Mammalian DNA methyltransferase DNMT1: structure and function analysis and development of novel inhibitors using in silico methods

Rijako, Morana

Master's thesis / Diplomski rad

2021

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:193:498195

UNIVERSITY OF RIJEKA

DEPARTMENT OF BIOTECHNOLOGY

Graduate programme

Drug research and development

Morana Rijako

Mammalian DNA methyltransferase DNMT1: structure and function analysis and development of novel inhibitors using in silico methods

Master's thesis

Rijeka, 2021.

UNIVERSITY OF RIJEKA

DEPARTMENT OF BIOTECHNOLOGY

Graduate programme

Drug research and development

Morana Rijako

Mammalian DNA methyltransferase DNMT1: structure and function analysis and development of novel inhibitors using in silico methods

Master's thesis

Rijeka, 2021.

Mentor: doc.dr.sc. Željko M. Svedružić, Ph.D.

SVEUČILIŠTE U RIJECI

ODJEL ZA BIOTEHNOLOGIJU

Diplomski sveučilišni studij

Istraživanje i razvoj lijekova

Morana Rijako

DNA metiltransferaza iz sisavaca DNMT1: razvoj novih inhibitora analizama strukture i funkcije enzima pomoću superračunala

Diplomski rad

Rijeka, 2021.

Mentor: doc.dr.sc. Željko M. Svedružić, Ph.D.

Acknowledgments

I would like to thank my mentor doc. dr. sc. Zeljko M. Svedruzic on sharing his knowledge and abilities with me. Many thanks for the motivation and every positive word. I was honored to work under your mentorship.

I would also like to thank all the members of the Laboratory for Biomolecular Structure and Function (BioSFLab) for sharing their knowledge, advice, and technical support.

Thanks to all friends and colleagues who supported me and participated in my accomplishments during my education.

I express my very profound gratitude to my parents and sister for providing me with support and continuous encouragement throughout my years of study. Special thanks and love to my little family Ivan and Irina, for their unconditional support, patience, and motivation through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

The thesis was defended on September 24, 2021 in front of committee members:

1. doc.dr.sc. Jelena Ban, Ph.D.

2. doc.dr.sc. Daniela Kalafatović, Ph.D.

3. doc.dr.sc. Željko M. Svedružić, Ph.D., mentor

The thesis has 55 pages, 22 images, 2 tables and 38 references.

Sažetak

Pozadina: DNA metilacija je osnovni i najdugotrajniji mehanizam u funkcionalnoj organizaciji ljudskog genoma. DNA metilacija se uspostavlja tijekom embriogeneze i traje do kraja života kao prvi korak epigenetske regulacije. Poremećaji DNA metilacije uzrokuju apoptozu, starenje te različita patološka stanja. Razvoj DNMT inhibitora je jedan od najvećih komercijalnih izazova za farmakologiju i biotehnologiju 21. stoljeća. DNA metiltransferaze su jedinstveno složen izazov jer se radi o velikim enzimima s višestrukim složenim regulatornim mehanizmima. Selektivna manipulacija DNA metilacije u zdravim tkivima može se koristiti za proizvodnju matičnih stanica. U ovom istraživanju smo koristili superračunalo kako bismo razvili inhibitore te analizirali strukturu i funkciju glavne ljudske DNA metiltransferaze DNMT1.

Rezultati: Mjesto vezanja kofaktora SAM u aktivnom mjestu enzima je evolucijski visoko očuvano što ukazuje da taj dio enzima najbolje poznajemo. 4WXX, 3PTA i 4DA4 su tri strukturalne forme DNMT1 iz PDB baze podataka. One prikazuju različite stupnjeve regulacije aktivnosti enzima sa suptilnim razlikama u aktivnom mjestu. Enzimski oblici su pokazali najveću mobilnost DNMT1 u RFTS, CXXC i BAH domenama. Različiti oblici enzima su stoga korišteni za pregled spojeva iz komercijalnih baza malih molekula u potrazi za spojevima koji se mogu najbolje vezati u aktivno mjesto DNMT1. Spojevi su početno selektirani metodom docking-a, čime su indetificirani spojevi koji stvaraju najveći broj interakcija i svojim oblikom najbolje odgovaraju aktivnom mjestu DNMT1. Najbolji spojevi su se koristili za opisivanje detaljnog mehanizma vezanja molekularnom dinamikom kroz 100 nanosekundi. Vezanje identificiranih spojeva se uspoređivalo s vezanjem kofaktora SAH kao referentnog liganda koji ima dobro poznat mehanizam vezanja.

Zaključak: ELITE12 i MYB1 su dva spoja iz komercijalnih baza malih molekula koji ne pokazuju značajno poboljšanje u stabilnosti i broju veznih interakcija u

odnosu na S-adenozil-homocistein. Razvoj malih molekula koje se vežu na aktivno mjesto DNMT1 ovisi o organskim sintezama i specifično dizajniranim spojevima.

Ključne riječi: DNA metiltransferaza DNMT1, inhibitori DNMT1, dizajn lijekova, molekularna dinamika

Abstract

Background: DNA methylation is considered a basic and longest-lasting mechanism of human genome organization. It is established during embryogenesis and lasts for the lifetime as the first step of epigenetic regulation. Dysregulation of DNA methylation causes apoptosis, aging, and various pathological conditions. The development of DNMT inhibitors is one of the greatest challenges for 21st-century pharmacology and biotechnology. DNA methyltransferases are a uniquely complex challenge considering the size and multiple complex regulatory mechanisms. Selective manipulation of DNA methylation in healthy tissues can be used to produce stem cells. In this study, we used a supercomputer to develop inhibitors and analyze the structure and function of the major human DNA methyltransferase DNMT1.

Results: The binding site of the SAM cofactor in the active site of the enzyme is evolutionarily highly conserved, indicating that this part of the enzyme is described the best. 4WXX, 3PTA, and 4DA4 are three structural forms of DNMT1 from the PDB database that show different degrees of enzyme activity regulation and subtle differences in the active site. DNMT1 showed the highest mobility in its RFTS, CXXC, and BAH domains. Different enzyme forms have been used to screen compounds from commercial small molecule libraries to find compounds with the best binding affinity for the DNMT1 active site. The compounds were initially selected by the docking method. By that, compounds that exhibit shape-complementarity with the DNMT1 active site and generate the largest number of interactions were identified. The best-ranked compounds were used to describe the detailed binding mechanism by MD over 100 ns. Binding modes of identified compounds were compared to the binding of cofactor SAH as a reference ligand with a well-known binding mechanism.

Conclusion: ELITE12 and MYB1 are two compounds from commercial compound libraries that do not show a significant improvement in stability and

number of binding interactions relative to SAH. The development of small molecules with binding affinity to the DNMT1 active site depends on organic syntheses and specifically designed compounds.

Keywords: DNA methyltransferase DNMT1, DNMT1 inhibitors, drug design, molecular dynamics

Table of contents

1. Introduction 1
   - 1.1. DNA methylation in mammalian cells 3
   - 1.2. DNA methyltransferases in human cells 6
     - 1.2.1. Structure of human DNA methyltransferase DNMT1 and the significance of its proper regulation 7
   - 1.3. Inhibitors of DNMT1 as potential drug candidates 10
   - 1.4. *In silico* drug design: Structure-Based Drug Design (SBDD) 12
   - 1.5. The role of Molecular Dynamics 13
   - 1.6. DNA-substrate dependent DNMT1 structures 15
2. Aim of the study 20
3. Materials and Methods 21
4. Results 24
   - 4.1. Insight into the SAH binding site 24
   - 4.2. Binding analysis of selected compounds to SAH binding site 27
     - 4.2.1. Molecular dynamics simulations point to binding modes of ELITE12 and Maybridge1 inhibitors 30
     - 4.2.2. Binding analysis of MYB1 and ELITE12 compounds at 4WXX relative to the reference SAH 30
     - 4.2.3. Binding analysis of SAH and MYB1 to 3PTA and 4DA4 35
   - 4.3. Coarse-Grained analysis of different forms of DNMT1 41
5. Discussion 44
   - 5.1. ELITE12 and MYB1 can bind to the SAH binding pocket 44
   - 5.2. Inhibitors show different stability in the protein-inhibitor complex when binding to different forms of DNMT1 45
   - 5.3. The high level of mobility is attributed to the sites that bind the DNMT1 interaction partners 46
6. Conclusion 48
7. References 50
8. Curriculum vitae 54

Introduction

Human body consists of about 200 different cell types that perform specific physiological function. Every cell in the human body, except erythrocytes, contains the entire genomic DNA, but only a small fraction of those genes is needed for cellular function. The cells specific functions are regulated by the functional organization of human genome. The epigenome is the network of DNA methylation and histone modification patterns. Both play a key role in regulating gene expression, and have to be faithfully maintained by the cell physiology and conserved during cell division (1).

Genetics, epigenetics, DNA replications, and DNA repair are the mechanisms that regulate gene activity. Epigenetics encompasses all processes which result in heritable changes in the cell phenotype and gene activity without changing the DNA sequence. There are various types of epigenetic processes, including acetylation, methylation, phosphorylation, ubiquitylation. Figure 1 shows major epigenetic modifications, including DNA methylation, histone modifications, and non-coding RNAs (ncRNAs). Epigenetic events play a role in maintaining normal body functions. Various environmental factors can disrupt the normal implementation of the processes. Dysfunction of epigenetic modifications can result in the development of pathological conditions, including cardiovascular, metabolic, or neurological disorders, as well as tumor development (2).

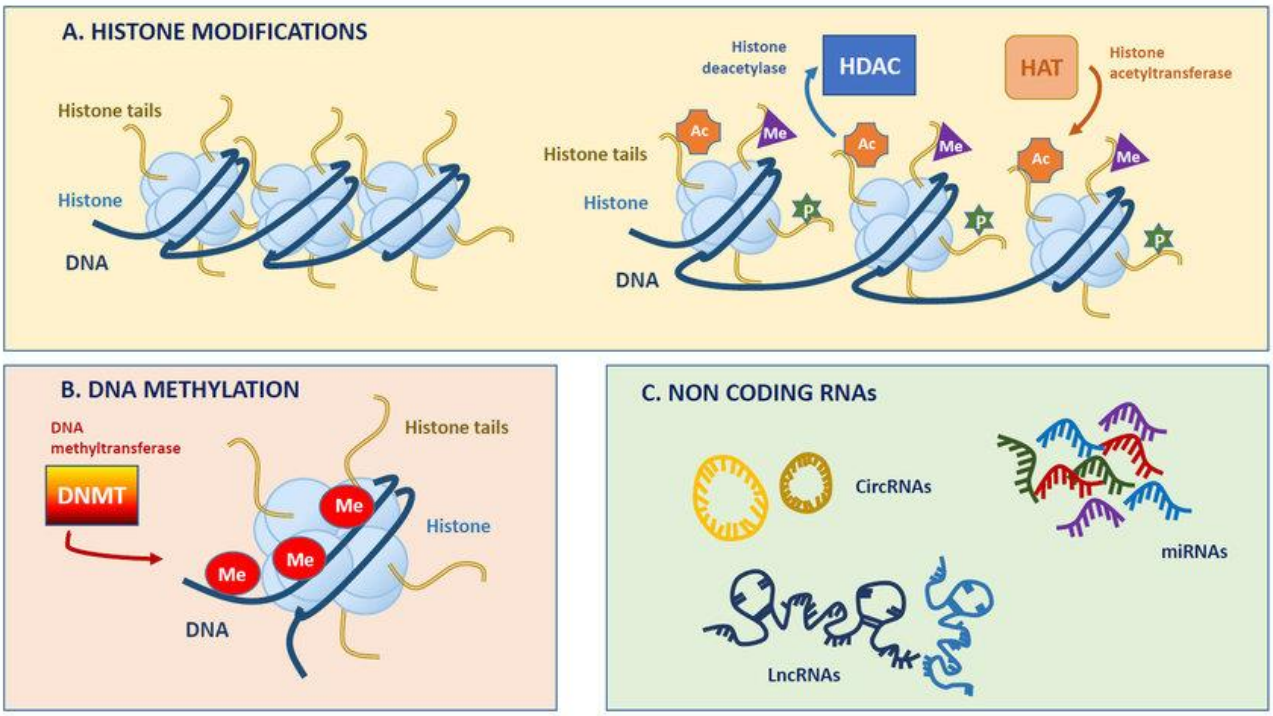


Figure 1. Overview of main epigenetic events. A) Histones are proteins around which DNA is packaged into nucleosomes. Post-translational modifications of histone tails (phosphorylation, acetylation) regulate chromatin transcription. Deacetylation of histone tails, with help of histone deacetylase (HDAC), ensures the availability of DNA for DNA methylation. B) DNA methylation process implies adding of a methyl group to the C5 position in cytosine by DNMT enzymes. C) Non-coding RNAs are functional RNA molecules grouped into small and long ncRNAs (LncRNAs). Their function is to control epigenetic pathways at the transcriptional and post-transcriptional levels. (Source: Coco C, Sgarra L, Potenza M, Nacci C, Pasculli B, Barbano R, et al. Can Epigenetics of Endothelial Dysfunction Represent the Key to Precision Medicine in Type 2 Diabetes Mellitus? International Journal of Molecular Sciences. 2019 Jun 17;20:2949. DOI:10.3390/ijms20122949. (3))

DNA methylation in mammalian cells

DNA methylation serves as the principal epigenetic modification for the establishment and maintenance of hereditary methylation pattern. Epigenetic changes at the DNA methylation level affect gene activation or silencing (4). The process originates from bacteria and has remained in most eukaryotic cells throughout evolution (5).

DNA methylation is catalyzed by a family of enzymes called DNA methyltransferases (DNMTs). It implies the transfer of the methyl group onto the C5 position of the cytosine (C) to form 5-methylcytosine (5mC). The position of 5mC in the DNA is limited mostly to the CpG dinucleotide sequences, regions of DNA that contain cytosine followed by guanine (G) along 5' → 3' direction (5). CpG areas are present in smaller quantities throughout the whole genome. Higher frequency has been recorded in short sections of DNA called CpG islands. They are defined as regions with at least 200 base pairs (bp), a GC percentage greater than 50%, and an observed to expected CpG frequency ratio greater than 60%. CpG islands are found to be located in the gene promoter region of approximately half of the human genes. In healthy mammalian cells, they are characterized by the absence of methylation (6). Although the content of CpG regions in the mammalian genome is reduced by approximately 5-fold relative to the expected amount, the level of CpG methylation reaches 70-80%. Since DNA methylation represses transcription, it is crucial for the lifetime silencing of germline-specific genes, imprinted genes, and genes on the inactive X chromosome. This ensures normal somatic cell function. Germline-specific genes are actively transcribed during embryogenesis, while their expression in somatic cells may result in tumor development (5). Except in embryogenesis, genomic imprinting also plays a major role in pre-natal and post-natal development. Each gene contains one paternal and one maternal allele. Genomic imprinting refers to

monoallelic gene expression, where one allele is silenced by DNA methylation depending on the parental background (7). DNA methylation also contributes to the inactivation of one X-chromosome in female mammals, probably as a final step in gene silencing (5).

Three stages of DNA methylation are described: establishment or de novo methylation, maintenance, and demethylation. DNMT3A and DNMT3B are enzymes responsible for the establishment of the initial CpG methylation patterns during embryogenesis through two periods of reprogramming: fertilization and germ cell specification (5). In somatic cells, already established methylation is maintained by DNMT1 during each cycle of mitosis. This assures proper epigenetic alterations in both DNA strands (Figure 2) (4). Demethylation of DNA may be conducted passively or actively. Passive demethylation of 5mC occurs as a result of DNMT1 absence through several generations of mitosis. Active DNA demethylation is carried out by a group of enzymes called TET methylcytosine dioxygenases. In this pathway, 5mC is converted to 5-hydroxylmethyl cytosine (5hmC), 5-formylcytosine (5fC), and 5-carboxylcytosine (5caC), respectively (5). The exchange of methylation and demethylation is of great importance in the regulation of gene expression during pre-natal human development (Figure 3). What is more, in somatic cells of adult individuals DNA methylation is globally decreased with aging (Figure 4). A possible reason for this is deregulated expression of genes encoding DNMTs (8).

