has RMSF values over four Å higher than those of 4DA4-Elite6.

All protein-SAH complexes formed more hydrogen bonds compared to protein-Elite6 complexes. Since neither ligand forms ionic bonds with protein, we can safely assume that protein-SAH bonds are stronger than that of protein-Elite6. Analyzing 2D graphs in LigPlot+ we found more hydrophobic interactions between SAH and protein than between Elite6 and protein, which is another evidence for stronger SAH binding to the protein than Elite6 binding to the protein.

Elite6 has one fewer rotatable bond compared to SAH, which is desirable in drug design, since more rotatable bonds the ligand has, it has a higher chance of binding to other proteins, thus increasing the risk of side effects. All tested ligands have higher lipophilicity compared to SAH, which indicates better absorption (Table 3).

6. Conclusion

DNMT1 is a large enzyme whose main purpose is DNA methylation. Dysregulations of this process can lead to various pathological conditions, including cancer. The structure of DNMT1 is known, but the function of each domain remains to be studied further.

We compared the binding of one ligand to three different DNMT1 structures to binding of SAH to the same DNMT1 structures. Due to fewer formation of hydrogen bonds and hydrophobic interactions in protein-Elite6 complexes compared to protein-SAH complexes, we concluded that Elite6 is not a viable candidate for further research.

Insight into SAH binding site gave us direction for further research. Future candidates should have more hydrogen bond acceptors and donors, to be able to compete with SAH, as well as taking advantage of potential π-π binding interactions.

7. References

1. Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of Epigenetic Drugs in Disease: An Overview. *Genet Epigenetics* 2014; 6: GEG.S12270.
2. Ren W, Gao L, Song J. Structural Basis of DNMT1 and DNMT3A-Mediated DNA Methylation. *Genes* 2018; 9: 620.
3. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* 2019; 20: 590–607.
4. Dhe-Paganon S, Syeda F, Park LC. DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. Int J Biochem Mol Biol 2: 58-66. *Int J Biochem Mol Biol* 2011; 2: 58–66.
5. Mazzio EA, Soliman KFA. Basic concepts of epigenetics: Impact of environmental signals on gene expression. *Epigenetics* 2012; 7: 119–130.
6. Song J, Rechkoblit O, Bestor TH, Patel DJ. Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. *Science* 2011; 331: 1036–1040.
7. Zhang Z-M, Liu S, Lin K, Luo Y, Perry JJ, Wang Y *et al.* Crystal Structure of Human DNA Methyltransferase 1. *J Mol Biol* 2015; 427: 2520–2531.
8. Yarychkivska O, Shahabuddin Z, Comfort N, Boulard M, Bestor TH. BAH domains and a histone-like motif in DNA methyltransferase 1 (DNMT1) regulate de novo and maintenance methylation in vivo. *J Biol Chem* 2018; 293: 19466–19475.
9. Andricopulo A, Salum L, Abraham D. Structure-Based Drug Design Strategies in Medicinal Chemistry. *Curr Top Med Chem* 2009; 9: 771–790.
10. Zardecki C, Dutta S, Goodsell DS, Voigt M, Burley SK. RCSB Protein Data Bank: A Resource for Chemical, Biochemical, and Structural Explorations of Large and Small Biomolecules. *J Chem Educ* 2016; 93: 569–575.

11. Lionta E, Spyrou G, Vassilatis D, Cournia Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. *Curr Top Med Chem* 2014; 14: 1923–1938.
12. Irwin JJ, Shoichet BK. ZINC − A Free Database of Commercially Available Compounds for Virtual Screening. *J Chem Inf Model* 2005; 45: 177–182.
13. de Azevedo WF (ed.). *Docking Screens for Drug Discovery*. Springer New York: New York, NY, 2019 doi:10.1007/978-1-4939-9752-7.
14. Martínez L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. *PLOS ONE* 2015; 10: e0119264.
15. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E *et al.* Scalable molecular dynamics with NAMD. *J Comput Chem* 2005; 26: 1781–1802.
16. Singh N, Li W. Recent Advances in Coarse-Grained Models for Biomolecules and Their Applications. *Int J Mol Sci* 2019; 20: 3774.
17. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-Based Mechanistic Insights into DNMT1-Mediated Maintenance DNA Methylation. *Science* 2012; 335: 709–712.
18. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. *J Mol Graph* 1996; 14: 33–38.
19. The PyMOL Molecular Graphics System. Schrödinger, L., & DeLano, W. (2020)
20. Pettersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC *et al.* UCSF Chimera?A visualization system for exploratory research and analysis. *J Comput Chem* 2004; 25: 1605–1612.
21. Kaushik M. A review of Innovative Chemical Drawing and Spectra Prediction Computer Software. *Mediterr J Chem* 2014; 3: 759–766.
22. Butt SS, Badshah Y, Shabbir M, Rafiq M. Molecular Docking Using Chimera and Autodock Vina Software for Nonbioinformaticians. *JMIR Bioinforma Biotechnol* 2020; 1: e14232.
23. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA *et al.* CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM,

and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. *J Chem Theory Comput* 2016; 12: 405–413.

24. Abraham MJ, Murtola T, Schulz R, Páll S, Smith JC, Hess B *et al.* GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. *SoftwareX* 2015; 1–2: 19–25.
25. B.J G, A.P.C R, K.M E, J.A M, L.S.D C. Bio3D: An R package for the comparative analysis of protein structures. *Bioinformatics* 2006; 22: 2695–2696.
26. Rijako,Morana. *Mammalian DNA methyltransferase DNMT1: structure and function analysis and development of novel inhibitors using in silico methods*. 2021.https://urn.nsk.hr/urn:nbn:hr:193:498195.
27. Laskowski RA, Swindells MB. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. *J Chem Inf Model* 2011; 51: 2778–2786.

8. Curriculum vitae

Viktor Morović

Date of birth: 17/05/1994 | **Nationality:** Croatian | **Gender:** Male | **Phone number:** (+385) 976092830 (Mobile) |

Email address: morovic.viktor@gmail.com | **Address:** Put Krčeva 12, 23000, Zadar, Croatia (Home)

EDUCATION AND TRAINING

10/2016 – 09/2019 Rijeka, Croatia

BSC IN BIOTECHNOLOGY AND DRUG RESEARCH Department of Biotechnology, University of Rijeka

Website biotech.uniri.hr

10/2019 – CURRENT Rijeka, Croatia

MSC IN MEDICINAL CHEMISTRY Department of Biotechnology, University of Rijeka

Website biotech.uniri.hr |

Thesis Analysis of DNA methyltransferase DNMT1 structure and function and development of noncovalent inhibitors using computational methods

WORK EXPERIENCE

05/2019 – 05/2019 Rijeka, Croatia

STUDENT INTERN JADRAN GALENSKI LABORATORIJ (JGL) AND INSTITUTE FOR PUBLIC HEALTH OF PGŽ

LANGUAGE SKILLS

Mother tongue(s): **CROATIAN**

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	B2	C1	B2	B2	B2

Levels: A1 and A2: Basic user; B1 and B2: Independent user; C1 and C2: Proficient user

DIGITAL SKILLS

Microsoft Office package: Microsoft Word, Excel, PowerPoint, Access | LIBREOFFICE | Molecular Visualization packages: UCSF Chimera, Visual Molecular Dynamics (VMD) | PyMol software for molecular visualization | CHARMM-GUI | LigPlot+ | GROMACS (Molecular Dynamics)

ADDITIONAL INFORMATION

DRIVING LICENCE

Driving Licence: B