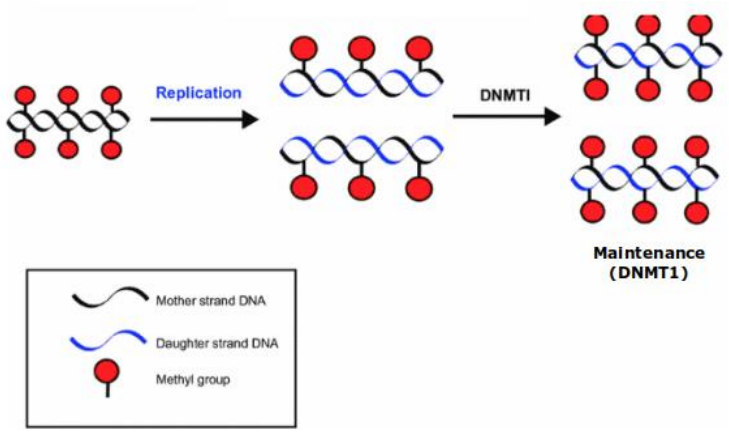


Figure 2. The basics of DNA methylation maintenance. During cell division DNA replication occurs. DNA molecules in daughter cells consist of one mother strand and one newly formed, daughter strand. DNMT1 uses the mother strand with an already established methylation pattern, as a template for methylating daughter strand. (Source: Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of Epigenetic Drugs in Disease: An Overview. Genetics & epigenetics. 2014 May 27;6:9–19. DOI:10.4137/GEG.S12270. (9))

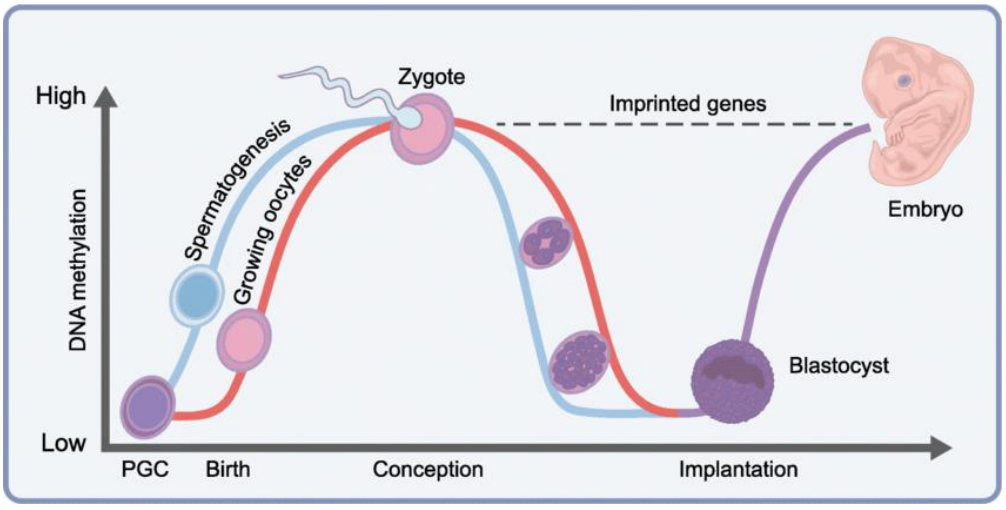


Figure 3. Changes in DNA methylation levels during embryo formation. Methylation patterns are shown for male (blue line) and female (red line) embryos. Fertilization, the process of zygote formation, is characterized by high levels of DNA methylation. After fertilization, passive demethylation occurs. During this period, the cells lack the establishment of an epigenetic pattern, which ensures the totipotency of the zygote. Imprinted genes retain high levels of methylation. Soon after implantation of the blastocyst, DNA methylation is re-established to ensure epigenetic regulation of the genome. (Source: Fernandez-Twinn D, Hjort

L, Novakovic B, Ozanne S, Saffery R. Intrauterine programming of obesity and type 2 diabetes. Diabetologia. 2019 Oct 1;62. DOI:10.1007/s00125-019-4951-9. (10))

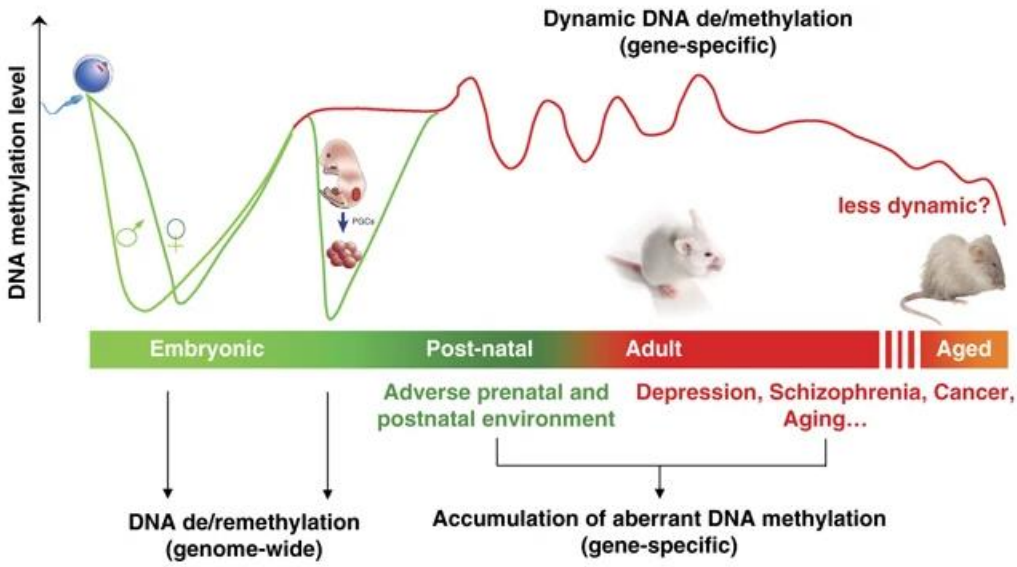


Figure 4. Dynamics of DNA methylation during embryonic development and adult life in animals. The embryonic development is followed by fetal development where the remethylation occurs due to cell specification. The exchange of DNA methylation and demethylation during adulthood is dynamic and depends on environmental factors. DNA methylation is thought to become less dynamic with age. Changes in DNA methylation levels are associated with a variety of behavioral and pathological conditions. (Source: Covic M, Karaca E, Lie DC. Epigenetic regulation of neurogenesis in the adult hippocampus. Heredity. 2010 Jul;105(1):122–34. (11)).

DNA methyltransferases in human cells

Several types of DNMTs are present in humans. Each enzyme is encoded by a separate gene. By the interaction of regulatory and catalytic domains and cofactors, DNMTs can change the location of action or regulation of genes. The human genome contains genes whose expression results in three catalytically active DNMT enzymes: DNMT1, DNMT3A, and DNMT3B (4,5). DNMT3A and DNMT3B carry out de novo DNA methylation. The main role of DNMT1 is to

maintain epigenetic methylation marks upon DNA replication. Catalytically inactive enzyme, methyltransferase-like protein (DNMT3L), interacts with DNMT3A and DNMT3B, thus enhancing their activity in the germline (5).

Structure of human DNA methyltransferase DNMT1 and the significance of its proper regulation

DNMT1 is the most abundant enzyme of the DNA methyltransferase family in humans and one of the largest cell enzymes. The human DNMT1 sequence contains 1616 amino acid residues. It is made up of two major subunits - regulatory and catalytic. Their mutual interaction results in a functional enzyme. Regulatory subunit takes amino terminus (N-terminus) of DNMT1, including residues 1-1139. Catalytic DNA methyltransferase domain (MTase) corresponds to residues 1140-1616. These two subunits are linked together by 12 glycine-lysine (Gly-Lys) alternating sequences (4). Structural insight into DNMT1 domains is shown in Figure 7.

On the cell level, the regulatory subunit interacts with more than 40 macromolecules, including proteins and DNA. There are several regulatory sub-domains: the DNA methyltransferase associated protein 1 (DMAP1)-binding domain, the Replication Foci Targeting Sequence (RFTS or RFD) domain, the CXXC zinc finger domain (zf-CXXC), and two Bromo-Adjacent Homology (BAH1, BAH2) domains. The DMAP domain consists of amino acid residues 1-115. It binds DMAP1, an interaction partner of DNMT1. DMAP1 serves as a co-repressor of transcription. It enhances the availability of DNA for the methylation process by interacting with histone deacetylase 2 (HDAC2). The RFTS domain (residues 350-600) appears to direct DNMT1 to replication foci during the S stage of the cell cycle when DNA replication occurs. It also interacts with E3 ubiquitin-protein ligase (UHRF1) whose SET- and RING-

associated (SRA) and ubiquitin-like (UBL) domains are the key mediators included in leading DNMT1 to the emerging DNA. When there is no catalytic activity, DNMT1 assumes an auto-inhibitory configuration characterized by RFTS domain embedded into MTase domain. The SRA domain recognizes and binds unmethylated CpGs. The UBL domain binds to the RFTS-MTase complex. That results in the separation of RFTS and MTase domains, thereby allowing a methylation process (4,5). The CXXC domain (residues 621-698) is rich in cysteine regions and contains a zinc-finger-binding domain. Its function is to enhance the binding of DNA substrate by binding hemimethylated CpG DNA portions and zinc fingers located within the catalytic site. The CXXC domain is characteristic of DNMT1. Its absence has been reported in DNMT3A and DNMT3B enzymes. CXXC and BAH domains are connected via a linker responsible for CXXC and RFTS-mediated autoinhibition (12). The absence of BAH domains, even in the presence of the RFTS domain, results in DNMT1 inactivity and thus impaired maintenance of DNA methylation. That confirms that BAH1 (residues 755-880) and BAH2 (residues 931-1100) are crucial for directing DNMT1 to sites of replication (13). On the N-terminal region, there are also short motifs that bind proteins and DNA. More precisely, residues 162-171 interact with proliferating cell nuclear antigen (PCNA), while three nuclear localization signal sites (NLS) occupy residues 191-211, 259-378, and 630-757. PCNA is a protein included in cell and DNA replication machinery. This indicates the importance of the PCNA-binding motif in the recruitment of DNMT1 to replication and repair sites during cell division and differentiation. The role of NLS is still being explored. Moreover, there are DNA-binding motifs included in the allosteric activation of DNMT1 by binding to methylated DNA (4).

The binding pocket on the catalytic subunit binds S-adenosylmethionine (SAM or AdoMet) (Figure 5) through residues 1150-1151, 1168-1169, and 1190-1191, known as the FxFxG motif (4,14). SAM is formed as a direct metabolite

of methionine. It serves as a methyl donor for almost all DNA and RNA methylation reactions in the body. Irreversible demethylation results with the conversion of SAM to S-adenosyl homocysteine (SAH or AdoHcy). SAH is the metabolic precursor of homocysteine and potent inhibitor of all methyltransferases (15).

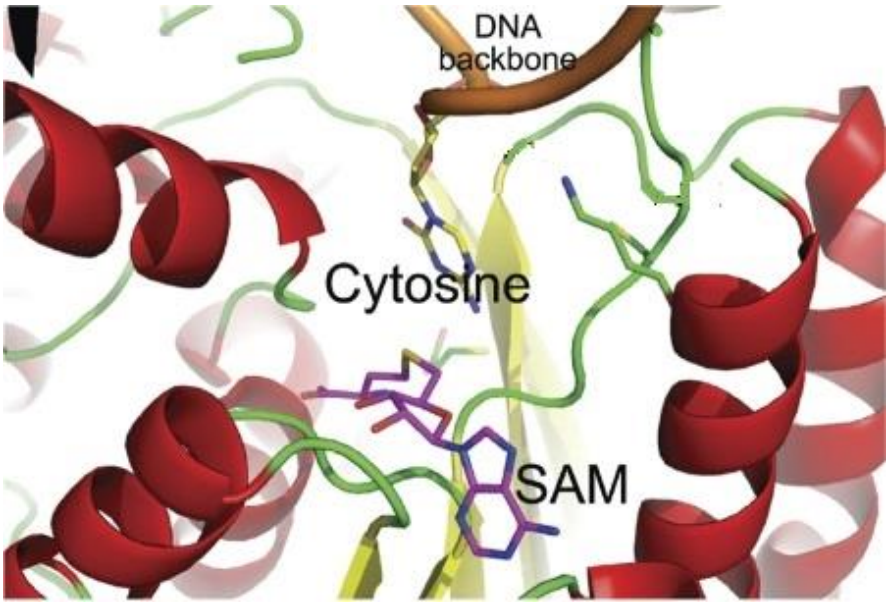


Figure 5. Representation of SAM-binding domain within catalytic subunit of DNMT1. The catalytic pocket of DNMT1 showing cytosine nucleobase turned outwards, precisely facing the site where the SAM is bound. Soon after methyl transfer, both SAM and 5mC leave the binding pocket. SAM is replaced by SAH which is released from the binding site soon after its formation allowing a new catalytic cycle. Protein is shown in the ribbon model. SAM and cytosine are shown in Licorice. (Source: Dhe-Paganon S, Syeda F, Park L. DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. Int J Biochem Mol Biol. 2011 Jan 30 ;2(1):58–66. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3180029/. (4))

DNMT1 plays an important role in establishing and maintaining epigenome features during development (imprinting and X-chromosome inactivation) and after birth. Any non-functional enzyme, whether resulting from dysregulation,

deficiency, or mutation of DNMT1 and DNMT3A/B, can cause developmental disorders such as Prader-Willi, Angelman, or Beckwith-Wiedemann syndrome. A large number of illnesses are mostly pronounced in adults, including psychiatric disorders, schizophrenia, obesity, and cancer (1). The increasing incidence of severe diseases in the population has led to great interest in the development of epigenetic drugs among which are DNMT1 inhibitors.

Inhibitors of DNMT1 as potential drug candidates

Dysregulation of epigenetic processes is the basis of various health disorders. Control of enzymes involved in the methylation machinery allows control of epigenetic events. Inhibition of methylation at the DNA level is of increasing interest. Research on DNMT1 inhibitors has been going on for the last 20-30 years. Its development precedes the study of the human epigenome and cell reprogramming and creation of induced pluripotent stem cells. DNMT1 inhibitors serve as anti-tumor therapy as well as therapy of viral, immunological, and neurological diseases (16).

Inhibitors of DNA methylation are divided into nucleoside analogs and non-nucleoside DNMT inhibitors. The first are nucleoside derivatives that most commonly mimic the cytidine, the nucleoside of cytosine. Nucleoside analogs are covalently incorporated into DNA and block the activity of DNMT when the cytosine analog is methylated. Azacitidine (5-azacytidine) and decitabine (5-aza-2-deoxycytidine) are the earliest discovered cytidine analogs with demethylating potential. Both drugs are approved by Food and Drug Administration (FDA) as a therapy for myelodysplastic syndrome (MDS) conditions (17,18). According to low specificity, their activity may cause severe unwanted toxic outcomes. Some of these are disruption of ribosome synthesis followed by cell death, neutropenia, and carcinogenic effects. In contrast, non-nucleoside inhibitors are considered to bind specifically to DNMT

enzymes thus blocking their activity directly. Thereby, these drugs should exhibit lower toxicity (17).

Over the past 30 years, more than 20 different DNMT1 inhibitors have been described. Most of them are used in cancer therapy and do not show the specificity of binding to a particular type of enzyme. In this way, it is possible to regulate the total DNA methylation at the genome level, but not the methylation of individual genes. The lack of specificity lies in the characteristics of the drugs. More precisely, inhibition of DNA methylation is not usually listed as the first indication which results in a higher risk of side effects. What is more, no compound with a half-maximal inhibitory concentration (IC50), and a measure of inhibitory effect, less than 1 μM has been identified to date (16). Drug repurposing and modifications of existing inhibitors are the main directions in the development of DNA methylation inhibitors. Modification of only certain parts of the molecules avoids de novo synthesis of the compounds, thus ensuring a reasonable price (Figure 6).

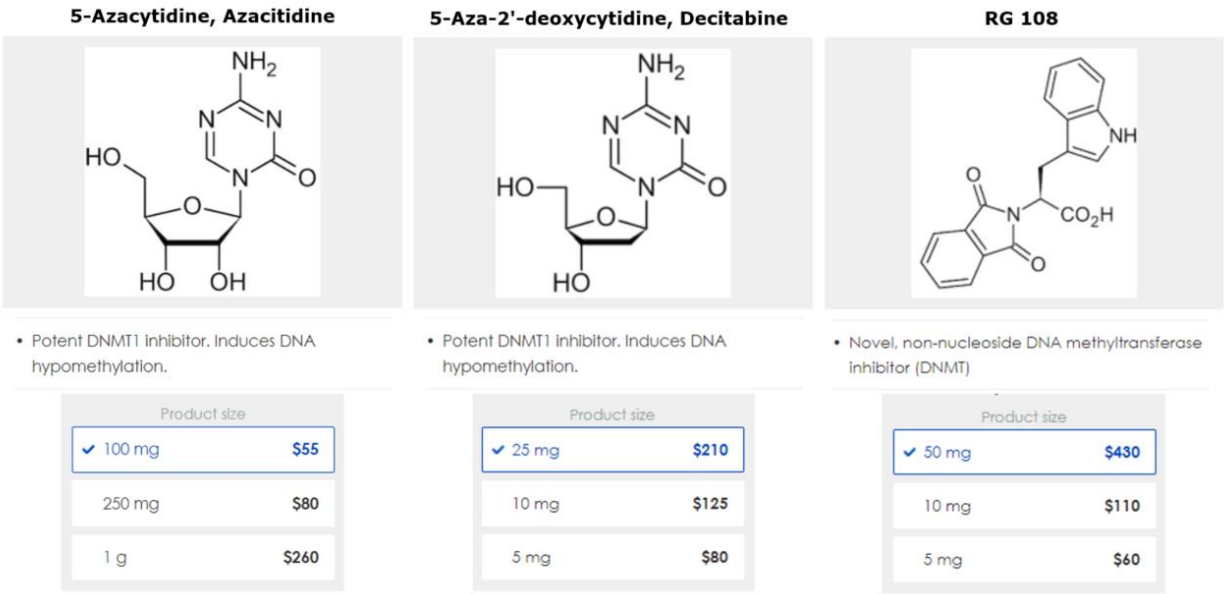


Figure 6. Chemical structures and quantity-dependent prices of DNMT1 inhibitors. Azacitidine and Decitabine are cytidine analogs that serve as approved potent DNMT1 inhibitors. RG 108 is a novel non-nucleoside DNMT inhibitor that directly blocks the DNMT1

active site. It inhibits DNA methylation in human cancer cells and reactivates tumor suppressor genes. The structures, descriptions of compounds, and prices per mg/g were obtained from the Abcam website (19).

In silico drug design: Structure-Based Drug Design (SBDD)

Solving DNMT domain structures and their functions, as well as computer development, has contributed to the research and discovery of novel potential DNMT1 inhibitors.

Structure-based drug design (SBDD) is a technique mostly used to design inhibitor molecules according to the known three-dimensional target structure. SBDD consists of few steps: knowledge of target structure, virtual screening (VS), binding site identification, and molecular docking followed by molecular dynamics (MD) simulations. Structural data of protein targets originates either from experimentally data obtained through Nuclear Magnetic Resonance (NMR) spectroscopy, X-ray crystallography, or from computational homology modeling (20). Protein Data Bank (PDB) (21) is a repository of biological macromolecular structures. Most PDB structure files do not contain information about protonation and tautomeric states, formal atomic charge, bond orders, or hydrogen atoms. Considering that, protein preparation is a necessity. PDB structure file may contain water molecules, cofactors, ligands, and metal ions, as well as several protein subunits. In this case, all these molecules also must satisfy a corresponding 3D geometry, bond orders, protonation, and tautomeric states before VS.

VS is a part of SBDD. It refers to computational screening of drug-like compounds contained in commercially available libraries, against targets of known structure. Docking programs, that explore the ligand conformational space, are used for this purpose. Compound libraries contain drug-like small molecules with desirable characteristics and the absence of toxic and

undesirable moieties. Drug-likeness is most often determined by following the Lipinski rule of five. The rule describes which molecular parameters the molecule would have to fit in to be drug-like. These include molecular weight lower (MW) than 500, lipophilicity (log P) lower than 5, less than five hydrogen bond donors, and less than 10 hydrogen bond acceptors.

Docking performance enables a detailed understanding of interactions between small molecules and proteins. That is the basis of drug design strategy. Molecular docking involves the creation of different docked positions with corresponding score values that may serve to predict the energy profile and stability of the protein-ligand complex. The result of molecular docking should be a stable docked complex - a system with minimized binding free energy (ΔG_{bind}). Compounds that are predicted to bind well *in silico* become hits. A good hit compound should satisfy properties given by the Lipinski rule, toxicological profile, and shape-complementarity with protein. What is more, compounds should exhibit high ligand efficiency and at least nanomolar IC50 values. Ligand efficiency is defined as a measure derived from scaling affinities by molecular size. By further hit optimization a promising drug candidate, lead, may arise. However, preclinical and clinical trials should be performed to determine the properties of absorption, distribution, metabolism, excretion, and toxicity (ADMET) of lead compounds (20).

The role of Molecular Dynamics

The molecular dynamics (MD) simulations opened a new capture in drug design enabling a more detailed examination of highly complex biological systems. Grouping of MD simulations refers to the division depending on the size of the biological system and the time scale of the dynamics. All-atom MD simulations (AAMD) are classic system simulations. There are two cases in which MD simulations are most used. Firstly, they are applied to observe the

conformational changes during the formation of biological complexes. MD simulations are also useful in observing the way proteins interact with other proteins, ligands, and DNA or RNA. On the time scale, dynamics are calculated in nanoseconds (ns) and microseconds (μs). Coarse-grained (CG) simulations denote a simple representation of the dynamics of large biological systems with a time scale in milliseconds (ms) (22). MD simulations mimic the motion of macromolecules under physiological conditions. Therefore, the most performed analysis are measurement of macromolecular motion and monitoring of intermolecular interactions. Analysis of structural fluctuation implies Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuations (RMSF) (23).

The RMSD is a measure of the average distance between atoms of the structure at a particular point in the simulation and the reference structure. The reference is usually the first simulation frame. It is defined as:

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

where the coordinates of N number of atoms in a structure at a particular point of simulation t ($x_i(t)$) and the coordinates of the reference structure x_i^{ref} show the mean square displacement for each atom (23).

The RMSF measures the deviation between the position of a single particle of biomolecule and the reference position, in a time-specific manner. RMSF is statistically defined as:

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

where T is the time in which RMSF is measured, r_i is the position of individual residue, r'_i presents atomic coordinates in residue after superimposition on the reference, and t_{ref} is the reference time.

In drug design, the RMSD provides information about stability of protein-bound drugs. RMSF values are used to show the mobility of each amino acid in a protein target. Comparing RMSD values of distinct tested compounds and reference inhibitor indicates better or worse binding relative to known inhibitor. The RMSD may also serve as a technique for the analysis of structural similarity. RMSF analysis provides insight into the protein parts affected by the bound drug. It also plays a role in the discovery of new potential binding sites, focusing on regions with high fluctuation rates whose function may be inhibited by the drug.

DNA-substrate dependent DNMT1 structures

Multiple crystal structures of human and mouse DNMT1 enzymes were solved and described until today. Among these are structures with PDB codes 4WXX, 3PTA, and 4DA4.

4WXX represents the crystal structure of human DNMT1 without DNA substrate (Figure 8A), showing residues 351-1600. In other words, 4WXX is comprised of RFTS, CXXC, BAH, and MTase domains, while the DMAP domain is missing (14). PDB code 3PTA stands for the structure of human DNMT1 (residues 646-1600) complexed with DNA whose cytosine base is not inverted yet to face SAM (Figure 8B) (12). 4DA4 is a structure of mouse DNMT1 (residues 731-1602) bound to hemimethylated CpG DNA (Figure 8C). Relative to the 3PTA structure, the difference is the inversion of the cytosine to the catalytic site (24). DMAP and RFTS segments are missing in the protein sequence of 3PTA. 4DA4 additionally lacks the CXXC domain and contains engineered residues 5CM and C49 which denote modified cytosine bases (12,24).

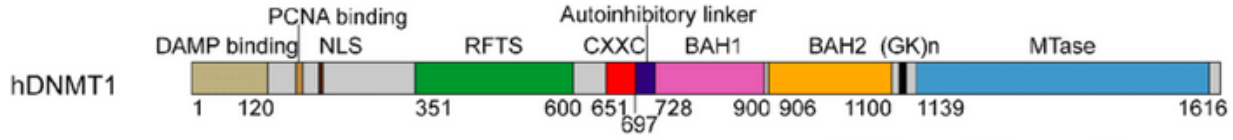


Figure 7. Domains of human DNMT1 enzyme. The enzyme is shown as a line. Domains differ in color. Numbers denote amino acid residues occupied by a particular domain. (Image is partly taken from: Ren W, Gao L, Song J. Structural basis of DNMT1 and DNMT3A-mediated DNA methylation. Genes. 2018 Dec 11;9:620. DOI:10.3390/genes9120620. *(25)*)

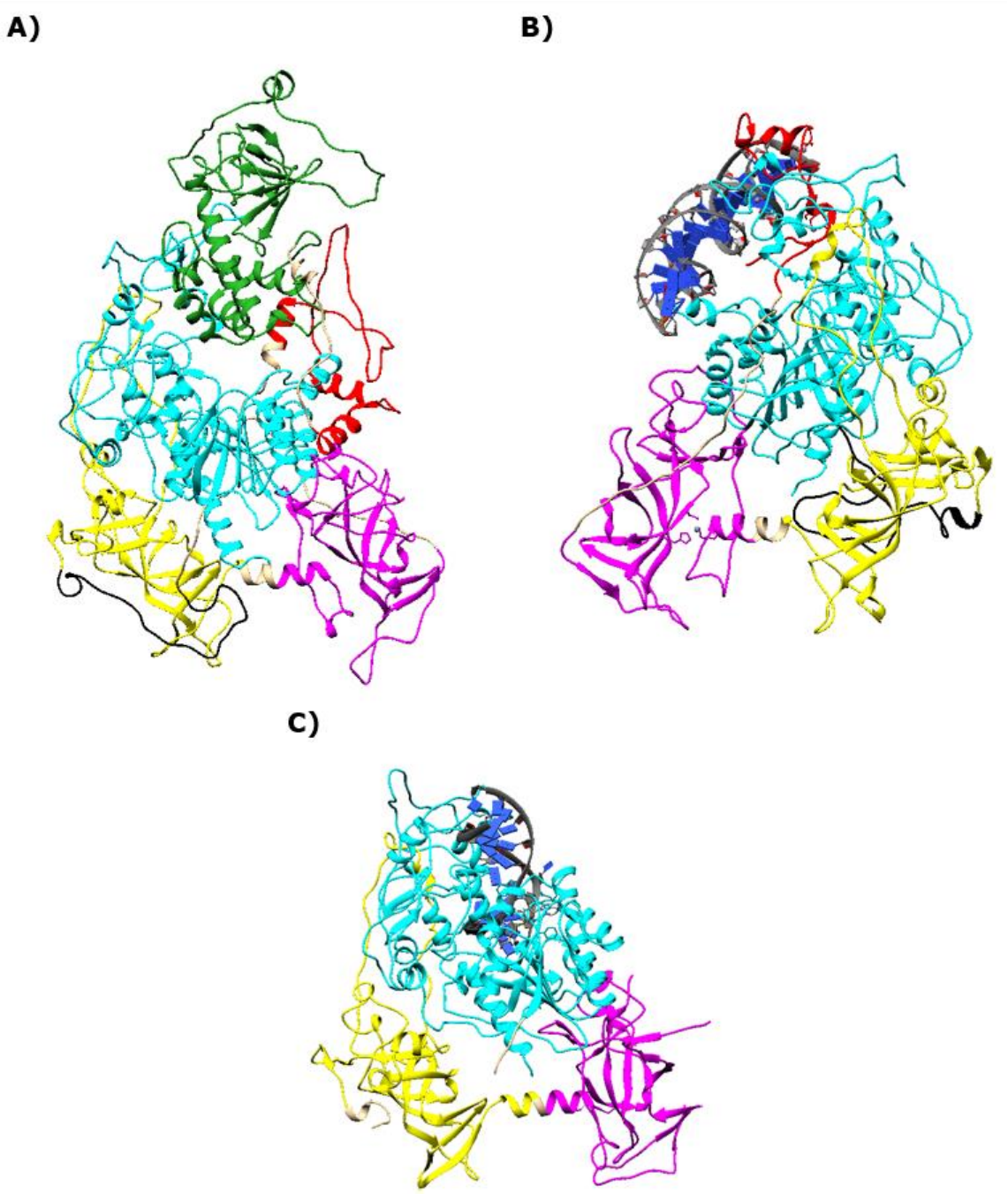


Figure 8. Display of crystal structure with colored domains for A) 4WXX, B) 3PTA and C) 4DA4. The images were created using UCSF Chimera *(26)* tools. Structures of DNMT1 were obtained from PDB. *(12,14,24)* Domains were colored according to Figure 7: RFTS domain in forest green, CXXC domain in red, BAH1 and BAH2 in magenta and yellow, and MTase domain in cyan. DNA duplex is colored in gray. Gly-Lys linker peptide is black.

Sequence similarity can be shown by structural alignment or using a phylogenetic tree. The latter is a branching diagram that depicts the evolutionary connection between various species, based on different parameters. There is noticeable conservation of residues within catalytic domains since they perform biochemical processes crucial for cell survival. Regions that are not similar in sequence are mostly located at the edges of structures. They are thought to be involved in protein-protein interactions that is distinct for each organism. Considering the significance of DNA methylation for the proper cell function, active sites of enzymes involved in this process have been preserved among different organisms through evolution (Figure 9). Human and murine DNMT1 share 80-90% (83.39%) of the amino acid sequence. Bacterial modification methylase M.HhaI, isolated from *Haemophilus parahaemolyticus*, consists only of the MTase domain. Thereby, the regulatory subunit and its domains represent the main structural characteristic to distinct human and mouse DNMT1 enzyme from its bacterial ancestor (Figure 9). What is more, the development of these domains through evolution has contributed to the regulation of DNMT1 and its recruitment to the site of action (14).

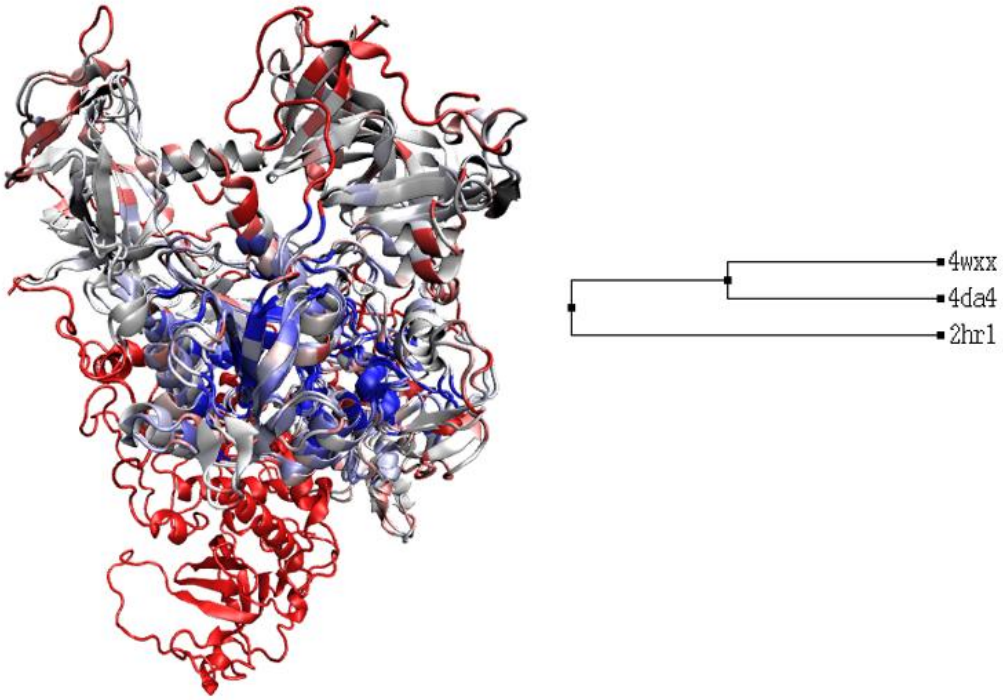


Figure 9. Evolutionary conservation of the catalytic domains of the bacterial, mouse, and human DNA methyltransferase DNMT1. The right side shows the image of aligned bacterial (2HR1), mouse (4DA4), and human (4WXX) DNMT1 structures (chain A) obtained from PDB *(12,14,24)* and the left side shows a phylogenetic tree, both generated using Multiseq analysis in VMD *(27)*. Structures are colored by sequence identity and similarity; blue is for similar and conserved protein regions, red is for distinct amino acid residues or parts where some residues are missing in one structure, compared to the other two proteins. The sequence tree is generated using percent identity. It shows high sequence similarity for 4WXX and 4DA4, while the amino acid sequence of 2HR1 differs due to the lack of N-terminal region.

Aim of the study

Previous studies report low specificity and bioavailability, and high toxicity of existing DNMT1 inhibitors. Various compounds were synthesized using *in silico* methods until now. In most cases, the high costs of drug synthesis for *in vitro* testing were an obstacle to further steps of the analysis. Because of that the development of inhibitors is unsuccessful or time-consuming (18). There are two main directions of this study: development of novel potential DNMT1 inhibitors using AAMD of protein-inhibitor complexes, and CG analysis for identification of protein functional domains.

The initial purpose of our study was to find commercially available and reasonably priced drug candidates, using Maybridge and ELITE databases. We aimed to find any compound, which may not be very potent, but exhibits drug-like properties and interacts with DNMT1. This strategy is economically acceptable for proving the concept of SBDD. It is not successful if we refer to the successful binding of the drug and its ability to inhibit the enzyme.

The results obtained during the development of candidate compounds directed the research towards CG analysis. Our goal was to give insight into DNMT1 parts with high structural fluctuations. By that, we laid the foundation for further analysis and possible discoveries of new DNMT1 binding sites.

Materials and Methods

The crystal structures of incompletely sequenced human DNMT1 (PDB codes: 4WXX and 3PTA) and murine DNMT1 (PDB code: 4DA4), in a complex with DNA and/or SAH, were obtained from PDB (12,14,24). Atomic coordinates of a protein, DNA, and ligand were visualized using UCSF Chimera. The missing amino residues were supplemented using the Modeller interface in UCSF Chimera (26). In the case of 4WXX and 3PTA, chain A was used for analysis. For 4DA4, chains A, C, and D were analyzed. The structures of the tested compounds were taken from Maybridge and ELITE databases and visualized in ChemAxon Marvin (28). A pKa based protonation state check was performed in MarvinSketch, at pH=7,2. Physicochemical descriptors, ADME parameters, and drug-likeness for potential inhibitors were computed and predicted using a website SwissADME (29).

Molecular docking of Maybridge1 and ELITE12 was carried out using Autodock Vina through the UCSF Chimera interface (30). DockPrep tool in UCSF Chimera (26) was used to add hydrogens, assign charges and specify net charges for ligand molecules. The compounds were docked at the SAH binding site. Water and other ligands were excluded to facilitate the observation of protein-compound interactions. For further analysis, the conformation of a compound with the lowest score value was selected.

AAMD and CG simulations were carried out using the GROMACS suite (31), version 2020.4. The obtained results are visualized in VMD from where the coordinates of the system are saved in pdb and dcd format. Data analysis was performed in VMD (27) and RStudio using the Bio3D package (32). RMSD and RMSF charts were overlapped using Microsoft Excel tools. The number of hydrogen bonds was analyzed in VMD (27). Thereafter, 2-D schematic representations of protein-ligand interactions were generated in LIGPLOT (33) using saved .pdb files of simulation frames as input. Images were created

using the UCSF Chimera tools. For MD analysis, all protein-SAH simulations were calculated at 150 million steps. The calculation for protein-inhibitor simulations was performed in 50 million steps. Input files for GROMACS were generated in CHARMM-GUI (34). Ligand parameters and the following structure files were created in the Ligand Reader & Modeler module based on protein-SAH structures and protein-docked compound structures. Ligand topology and parameter files were loaded in the Solution builder module where models of protein-DNA-ligand complexes in an aqueous solution were generated. A 0.15M potassium chloride solution (KCl) was used in all parameterizations. The number of ions was calculated depending on the water box size. A rectangular water box was set up in all simulations using the TIP3 model for water molecules. For 4WXX, 3PTA, and 4DA4 in a complex with SAH, the distance from the edge of the protein to the edge of the box was 10 Å. For proteins in a complex with inhibitors, system size was determined with coordinates X:180, Y:130, and Z:130. The system was simulated at a temperature of 303.15K and a pressure of 1 bar. In the 4DA4 structure, that contains DNA, engineered nucleotide bases 5CM and C49 were replaced by cytosine (CYT). For CG analysis, 3PTA and 4DA4 simulations were calculated at one billion steps. 4WXX simulation was calculated at 250 million steps due to unstable simulation. Since CHARMM-GUI does not support DNA, the nucleic acid was excluded from the analysis. PDB entities were modified to a format readable in CHARMM through the CHARMM-GUI PDB Reader module. The resulting output was loaded into Martini Maker - Solution builder. The water box was fitted to protein size with an edge distance of 70 Å. A 0.15 M sodium chloride (NaCl) solution was used. The temperature and pressure of a system were set up at 303,15 K and 1 bar.

All calculations were performed on Bullx DLC720 supercomputer Bura using 10-20 nodes for each calculation. Binding interactions were performed over a time range of 300 ns for protein-SAH simulations and 100 ns for protein-

compound interactions, with a timestep of 2 femtoseconds (fs). The duration of CG simulations was 5 μs for 4WXX, and 20 μs for 3PTA and 4DA4. As a result, 100-250 frames of one simulation were created. For ease of analysis, only protein and inhibitor are shown in the results.

Results

Insight into the SAH binding site

SAH is a potent inhibitor of DNMT1. Its structure serves for the discovery and synthesis of new, more potent inhibitors for many years. Another way for the development of DNMT1 inhibitors is the use of the binding pocket where the SAM/SAH riboswitch occurs (18). We noticed that there are several different forms of DNMT1 depending on the stage of the DNA methylation process. Therefore, we studied the binding of potential inhibitors to all three structures (4WXX, 3PTA, and 4DA4). What makes a molecule a good drug candidate are stable interactions with all three forms of protein.

We analyzed the binding site of SAH on 4WXX, 3PTA, and 4DA4 to compare the binding of drug candidates with SAH binding below. The cofactor SAH binds to all three forms of DNMT1. For 4WXX and 3PTA, the binding domain includes amino acid residues from position 1145 to position 1247, also including Val1580. For 4DA4, it is shifted by 3 amino acids forward. Since SAH is a relatively hydrophilic molecule (35), the binding pocket is hydrophobic (Figure 10).

Furthermore, we analyzed protein-SAH interactions. SAH forms hydrogen bonds and hydrophobic interactions with several amino acids within the active site (Figure 11). These interactions break and arise during the protein-ligand motion. The number of hydrogen bonds formed by SAH within the binding pocket of the 4WXX, 3PTA, and 4DA4 in the first frame of simulation is seven, six, and seven, respectively. Amino acids involved in the formation of hydrogen bonds are alanine (Ala), asparagine (Asn), aspartic acid (Asp), cysteine (Cys), glutamic acid (Glu), glycine (Gly), leucine (Leu), methionine (Met), proline (Pro), serine (Ser), and valine (Val). All hydrogen bonds are characterized by a donor-acceptor distance of 2.50-3.14 Å. Hydrophobic interactions include Ser1146, Met1169, Glu1189, Pro1225, Leu1247, Asn1578, and Ala1579.

Additionally, we confirmed that phenylalanine (Phe) is located within the binding site and may create π-π stacking interaction with the aminopurine ring of SAH through its aromatic moiety.

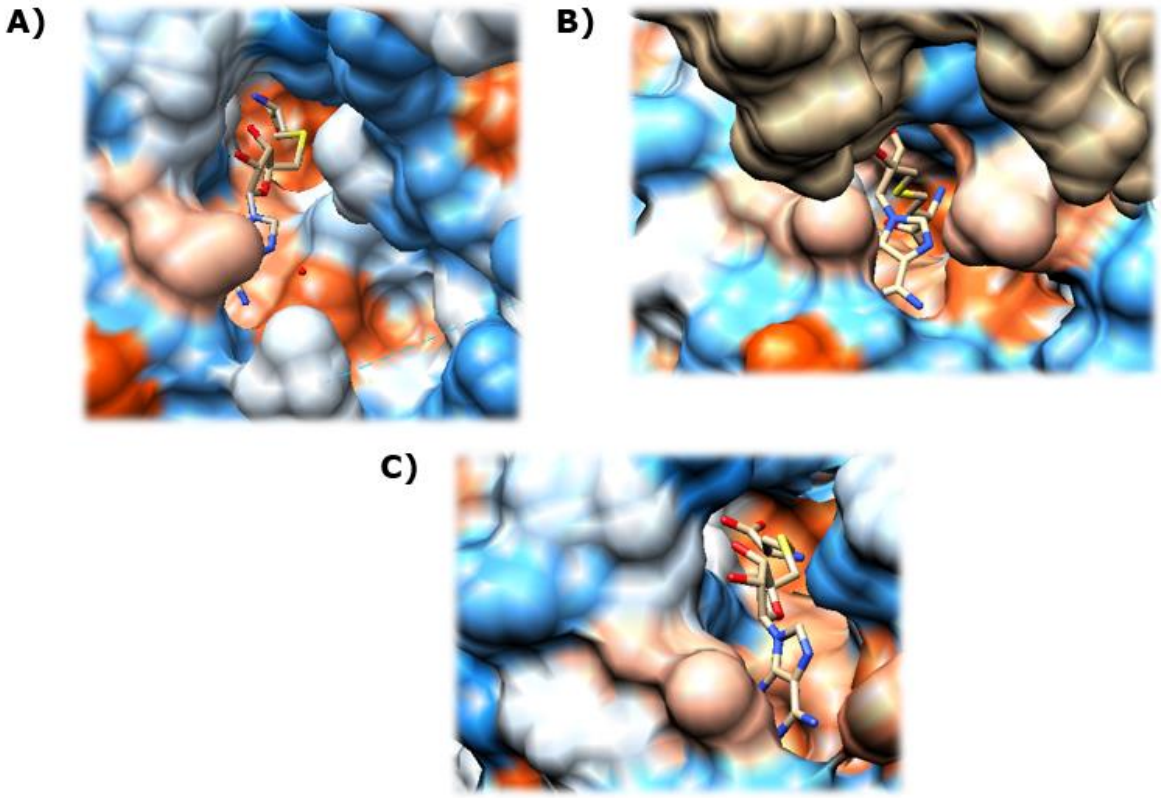


Figure 10. The binding pocket of DNMT1 shows hydrophobic properties. Images were created in UCSF Chimera *(26)*. Protein is depicted by hydrophobicity surface (the white area is for non-polar, green for polar, red for negatively charged, and blue for positively charged amino acids). The SAH structure is depicted by the atom type (sulfur in yellow, nitrogen in blue, oxygen in red). A) Binding pocket of 4WXX *(14)*. B) Binding pocket of 3PTA *(12)*. C) Binding pocket of 4DA4 *(24)*.

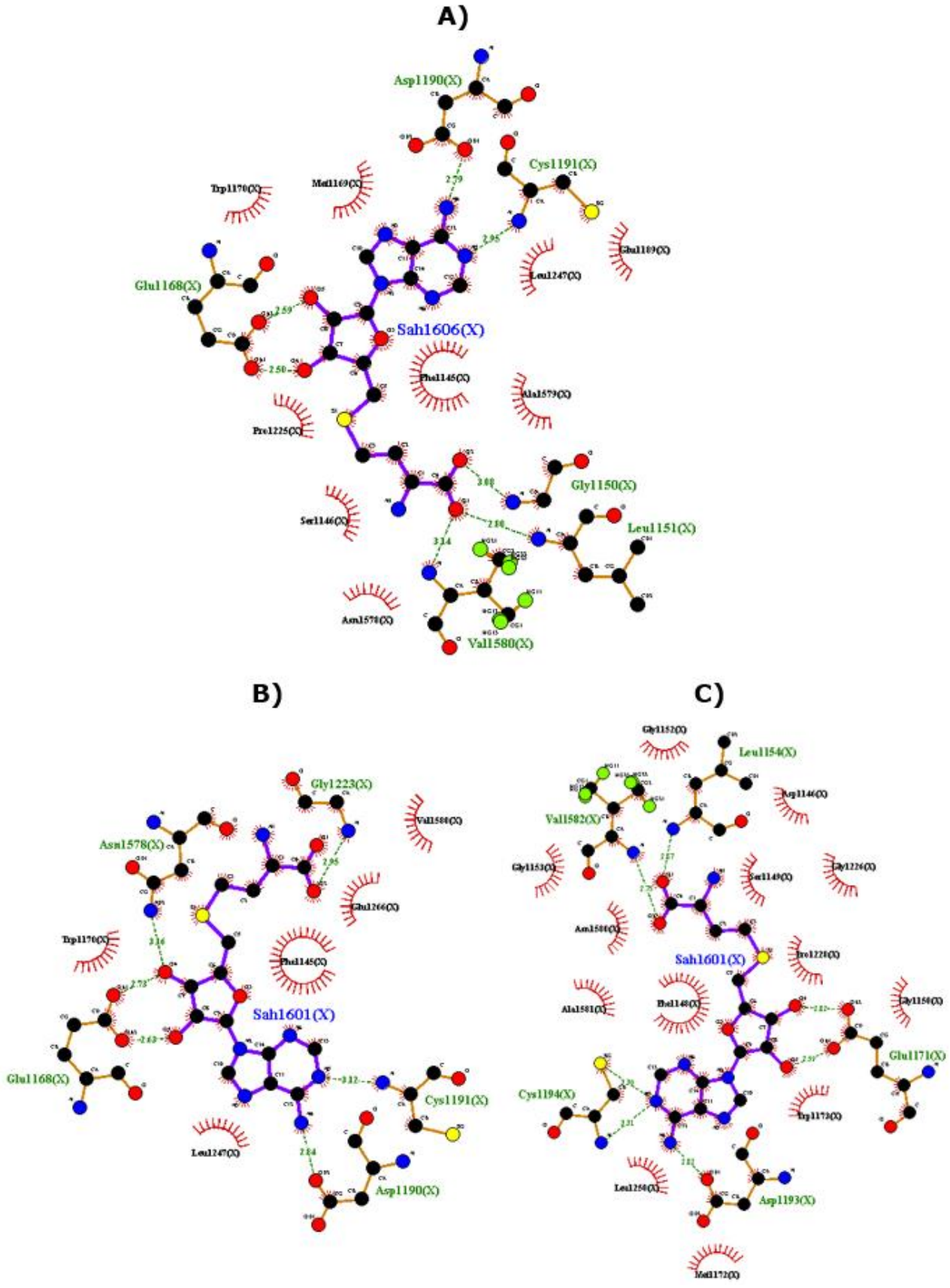

A)
Asp1190(X)
Cys1191(X)
Trp1170(X)
Met1169(X)
Glu1168(X)
Leu1247(X)
Sah1606(X)
Phe1145(X)
Ala1579(X)
Pro1225(X)
Gly1150(X)
Ser1146(X)
Leu1151(X)
Asn1578(X)
Val1580(X)
B)
Gly1223(X)
Val1580(X)
Asn1578(X)
Glu1266(X)
Trp1170(X)
Phe1145(X)
Sah1601(X)
Glu1168(X)
Cys1191(X)
Leu1247(X)
Asp1190(X)
C)
Gly1152(X)
Leu1154(X)
Val1582(X)
Asp1146(X)
Gly1153(X)
Ser1149(X)
Asn1581(X)
Sah1601(X)
Ala1581(X)
Phe1148(X)
Glu1171(X)
Cys1194(X)
Leu1254(X)
Asp1193(X)

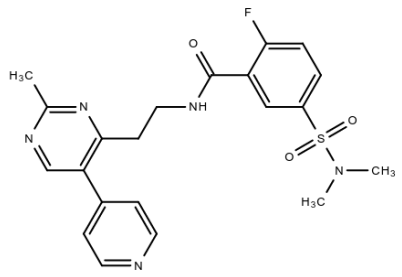
Figure 11. 2-D representation of protein-SAH interactions for A) 4WXX, B) 3PTA, and C) 4DA4. 2D graphs were generated in the LIGPLOT program *(33)*. Interactions were captured at the first frame of the simulation. The structure of SAH is colored in purple. Hydrogen bonds are indicated by a green dashed line. The number next to the line stands for the donor-acceptor distance. Amino acid residues involved in hydrophobic interactions are visible as curved red lines.

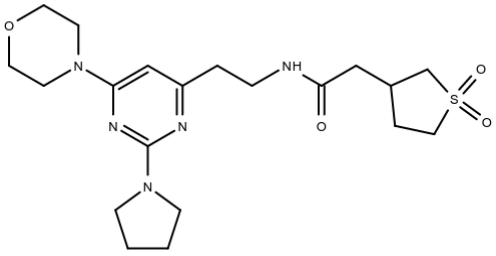
Binding analysis of selected compounds to SAH binding site

Based on the structure of the SAH binding pocket, we took candidate compound structures from the Maybridge and ELITE databases. We performed molecular docking using crystal protein structures of 4WXX, 3PTA, and 4DA4. The structures of the compounds and their chemical names are shown in Table 1. ELITE12 and Maybridge1 bind to the same binding site as the already known SAH inhibitor (Figure 12).

Next, we calculated pharmacokinetics, drug-likeness, medicinal chemistry, and physicochemical properties for each inhibitor and reference molecule. For that purpose, we used the website SwissADME. Absorption, distribution, metabolism, and excretion (ADME) properties are shown in Table 2.

Table 1. Structures and names of the tested inhibitors. The structures were drawn using the tools of the ChemSpider *(36)* website. The chemical name according to IUPAC was found in the ChemSpider database of chemical structures. We used an abbreviated name in the paper for easier reference.

	IUPAC name:	5-(Dimethylsulfamoyl)-2-fluoro-N-{2-[2-methyl-5-(4-pyridinyl)-4-pyrimidinyl]ethyl}benzamide
	Name used in a study:	ELITE12

	IUPAC name:	2-(1,1-Dioxidotetrahydro-3-thiophenyl)-N-{2-[6-(4-morpholinyl)-2-(1-pyrrolidinyl)-4-pyrimidinyl]ethyl}acetamide
	Name used in a study:	Maybridge1, MYB1

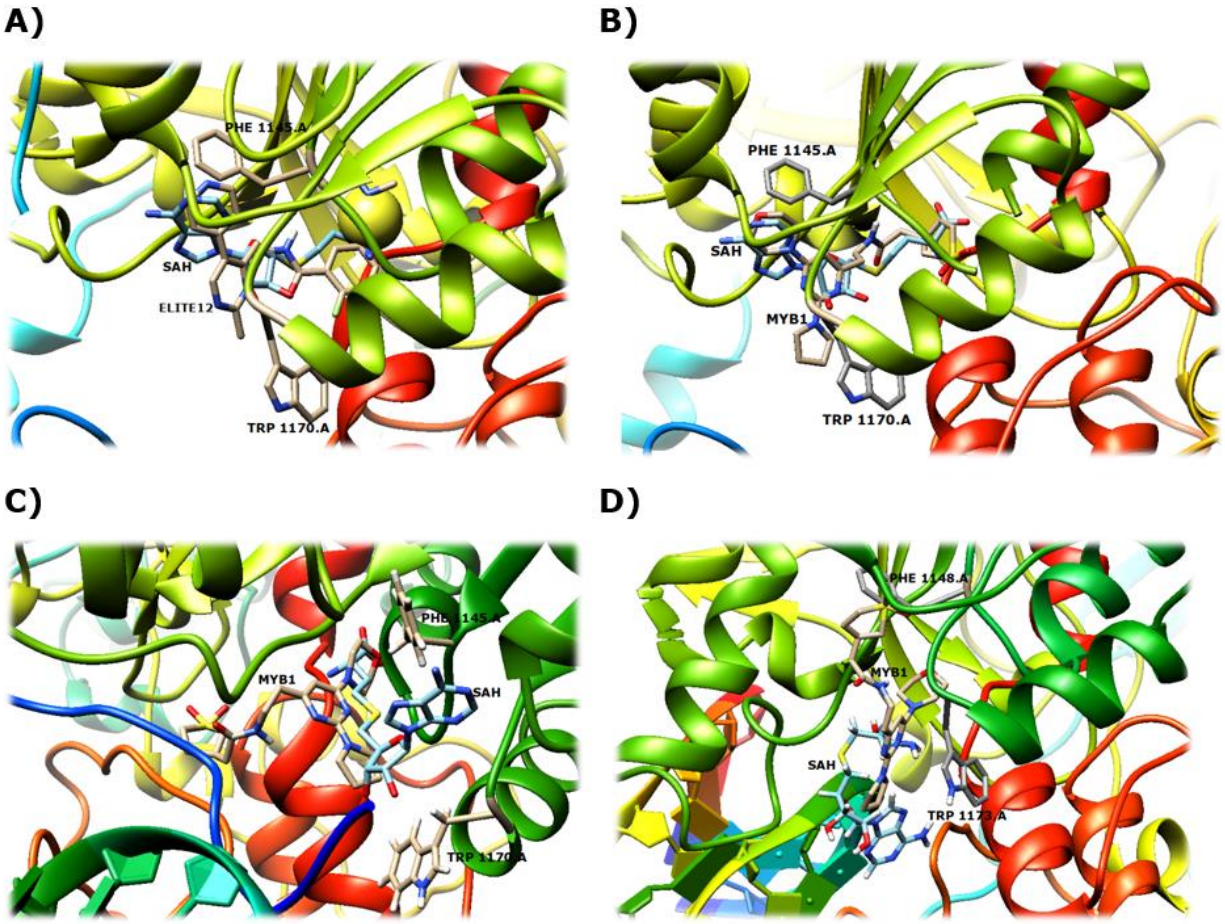
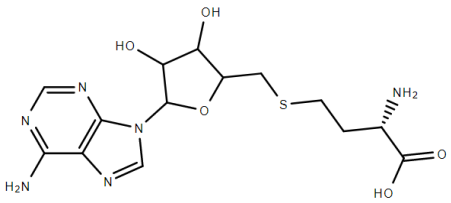
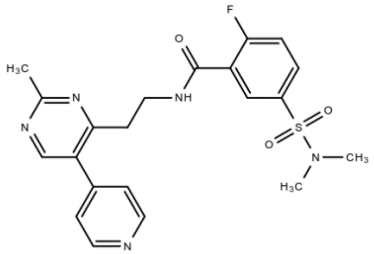
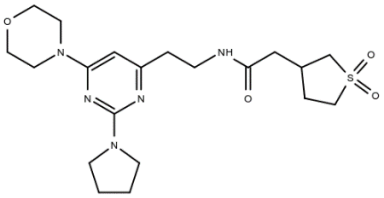
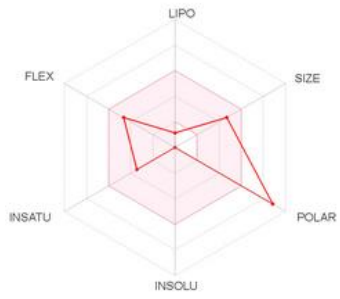
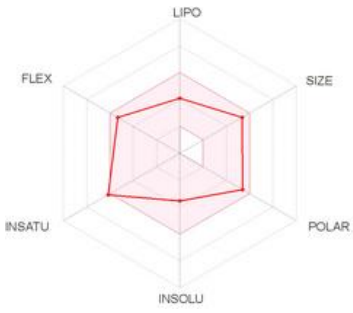
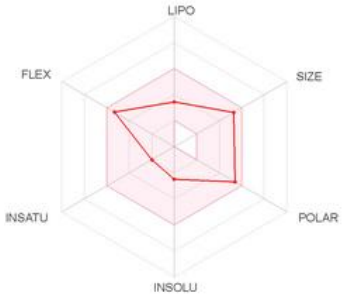


Figure 12. ELITE12 and Maybridge1 fit into the same binding pocket as SAH. The enlarged view of the 4WXX *(14)*, 3PTA *(12)*, and 4DA4 *(24)* binding pocket with bound ligands was generated using UCSF Chimera *(26)*. Protein is depicted using the ribbon model and

colored in Rainbow. Atoms and bonds of SAH, ELITE12, and MYB1 structures are shown as sticks. The hydrocarbon skeleton is colored in cyan for the reference structure (SAH), and in beige for ELITE12 and MYB1 structures. Phe1145 and Trp1170 (Phe1148 and Trp1173 for 4DA4) are labeled as amino acids that can participate in the formation of π-π stacking interactions. A) ELITE12 and SAH bound to 4WXX. B) MYB1 and SAH bound to 4WXX. C) MYB1 and SAH bound to 3PTA. D) MYB1 and SAH bound to 4DA4.

Table 2. ADME properties for SAH and drug candidates ELITE12 and Maybridge1. The table presents structures of SAH, ELITE12, and MYB1, drawn using the ChemSpider *(36)* website, and physicochemical properties for oral bioavailability obtained from the SwissADME *(29)* website. Physicochemical descriptors are fitted into a bioavailability radar: LIPO for lipophilicity, FLEX for flexibility, INSATU for insaturation, INSOLU for insolubility, POLAR for polarity, and SIZE.

SAH	ELITE12	MYB1
		
		

Molecular dynamics simulations point to binding modes of ELITE12 and Maybridge1 inhibitors

To show interactions between different forms of DNMT1 and inhibitor compounds ELITE12 and Maybridge1, we calculated AAMD showing the protein-ligand system in an aqueous environment. We obtained RMSD values for each ligand, including the reference inhibitor SAH.

Binding analysis of MYB1 and ELITE12 compounds at 4WXX relative to the reference SAH

The comparison of SAH, ELITE12, and Maybridge1 RMSD values bound to 4WXX is shown in Figure 13. RMSD values indicate relatively equal protein-SAH movement. The amplitude of the variation of SAH motion within the binding site is approximately 2.5 Å. The values for Maybridge1 are similar to those of SAH. The values for ELITE12 are significantly increased.

After an indication that there is a strong ELITE12 motion within the binding site, we examined the stability of the resulting protein-ligand complex. For that purpose, we used the first and the last simulation frame. We found that the ELITE12 compound did not remain bound within the binding pocket during 100 ns of simulation (Figure 14). Therefore, docking of this compound to forms 3PTA and 4DA4 was not performed below. The binding of the Maybridge1 molecule within the catalytic site is partially maintained over the 100 ns of simulation (Figure 15).

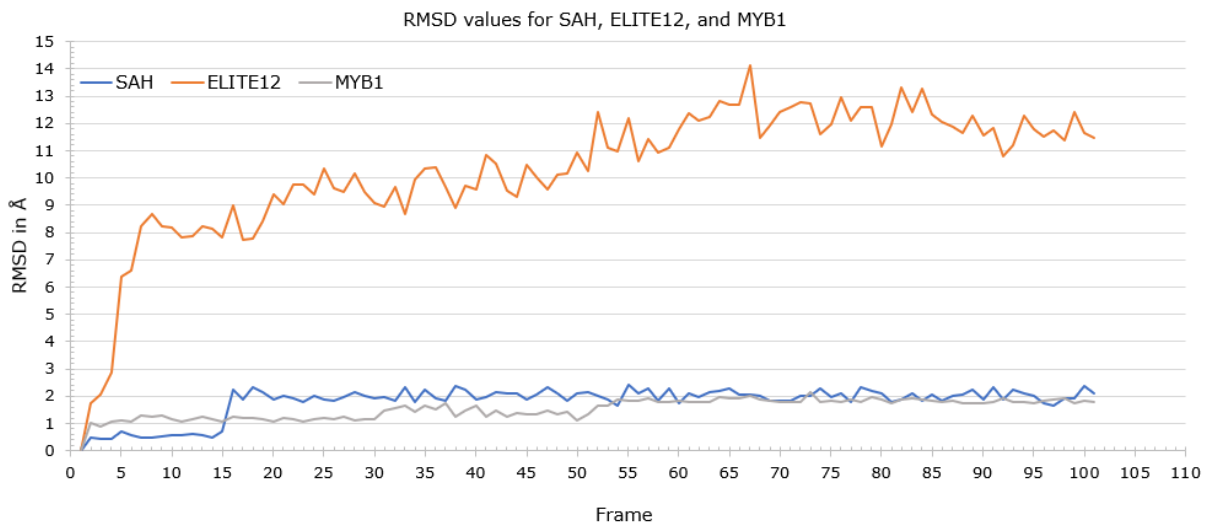


Figure 13. ELITE12 and Maybridge1 show opposite structural fluctuation compared with SAH. The RMSD values were obtained through RMSD Trajectory Tool in VMD *(27)*. The chart was generated using Excel. X-axis stands for the number of frames in simulation. Y-axis presents RMSD values in Å for each ligand: blue line for SAH, gray line for Maybridge1, and orange line for ELITE12.

The 90th frame is a part of the simulation in which the motion of all three ligands is relatively steady. Based on that, we used it to capture interactions using the LIGPLOT program. We compared the number of hydrogen bonds created for 4WXX-ELITE12 and 4WXX-Maybridge1 with the number of bonds created in the case of 4WXX-SAH. Five out of seven hydrogen bonds that SAH forms at the beginning of the simulation are found to be stable in the 90th frame (Figure 16A and 16B). ELITE12 forms a single hydrogen bond within the binding site (Figure 16C). Despite the previously demonstrated retention of one part of the molecule within the 4WXX binding site, Maybridge1 does not interact with protein in the 90th frame (Figure 16D).

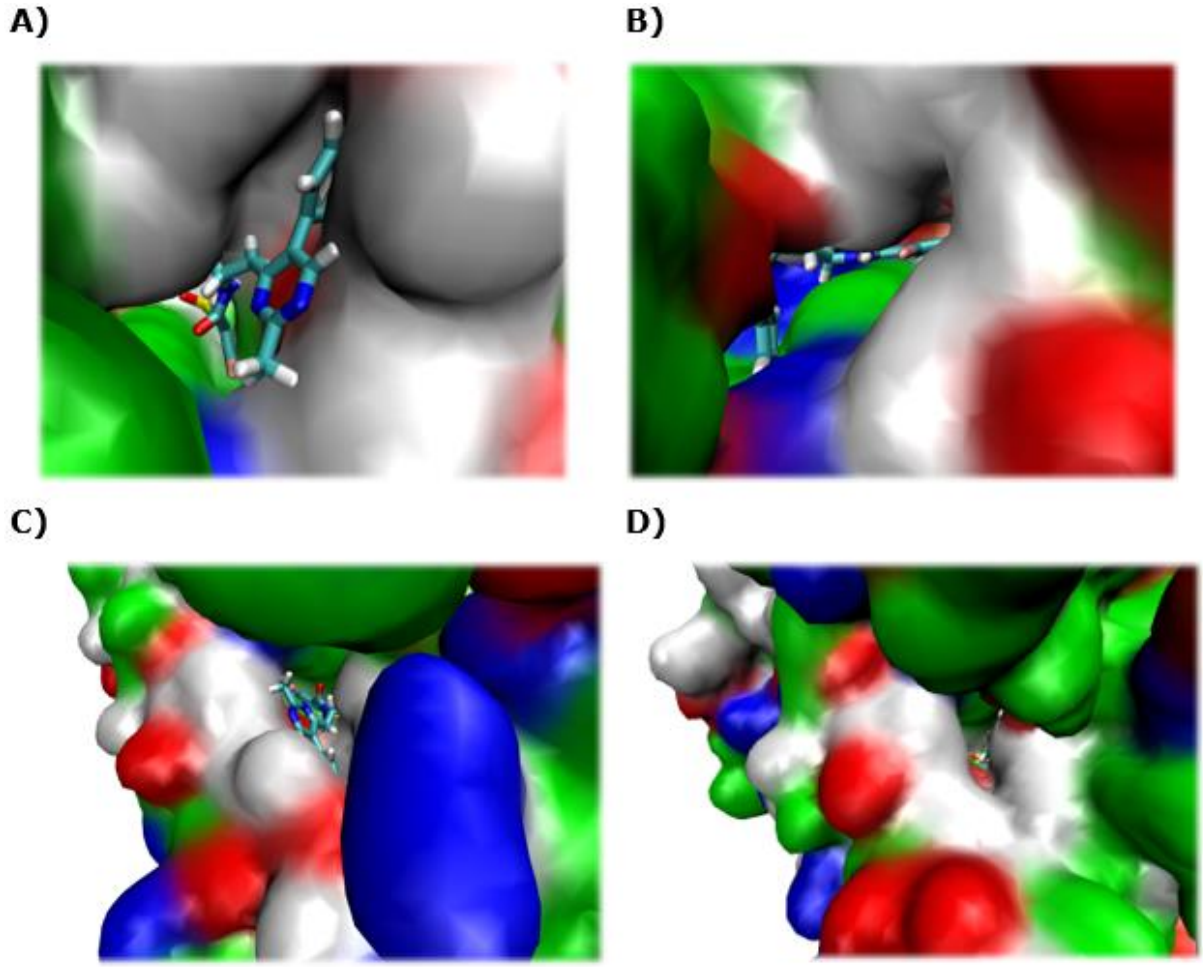


Figure 14. ELITE12 drops out of the binding site during 100 ns simulation. 4WXX-ELITE12 binding mode was analyzed in VMD (27). Display of ELITE12 in the binding pocket of 4WXX was captured in the first (A, C) and the last frame of simulation (B, D) from different angles. The protein surface is shown by atom density (QuickSurf). The ELITE12 structure is shown using the Licorice representation method. The protein surface is colored according to the type of amino acid residues (ResType) - hydrophobic amino acids are colored in white, hydrophilic in green, positively charged amino acids in blue, and negatively charged in red. The ELITE12 structure is colored by atom type (sulfur in yellow, nitrogen in blue, oxygen in red).

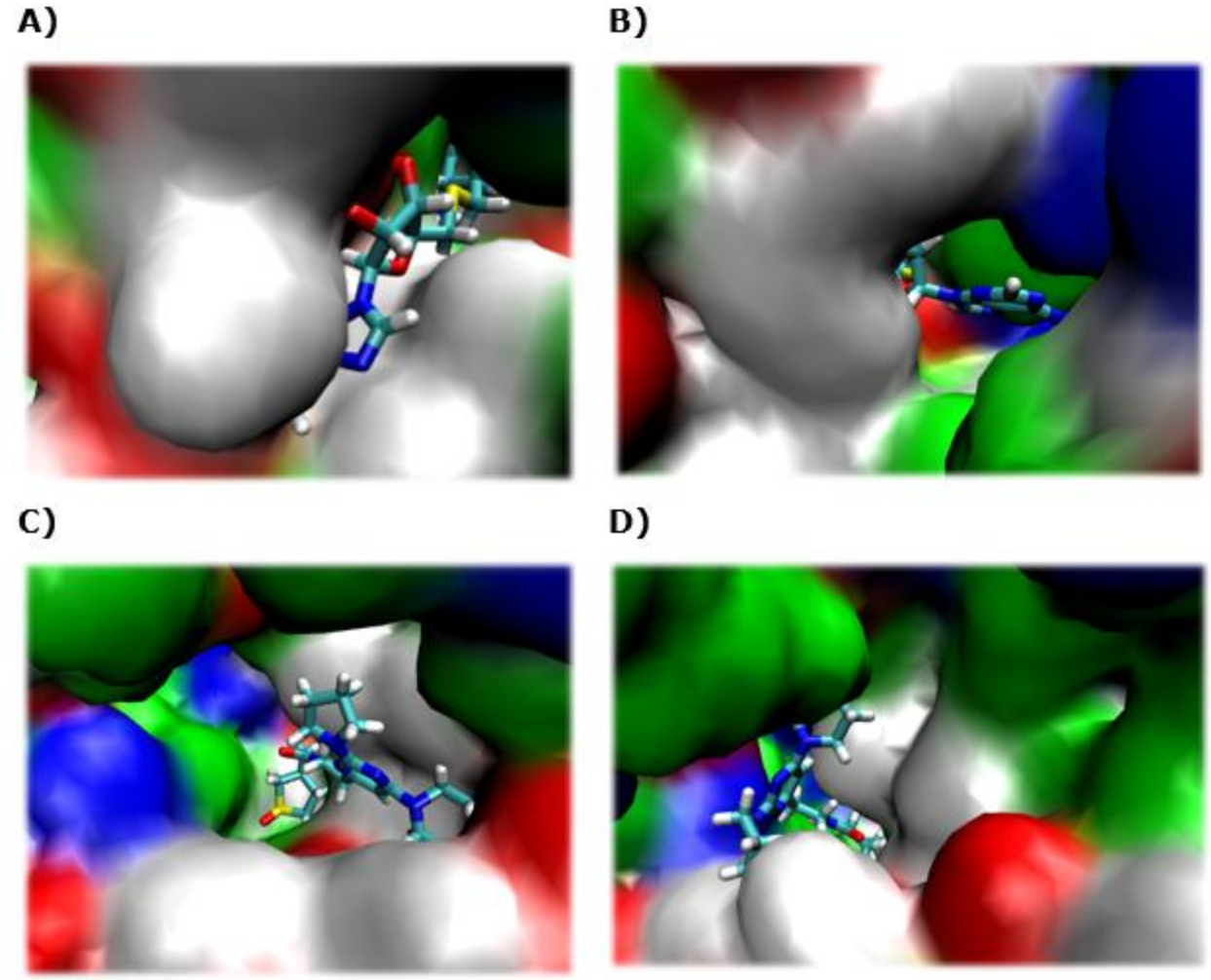


Figure 15. Maybridge1 remains bound to 4WXX during 100 ns simulation. 4WXX-Maybridge1 and 4WXX-SAH binding modes were analyzed in VMD *(27)*. The first and the last frame of the simulation were captured for SAH (A, B) and Maybridge1 (C, D). The protein surface is shown by atom density (QuickSurf) and colored by residue type (ResType). Hydrophobic amino acids are colored in white, hydrophilic in green, positively charged amino acids in blue, and negatively charged in red. The ligand structures are shown using the Licorice representation method and colored by atom type (sulfur in yellow, nitrogen in blue, oxygen in red).

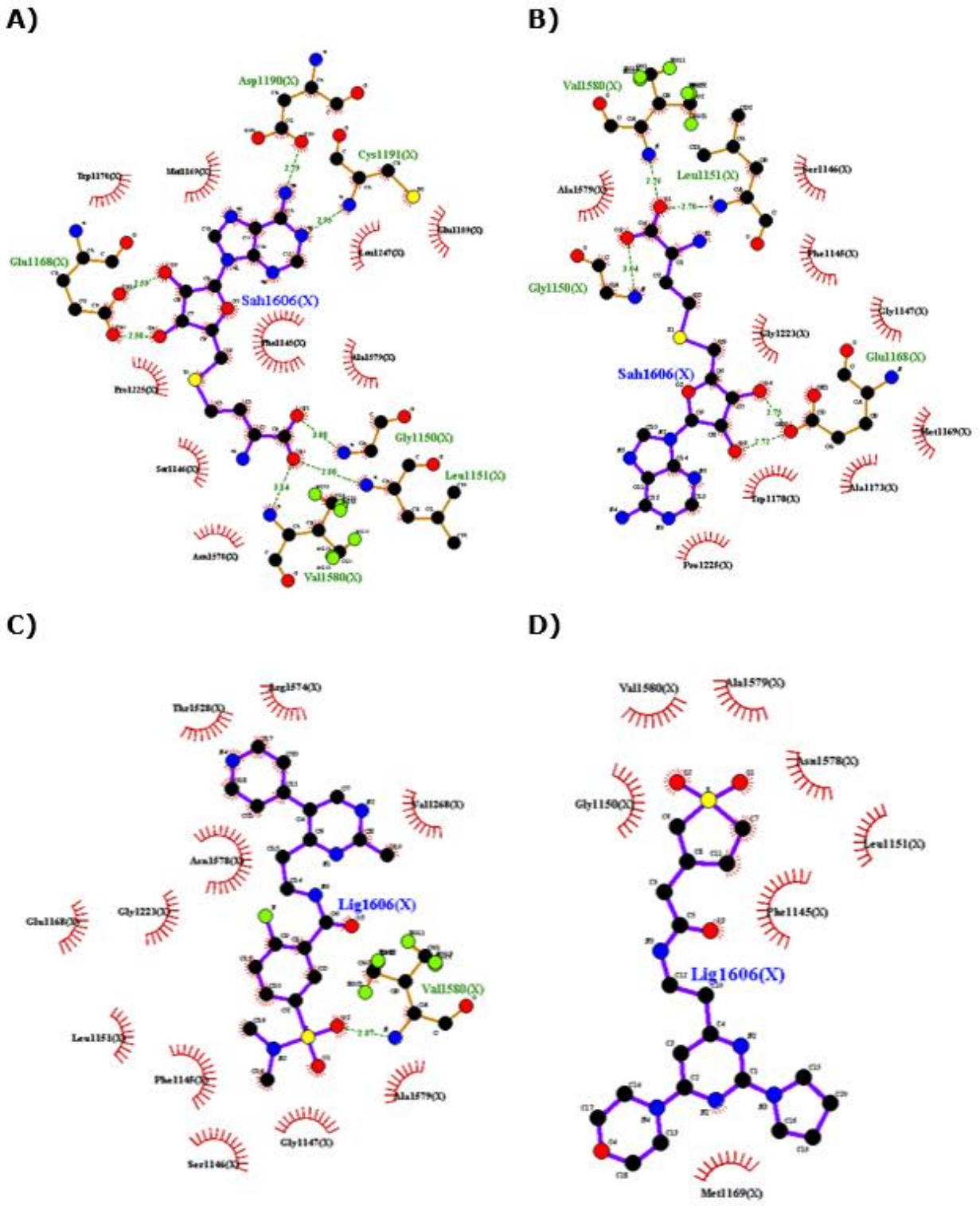


Figure 16. 2-D representation of 4WXX - SAH (A, B), 4WXX - ELITE12 (C), and 4WXX - MYB1 interactions (D). 2D graphs were generated using the LIGPLOT program. Interactions were captured at the first and the 90th frame for 4WXX – SAH, and only at the

90^{th} frame of simulation in the case of ELITE12 and MYB1. The ligand structures are colored in purple. SAH is labeled as Sah1606. ELITE12 (C) and MYB1 (D) are labeled as Lig1606. Hydrogen bonds are indicated by a green dashed line. The number next to the line stands for the donor-acceptor distance. Amino acid residues involved in hydrophobic interactions are visible as curved red lines.

Binding analysis of SAH and MYB1 to 3PTA and 4DA4

We repeated the procedure for the analysis of MD simulations of 3PTA and 4DA4 in a complex with SAH and Maybridge1 compound. We analyzed structural fluctuation within binding sites for SAH and Maybridge1 using RMSD values.

In a complex with 3PTA, Maybridge1 shows significantly less movement within the binding site compared to SAH. SAH fluctuates with the amplitude of variation of approximately 2 Å. The amplitude of inhibitor movement is reduced by 50 % (Figure 17A). Figure 17B presents the chart with RMSD values for SAH and Maybridge1 movement when bound to the protein form 4DA4. The SAH moves relatively uniformly within the binding pocket of the 4DA4. The amplitude of the variation is approximately 0.3 Å. The motion of Maybridge1 differs significantly due to higher RMSD values where the amplitude of variation is around 2 Å.

Then we analyzed the stability of 3PTA-Maybridge1 interactions over the whole MD simulation. We observed the first and the last simulation frame in VMD to see if the inhibitor remained bound to the enzyme. We also repeated the observation of the mutual motion of 4DA4 and MYB1 at the beginning and the end of the simulation. We compared the obtained results with the results from the analysis of the reference complexes 3PTA-SAH and 4DA4-SAH.

Figure 17. A) RMSD values of Maybridge1 indicate better binding to 3PTA compared to SAH. B) High RMSD values of Maybridge1 point to poorer binding to 4DA4 in relation to SAH. The RMSD values were obtained using RMSD Trajectory Tool in VMD *(27)*. The chart was generated using Excel. X-axis stands for the number of frames in simulation. Y-axis presents RMSD values in Å: blue line for SAH, orange line for MYB1 compound.

In the 3PTA-Maybridge1 case, we found that the protein-ligand complex is stable during 100 ns. More precisely, the compound forms interactions at the beginning and end of the simulation and does not exit the binding site. What

is more, aromatic moieties of the molecule are buried deeper in the binding pocket, compared to SAH (Figure 18).

When bound to 4DA4, SAH shows satisfactory retention within the binding site over the 300 ns simulation. On the contrary, the compound Maybridge1 shows good binding at first but drops out of the binding site after 100 ns simulation (Figure 19).

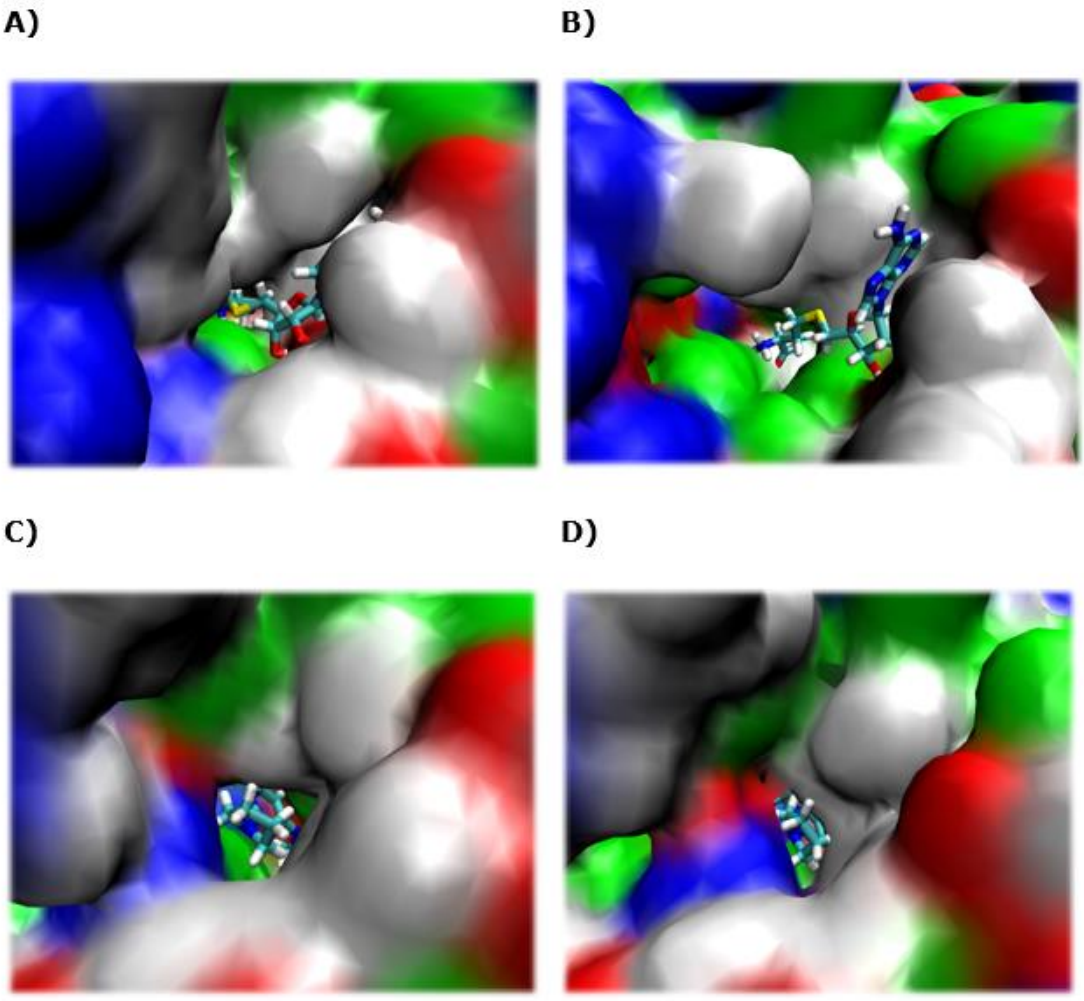


Figure 18. Similar binding of SAH and Maybridge1; Maybridge1 remains bound to 3PTA during the 100 ns simulation. MD simulations were analyzed in VMD *(27)*. Zoomed display of 3PTA binding site with bound SAH was captured in frame 1 (A) and frame 101 (B). Zoomed display of 3PTA-Maybridge1 was captured in frames 1 and 97 (C, D). The protein

surface is depicted by atom density (QuickSurf), and colored by the type of amino acid residues (ResType) - hydrophobic amino acids are colored in white, hydrophilic in green, positively charged amino acids in blue, and negatively charged in red. The Maybridge1 structure is shown using the Licorice representation method and colored by atom type (sulfur in yellow, nitrogen in blue, oxygen in red).

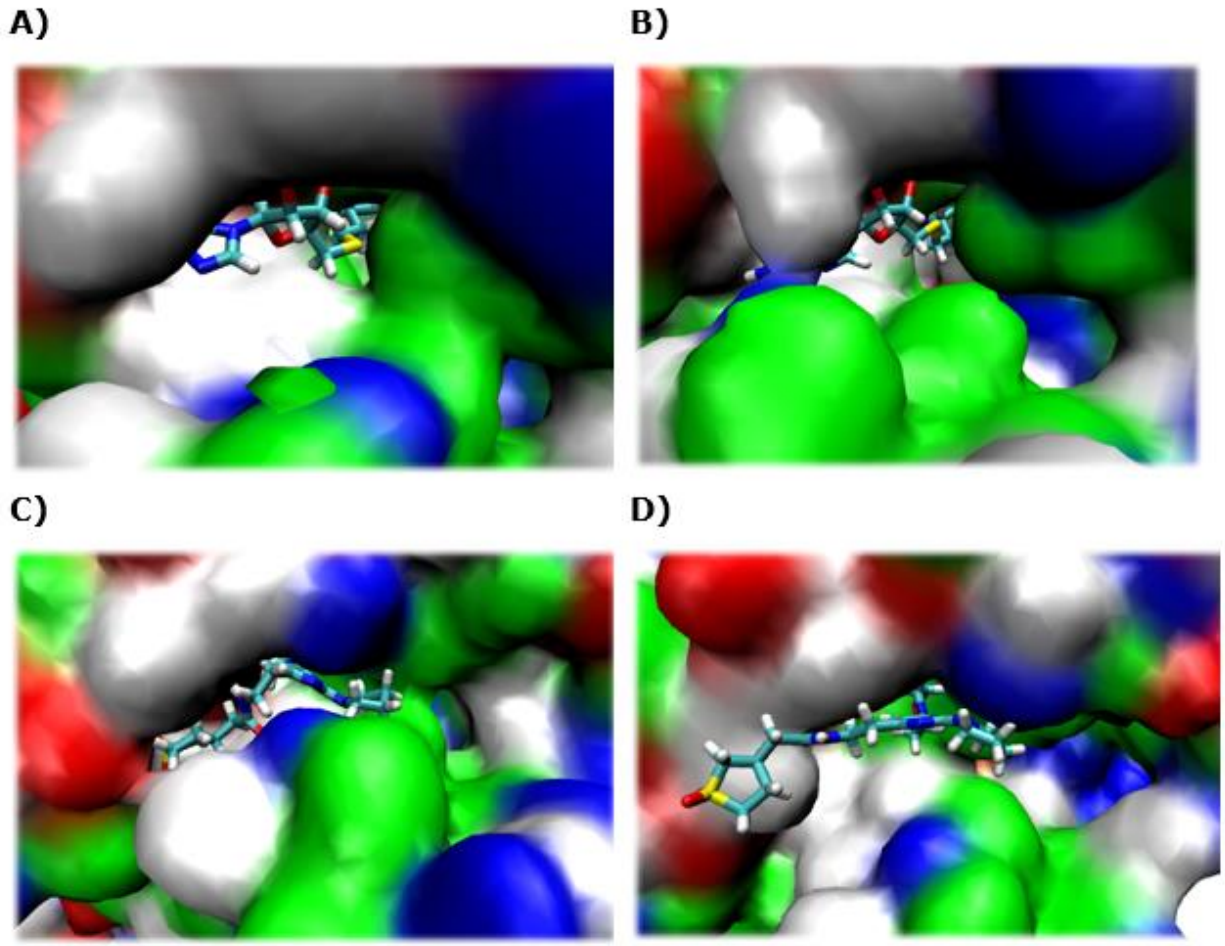


Figure 19. Maybridge1 does not remain bound to 4DA4 after 100 ns simulation, compared to SAH. MD simulations were analyzed in VMD *(27)*. Enlarged display of 3PTA binding site with SAH was captured in frame 1 (A) and frame 101 (B). The binding mode of MYB1 in the first and last frames is shown in C) and D). The protein surface is depicted by atom density (QuickSurf) and colored by the type of amino acid residues (ResType) – white for hydrophobic, green for hydrophilic, blue for positively charged amino acids, and red for negatively charged amino acids. The Maybridge1 structure is shown using the Licorice representation method and colored by atom type (sulfur in yellow, nitrogen in blue, oxygen in red).

Again, we studied interactions in LIGPLOT using the 90th simulation frame, as explained above. In LIGPLOT we showed the formed hydrogen bonds. The number of these polar bonds is three, both for 3PTA-SAH and 3PTA-Maybridge1. Amino acid residues included in the interaction are different for each case. In addition to good binding compared to SAH, Maybridge1 interacts with more amino acids via hydrophobic interactions (Figure 20). This is explained by the fact that the hydrophobic character within the binding site is more pronounced than on the surface.

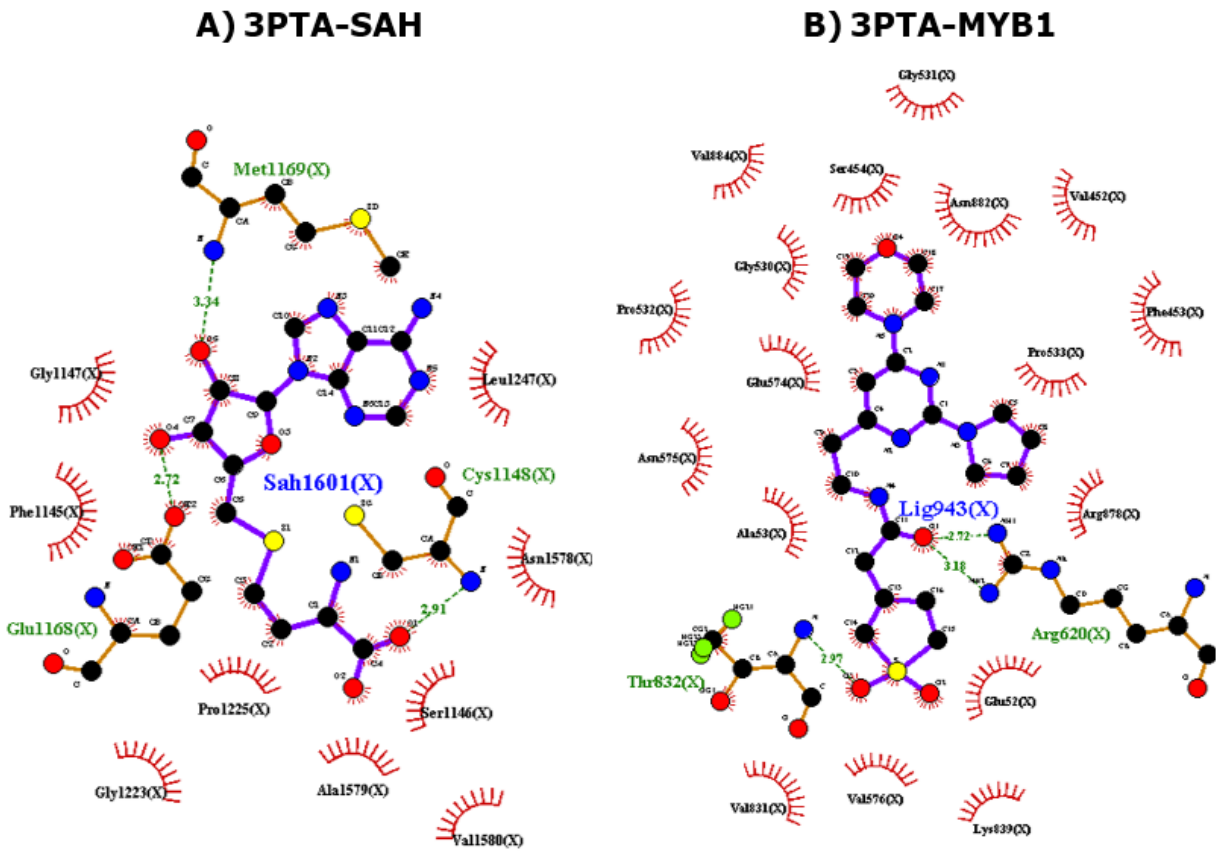


Figure 20. 2-D representation of 3PTA-SAH and 3PTA-MYB1 interactions. 2D graphs were generated using the LIGPLOT *(33)* program. Interactions were captured at the 90th of the simulation. The ligand structures are colored in purple. SAH is labeled as Sah1601 and MYB1 as Lig943. Green dashed line shows hydrogen bonds. The number next to the line stands for the donor-acceptor distance. Amino acid residues involved in hydrophobic interactions are shown as curved red lines. Due to docking, the 3PTA-MYB1 complex shows residues 1-943 so that Arg620 and Thr832 denote Arg1266 and Thr1478 in the original amino acid sequence.

SAH forms many hydrogen bonds within the binding pocket of 4DA4, even eight. Amino acids included in polar interactions are Gly1152, Gly1153, Leu1154, Glu1171, Asp1193, Cys1194, and Val1582. We noticed that the tested compound forms only one hydrogen bond with Gly1150 (Gly384). Common hydrophobic interactions are created between ligands and Met1172 (Met406), Trp1173 (Trp407), and Asn1580 (Asn814) (Figure 21).

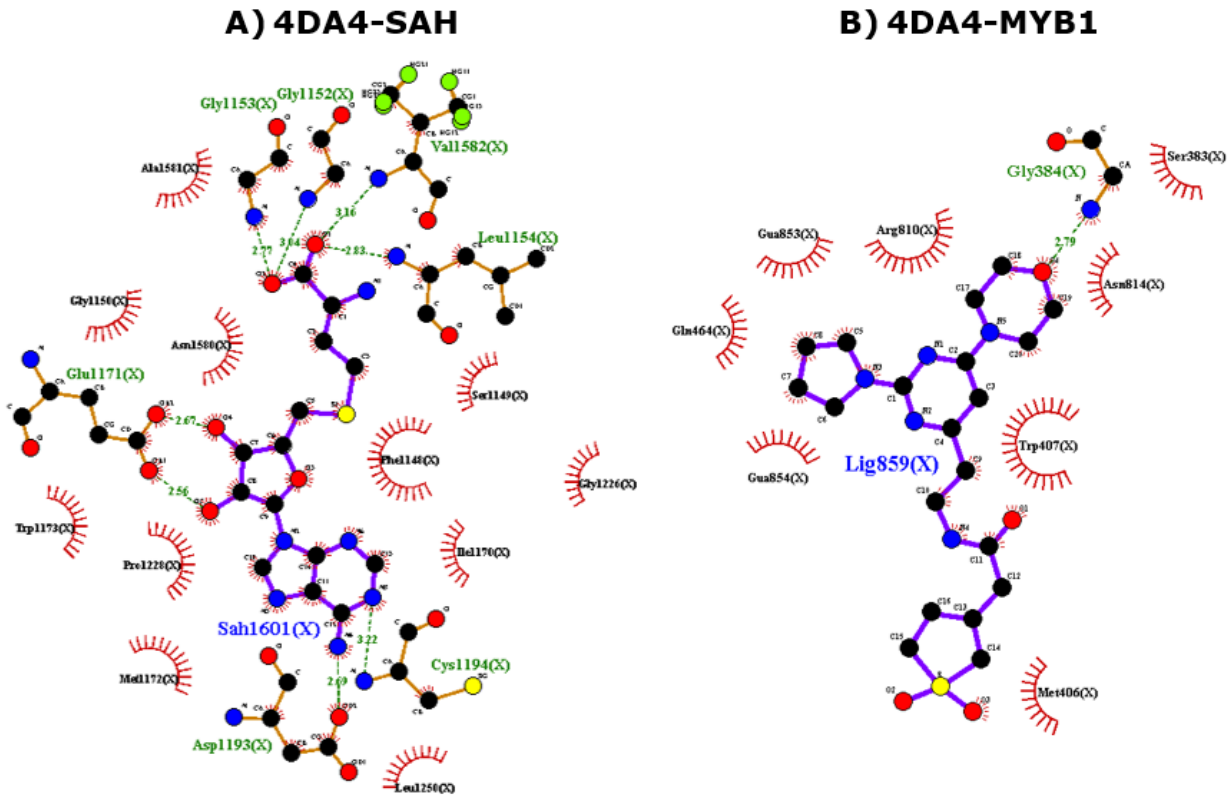
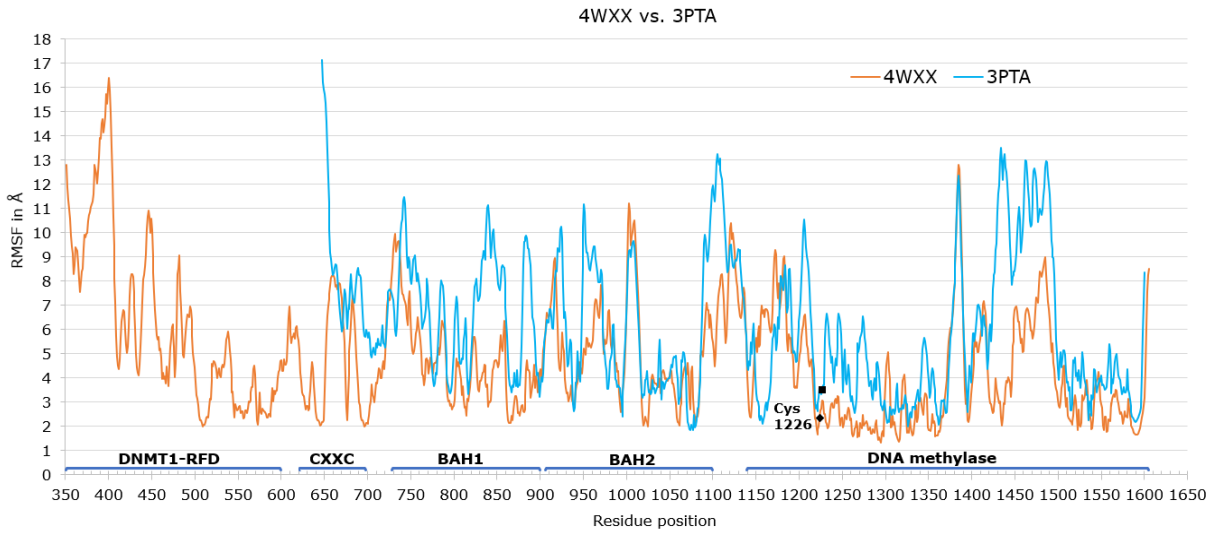


Figure 21. 2-D representation of 4DA4-SAH and 4DA4-MYB1 interactions. 2D graphs were generated using the LIGPLOT *(33)* program. Interactions were captured at the 90th of simulation. The ligand structures are colored in purple. SAH is labeled as Sah1601 and MYB1 as Lig859. Green dashed line are for hydrogen bonds. The number next to the line is for the donor-acceptor distance. Amino acid residues involved in hydrophobic interactions are shown as curved red lines. Due to docking, the 4DA4-MYB1 complex shows residues 1-859 so that Gly384 denote Gly1115 in the original amino acid sequence.

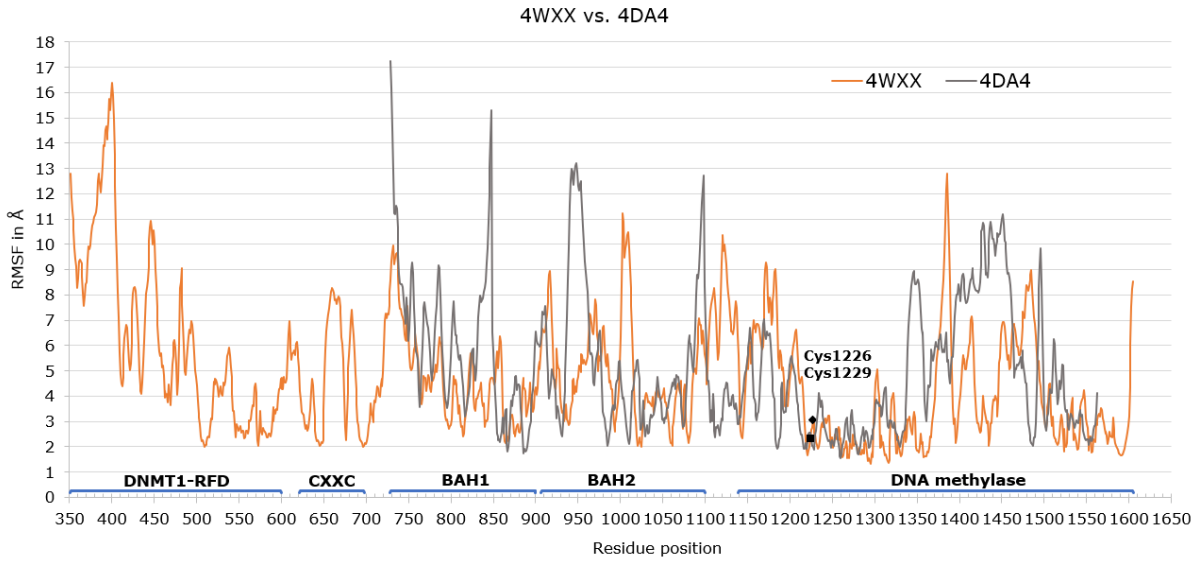
Coarse-Grained analysis of different forms of DNMT1

DNMT1 crystal structures used in this study bind to different forms of DNA (3PTA and 4DA4) or are not complexed with DNA at all (4WXX). We considered them as three different forms of DNMT1 depending on the stage of the DNA methylation process. In addition, previous studies report several turnovers in the DNMT1 catalytic reaction with unmethylated DNA (37). That also points to the existence of different enzyme forms in the reaction. We performed CG data analysis to find similarities between 4WXX, 3PTA, and 4DA4 by comparing the RMSF values of the individual structures. This could provide insight into the mode of transition from one form of protein to another during the methylation process. What is more, the identification of increased residual mobility constitutes an introduction to the discovery of new binding sites.

A)



B)



C)

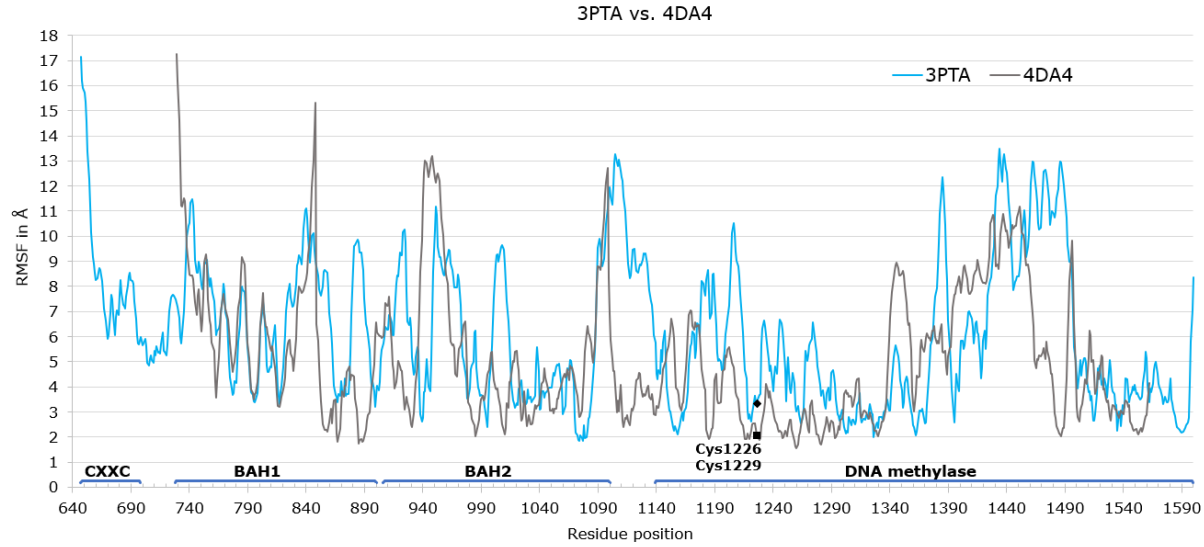


Figure 22. Comparison of residual movement within 4WXX, 3PTA, and 4DA4. RMSF chart for each protein structure was obtained in RStudio using Bio3D package *(32)*. X-axis shows residue position and Y-axis stands for RMSF values in Å. Graphs were overlapped using Excel tools: A) RMSF comparison for 4WXX and 3PTA, B) RMSF comparison for 4WXX and 4DA4, C) RMSF comparison for 3PTA and 4DA4. Orange line is for 4WXX, blue line is for 3PTA and gray line is for 4DA4. Domains are named under the RMSF line. Active site cysteine (Cys1226, Cys1229) is marked with black dot on each line.

The comparison of RMSF values for each structure is shown in Figure 22. In the 4WXX structure, high mobility is present in part of the RFTS domain (residues 397-401), where Glu400 exhibits the highest movement. Lys1003, Leu1384, Glu1385, and Ile1386 are characterized by medium mobility. Lys1003 is part of the BAH2 domain, while residues 1384-1386 belong to the catalytic domain. When comparing 3PTA and 4DA4 with 4WXX, there are noticeable differences in the movement of residues within BAH domains, Gly-Lys linker region, and MTase domain. Ala647, as part of the CXXC domain, shows weak motion in 4WXX. In 3PTA structure Ala647 represents the residue with the highest mobility. Residue on 1385th position (Glu1385) shows relatively equal movement for 3PTA and 4WXX. The same residue in 4DA4 structure (Glu1388) shows significantly less mobility with an amplitude of variation of 7 Å. Both 3PTA and 4DA4 exhibit higher residual mobility at residues 1433-1440 (1436-1443 in 4DA4) within the methyltransferase domain. There are visible differences in BAH domains of 3PTA and 4DA4, referring to a higher fluctuation rate of residues at positions 731 and 851 (Gly851) in 4DA4. Next, there are regions with pronounced opposite residual movement within all shared domains. According to RMSF values, we noticed that 3PTA has the greatest mobility compared both to 4WXX and 4DA4. The catalytic site nearby active site cysteine is not very mobile in 4WXX and 4DA4 structures. The mobility of active site cysteine is slightly increased in 3PTA.

Discussion

ELITE12 and MYB1 can bind to the SAH binding pocket

We found two compounds, named ELITE12 and Maybridge1, that bind to already known and identified SAH binding site. The SBDD technique according to the DNMT1 active site was used, as we did not prove whether there are other binding sites. The SAM/SAH binding pocket occupies the part of the methyltransferase domain is hydrophobic in nature, meaning it is composed mainly of amino acids with a hydrophobic side chain. Therefore, ligands that bind within the binding pocket should have hydrophilic properties. The presence of phenyl and indole rings in Phe1145 and Trp1170 side chains allows π-π stacking interactions due to delocalized electrons. π-π stacking interactions are classified as noncovalent interactions between aromatic moieties with energy lower than 10 kJ/mol. That makes them weaker than covalent bonds, whose energy is more than 200 kJ/mol. π-π stacking interactions are also weaker than noncovalent hydrogen bonds with the energy of 25-40 kJ/mol. It is shown that π-π stacking interactions may significantly contribute to the stability of protein-protein or protein-ligand complexes, along with hydrogen bonds (38). Previous studies reported that the compound structures bind better to catalytic pocket if they correspond to the long SAM/SAH scaffold (18).

Except for similar scaffolding, both molecule structures also contain aromatic moieties due to which π-π stacking interactions with Phe1145 and Phe1148 could be maintained in all three forms of DNMT1 (Figure 12). ELITE12 and Maybridge1 have relatively rigid structures considering pyridine, pyrimidine, and benzamide rings in ELITE12, and pyrimidine ring in Maybridge1. These aromatic rings contain conjugated double bonds around which the rotation is limited. The peptide bond in both structures further increases the rigidity of molecules. Despite this, structures are still quite flexible taking into account

more than seven rotatable bonds. Pharmacokinetic properties indicate satisfactory bioavailability of tested inhibitors, with a deviation in polarity relative to SAH (Table 2). Our compounds do not show violations of the Lipinski rule of five. Nevertheless, water solubility is poorer due to lower polarity, and flexibility is higher for ELITE12 and Maybridge1 relative to SAH. Molecular weight higher than 350 g/mol and the number of freely rotating bonds are characteristics that violate lead likeness. On contrary, inhibitors show higher lipophilicity, indicating better gastrointestinal absorption and facilitated passage through phospholipid bilayers, when compared to SAH.

Inhibitors show different stability in the protein-inhibitor complex when binding to different forms of DNMT1

With the results of MD simulation analysis, we showed that ELITE12 and Maybridge1 are not good drug candidates. Both ELITE12 and Maybridge1 do not interact equally with 4WXX, 3PTA, and 4DA4. The mode of movement of each inhibitor within the binding site was monitored via RMSD values. High values were interpreted as strong motility and vice versa. We initially examined binding to DNMT1 without bound DNA substrate (4WXX) as it is the simplest form of the enzyme, among the three we observed. Despite the peptide bond and large proportion of double bonds, ELITE12 exhibits a high level of motility which makes it difficult to stably interact with the protein. Based on that, ELITE12 is classified as a poorer inhibitor relative to SAH. MD simulations, as well as analysis of established interactions, confirm the weak ability of this compound to remain within the 4WXX binding site (Figure 14, 16). On the other hand, the Maybridge1 compound shows different degrees of stability when binding to a particular enzyme form. This points to the necessity of inhibition of all protein structures present in reaction to develop a successful DNMT1 inhibitor. Maybridge1 has a low affinity for DNMT1 structures 4WXX

and 4DA4 but satisfying binding mode in 3PTA. Considering lower rates of fluctuation in 4WXX and 3PTA binding sites, more stable maintenance of protein-ligand interactions was expected relative to SAH. The interaction of the Maybridge1 inhibitor with the enzyme structure 3PTA confirms the expected binding. Opposite of that, the assumption that Maybridge1 remains buried in the hydrophobic interior of the binding pocket of 4WXX is refuted by a more detailed analysis of interactions. However, due to deeper penetration into the binding cavity, Maybridge1 has the potential to achieve a greater number of hydrophobic interactions. On the other hand, in the part of the simulation, it does not interact with 4WXX via hydrogen bonds. To date, it has not been determined whether hydrophobic interactions are sufficient for the stability of biological complexes. They are classified as relatively weak interactions. They are known to cooperate with other noncovalent interactions, including van der Waals forces, hydrogen bonds, π-π stacking and electrostatic interactions. Hydrogen bonds are considered to be the most important interactions in drug design. These interactions determine ligand conformation best suited to the binding site and the affinity of the ligand for the target (38). The absence of hydrogen bonds between the protein target and the inhibitor indicates that the compound has dropped out of the binding site, which could result in incomplete inhibition of the enzyme. Considering the binding potential for DNTM1 and calculated ADME properties, further research should include experiments on cells to confirm the results obtained using *in silico* methods.

The high level of mobility is attributed to the sites that bind the DNMT1 interaction partners

The coarse-graining models we have created provide insight into a large number of DNMT1 parts and domains. This confirms the size of DNMT1 and the high level of interactions with other macromolecules. There is several

previous studies where CG analysis was used for studying protein folding and aggregation, various interaction networks, mechanism of action, and so on (22). The mode of movement of individual protein domains (Figure 22) represent useful tool to identify parts of the protein suitable to serve as a template for drug design. Sites with enhanced motility have a pronounced specific function. The function is thought to be the regulation of DNA methylation process and DNMT1 itself. The inhibition of these moving parts of DNMT1 mainly results in the induction of hypomethylation. It also may prevent pathological conditions caused by the evolutionary problem of protein unfolding at highly mobile sites. Since each enzyme structure (4WXX, 3PTA, and 4DA4) starts from a different domain, we could not compare the motion of all domains. From the results we were able to obtain, we concluded that these three forms are permeated in action during the DNA methylation process. 4WXX as the initial form shows the greatest movement in the domains responsible for directing the enzyme to the site of action. Although not complexed with DNA, inhibition of this form of the enzyme is necessary given the combined action of all three forms. 3PTA and 4DA4 differ in the position of the cytosine base, according to which 3PTA should be the form following 4WXX. The 3PTA structure starts from the CXXC domain, which shows the strongest motion. This could be explained by the property of CXXC domain to bind hemimethylated sites with high affinity. The highest motility of residues in BAH domains of 4DA4 structure is probably due to its function to recruit DNMT1 to replication foci and to retain it there.

Subsequent studies based on CG models can go in two directions. Firstly, macromolecules considered to be interaction partners should be docked to mobile amino acid residues. Another possibility is the design of drugs precisely for motile protein sites, aiming to change the way of DNA methylation by inhibition of protein-protein or protein DNA/RNA interactions.

Conclusion

Epigenetic changes on the level of DNA methylation affect gene expression thus activating or silencing genes. Dysregulation of DNA methylation can lead to the various pathological conditions. DNMT1 is a large complex enzyme with multiple regulatory mechanisms that interacts with over 40 different proteins, DNA and RNA molecules. Several forms of DNMT1 structure are known, but the regulatory functions of different domains is not known, and thus, cannot be exploited in different drug-design studies.

We docked two compounds from commercial compound libraries ELITE and Maybridge to three different DNMT1 structures. Docking interactions and matching shape between the protein and the compounds has been described using several computational techniques. The interactions have been compared to the binding of the cofactor S-adenosyl-homocysteine that has been well described in the literature. We found poor binding potential for the ELITE compound, and the variable nature of binding depending on the form of the enzyme for the Maybridge compound. Following *in silico* studies, *in vitro* testing is required.

Previous results have led us to CG analysis. According to the RMSF charts, we concluded that the activity of the three forms 4WXX, 3PTA, and 4DA4 is related. We noticed the greatest mobility in RFTS, CXXC, and BAH domains. Based on this, we made the assumption that highly mobile parts bind other proteins. This should be confirmed by identifying and docking these macromolecules to DNMT1 sites to prove interactions and better interpret the role of DNMT1 domains.

Recent insights into the structure and function of DNMT1, as well as the development of novel human DNMT1 inhibitors, opens up new possibilities in the treatment of diseases caused by errors in the regulation of gene activity. What is more, the obtained results form the basis for further research in the

field of the human epigenome, cell reprogramming, and creation of induced pluripotent stem cells.

References

1. Mazzio EA, Soliman KF. Basic concepts of epigenetics. Epigenetics. 2012 Feb;7(2):119–30.

2. Weinhold B. Epigenetics: The Science of Change. Environ Health Perspect. 2006 Mar;114(3):A160–7.

3. Coco C, Sgarra L, Potenza M, Nacci C, Pasculli B, Barbano R, et al. Can Epigenetics of Endothelial Dysfunction Represent the Key to Precision Medicine in Type 2 Diabetes Mellitus? Int J Mol Sci. 2019 Jun 17;20:2949.

4. Dhe-Paganon S, Syeda F, Park L. DNA methyl transferase 1: regulatory mechanisms and implications in health and disease. Int J Biochem Mol Biol. 2011 Jan 30;2(1):58–66.

5. Greenberg MVC, Bourc'his D. The diverse roles of DNA methylation in mammalian development and disease. Nat Rev Mol Cell Biol. 2019 Oct;20(10):590–607.

6. Jung M, Pfeifer GP. CpG Islands. In: Maloy S, Hughes K, editors. Brenner's Encyclopedia of Genetics (Second Edition) [Internet]. San Diego: Academic Press; 2013 [cited 2021 Aug 26]. p. 205–7. Available from: https://www.sciencedirect.com/science/article/pii/B9780123749840003491

7. Elhamamsy AR. Role of DNA methylation in imprinting disorders: an updated review. J Assist Reprod Genet. 2017 May;34(5):549–62.

8. Xiao F-H, Wang H-T, Kong Q-P. Dynamic DNA Methylation During Aging: A "Prophet" of Age-Related Outcomes. Front Genet. 2019;10:107.

9. Heerboth S, Lapinska K, Snyder N, Leary M, Rollinson S, Sarkar S. Use of Epigenetic Drugs in Disease: An Overview. Genet Epigenetics. 2014 May 27;6:9–19.

10. Fernandez-Twinn D, Hjort L, Novakovic B, Ozanne S, Saffery R. Intrauterine programming of obesity and type 2 diabetes. Diabetologia. 2019 Oct 1;62.

11. Covic M, Karaca E, Lie DC. Epigenetic regulation of neurogenesis in the adult hippocampus. Heredity. 2010 Jul;105(1):122–34.

12. Song J, Rechkoblit O, Bestor TH, Patel DJ. Structure of DNMT1-DNA Complex Reveals a Role for Autoinhibition in Maintenance DNA Methylation. Science. 2011 Feb 25;331(6020):1036–40.

13. 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 Dec 14;293(50):19466–75.

14. 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 Jul 31;427(15):2520–31.

15. Shrubsole MJ, Wagner C, Zhu X, Hou L, Loukachevitch LV, Ness RM, et al. Associations between S-adenosylmethionine, S-adenosylhomocysteine, and colorectal adenoma risk are modified by sex. Am J Cancer Res. 2014 Dec 15;5(1):458–65.

16. Miletić V, Odorčić I, Nikolić P, Svedružić ŽM. In silico design of the first DNA-independent mechanism-based inhibitor of mammalian DNA methyltransferase Dnmt1. PLOS ONE. 2017 Apr 11;12(4):e0174410.

17. Medina-Franco JL, Yoo J, Dueñas-González A. Chapter 13 - DNA Methyltransferase Inhibitors for Cancer Therapy. In: Zheng YG, editor. Epigenetic Technological Applications [Internet]. Boston: Academic Press; 2015 [cited 2021 Aug 26]. p. 265–90. Available from: https://www.sciencedirect.com/science/article/pii/B9780128010808000132

18. Juárez-Mercado KE, Prieto-Martínez FD, Sánchez-Cruz N, Peña-Castillo A, Prada-Gracia D, Medina-Franco JL. Expanding the Structural Diversity of DNA Methyltransferase Inhibitors. Pharmaceuticals. 2021 Jan;14(1):17.

19. Antibodies, Proteins, Kits and Reagents for Life Science | Abcam [Internet]. [cited 2021 Sep 10]. Available from: https://www.abcam.com/

20. Lionta E, Spyrou G, Vassilatis DK, Cournia Z. Structure-Based Virtual Screening for Drug Discovery: Principles, Applications and Recent Advances. Curr Top Med Chem. 2014 Aug;14(16):1923–38.

21. 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 Mar 8;93(3):569–75.

22. Singh N, Li W. Recent Advances in Coarse-Grained Models for Biomolecules and Their Applications. Int J Mol Sci. 2019 Jan;20(15):3774.

23. Martínez L. Automatic Identification of Mobile and Rigid Substructures in Molecular Dynamics Simulations and Fractional Structural Fluctuation Analysis. PLoS ONE. 2015 Mar 27;10(3):e0119264.

24. Song J, Teplova M, Ishibe-Murakami S, Patel DJ. Structure-based mechanistic insights into DNMT1-mediated maintenance DNA methylation. Science. 2012 Feb 10;335(6069):709–12.

25. Ren W, Gao L, Song J. Structural basis of DNMT1 and DNMT3A-mediated DNA methylation. Genes. 2018 Dec 11;9:620.

26. Pettersen E, Goddard T, Huang C, Couch G, Greenblatt D, Meng E, et al. UCSF Chimera – A Visualization System for Exploratory Research and Analysis. J Comput Chem. 2004 Oct 1;25:1605–12.

27. Hsin J, Arkhipov A, Yin Y, Stone JE, Schulten K. Using VMD - An Introductory Tutorial. Curr Protoc Bioinforma Ed Board Andreas Baxevanis Al. 2008 Dec;CHAPTER:Unit-5.7.

28. Kaushik M. A review of Innovative Chemical Drawing and Spectra Prediction Computer Software. Mediterr J Chem. 2014 Apr 4;3(1):759–66.

29. Daina A, Michielin O, Zoete V. SwissADME: a free web tool to evaluate pharmacokinetics, drug-likeness and medicinal chemistry friendliness of small molecules. Sci Rep. 2017 Mar 3;7(1):42717.

30. Butt SS, Badshah Y, Shabbir M, Rafiq M. Molecular Docking Using Chimera and Autodock Vina Software for Nonbioinformaticians. JMIR Bioinforma Biotechnol. 2020 Jun 19;1(1):e14232.

31. 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 Sep 1;1–2:19–25.

32. Skjærven L, Jariwala S, Yao X-Q, Idé J, Grant BJ. The Bio3D Project: Interactive Tools for Structural Bioinformatics. Biophys J. 2016 Feb 16;110(3):379a.

33. Laskowski RA, Swindells MB. LigPlot+: Multiple Ligand–Protein Interaction Diagrams for Drug Discovery. J Chem Inf Model. 2011 Oct 24;51(10):2778–86.

34. 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 Jan 12;12(1):405–13.

35. Michelin O, Zoete V. SwissADME [Internet]. SwissADME. [cited 2021 Aug 26]. Available from: http://www.swissadme.ch/

36. Pence HE, Williams A. ChemSpider: An Online Chemical Information Resource. J Chem Educ. 2010 Nov 1;87(11):1123–4.

37. Svedruzic Z. Mammalian Cytosine DNA Methyltransferase Dnmt1: Enzymatic Mechanism, Novel Mechanism-Based Inhibitors, and RNA-directed DNA Methylation. Curr Med Chem. 2008 Jan 1;15(1):92–106.

38. Deng J-H, Luo J, Mao Y-L, Lai S, Gong Y-N, Zhong D-C, et al. π-π stacking interactions: Non-negligible forces for stabilizing porous supramolecular frameworks. Sci Adv. 2020 Jan 1;6(2):eaax9976.

Curriculum vitae

Morana Rijako

Date of birth: 19/02/1997 | (+385) 997500468 | rijako.morana@gmail.com | Kragujski put 1, 34550, Pakrac, Croatia

WORK EXPERIENCE

02/2021 – 09/2021
MASTER'S THESIS

Thesis: *Mammalian DNA methyltransferase DNMT1: structure and function analysis and development of novel inhibitors using in silico methods*
Mentor: doc.dr.sc. Željko M. Svedružić, Ph.D.
Department of Biotechnology | 51000, Rijeka, Croatia

02/2019 – 02/2019
PEER LEARNING ASSISTANT IN PHYSICAL CHEMISTRY COURSE

Lecturer: doc.dr.sc. Duško Čakara, assistant professor
Department of Biotechnology | 51000, Rijeka, Croatia

07/2018 – 07/2018
PRACTICAL WORK – JGL D.D.

Vigilance and Medical Department: Global Registrations | 51000, Rijeka, Croatia

07/2014 – 08/2014
PRACTICAL WORK – PHARMACY PROHASKA-TOME

34550, Pakrac, Croatia

EDUCATION AND TRAINING

2015 – 2018 – Rijeka, Croatia
BACHELOR OF BIOTECHNOLOGY AND DRUG RESEARCH – University of Rijeka, Department of Biotechnology, Biotechnology and Drug Research

2011 – 2015 – Bjelovar, Croatia
PHARMACY TECHNICIAN – Medical school Bjelovar

LANGUAGE SKILLS

Mother tongue(s): CROATIAN

Other language(s):

	UNDERSTANDING		SPEAKING		WRITING
	Listening	Reading	Spoken production	Spoken interaction	
ENGLISH	B2	C1	B2	B1	B2

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

DIGITAL SKILLS

MS Office package

Microsoft Word | Microsoft Powerpoint | Microsoft Excel | Outlook

Basic knowledge in:

UCSF Chimera | VMD (Visual Molecular Dynamics) | CHARMM | PyMoL | MarvinSketch | GROMACS package for Molecular Dynamics simulations | GAMESS | LigPlot | AVOGADRO

DRIVING LICENCE

Driving Licence: B