Přírodovědecká fakulta
Faculty of Science

Jihočeská univerzita v Českých Budějovicích
University of South Bohemia in České Budějovice

Isolation and molecular characterization of the histone methyltransferase DOT1L from the soft tick *Ornithodoros moubata*

Bachelor thesis

Laboratory of Host-Pathogen interaction Evolution
Institute of Parasitology, Biological Centre,
Academy of Sciences of Czech Republic, Faculty of Science,
University of South Bohemia

Julia Gobl
Supervisor
Dr. Alejandro Cabezas-Cruz
Co-supervisor
Dr. Ryan O. M. Rego

České Budějovice 2018

Annotation:

The aim of this thesis was to isolate the histone methyl transferase *dot1l* from the organism *Ornithodoros moubata* and to further characterize this enzyme in our model organism.

Affirmation:

I hereby declare that I have worked on the submitted bachelor thesis independently. All additional sources are listed in the bibliography section. I hereby declare that, in accordance with Article 47b of Act No. 111/1998 in the valid wording, I agree with the publication of my bachelor thesis, in full form to be kept in the Faculty of Science archive, in electronic form in publicly accessible part of the STAG database operated by the University of South Bohemia in České Budějovice accessible through its web pages. Further, I agree to the electronic publication of the comments of my super visor and thesis opponents and the record of the proceedings and results of the thesis defense in accordance with aforementioned Act No. 111/1998. I also agree to the comparison of the text of my thesis with the Theses.cz thesis database operated by the National Registry of University Theses and plagiarism detection system.

……... ……...

České Budějovice, Date Julia Gobl

Acknowledgment

There are a lot of people whom I really want to thank for all their support during my studies and the work on this thesis.

First of all I am really grateful for the support of my parents, who even made it possible for me to join this study program. Together with my brother they always supported me on the long way of this study, showed a lot of understanding when I was busy and needed their help and always knew how to cheer me up. Danke für all die Unterstützung die ihr mir über all diese Jahre zukommen habt lassen!

I want to express my outmost gratitude to my supervisor Dr. Alejandro Cabezas-Cruz who enabled it for me to get a lot of insights in the progress of scientific work and therefore also made it possible for me to collect a lot of experiences on the way to the final thesis. I also want to thank him for his sympathy when time for the thesis was limited due to other duties. Thank you for all your motivation and enthusiasm that you passed onto me as well as the confidence you gave me!

I want to thank my co-supervisor Dr. Ryan O. M. Rego, who always was there to help me with his experience and expertise when I got stuck in any experiment, as well as to leaven the atmosphere with his delightful humour. Thank you for the time you always took for discussion and aid for this work.

Big thanks to James Valdés who enabled us to do homology protein modelling due to his knowledge, experience and expertise in this field. Thank you for all the nice meetings and time you spent to show me what to do. Also I want to thank you a lot for showing me the bright side of České Budějovice.

I also want to thank the staff of the Parasitology institute who offered their experience and knowledge to make certain experiments possible.

I need to thank my former study colleagues, flatmates and now great friends Marnol and Anna for all their support, help, understanding and everything that they have done for me over the past years of studying. You two really helped me out so many times during this time and I cannot say how grateful I am. Not only for your help, but also to have met you!

I want to thank all my other study colleagues for the solidarity and support during this study. It is not taken for granted to always help each other out with hard exams etc.

Last, but obviously not least, I want to express a special thanks to Prof. Libor Grubhoffer. You always helped us students, no matter how big or small our problem was. Thank you for all your passion for Biological Chemistry that you passed on to us.

And then finally last I want to pass on some extra thanks to all my friends, that did not give me up, despite the fact that I could not spend the time with them they deserved.

Table of Content

1. Abstract

The research field of epigenetics has gained more and more attention over the past decade. Studies are published concerning different functions of epigenetic modifying enzymes, like gene expression regulation, somatic cell reprogramming and embryonic development. Seldom are those investigation done on models from the subphylum of Chelicerata. Dysregulation of histone modifications may cause cancer, and this might happen in several ways. Alternation of gene expression programs, like regulation of tumor suppressors for example. Since DOT1L bears a SAM binding catalytic site, which is unique in regard to histone methyltransferases, it could be an important target for future cancer treatments. There exist indications that epigenetic mechanisms modulate the interactions between hosts and parasites. For example it is known that bacteria alternate the levels of histone modifying enzymes in ticks to ease pathogen infections. Therefore research in this field would be necessary to gain more information about epigenetics, also in ticks. This may lead to important insights that could help in the development of anti-tick vaccines or new acaricides.

Our model organism *O. moubata* has certain advantages over other tick species from the Ixodidae family. It can be reproduced faster and easier since female *O. moubata* only need minutes to hours to be able to lay eggs, in contrast to Ixodids that need days for feeding. Also, offspring can be obtained from a single female for several times with only one mating event. This makes it easy to get a stable tick supply to enable various experiments.

Our initial objective was to test the presence of *dot1l* in *O. moubata*. After the presence of this enzyme was confirmed, various tools, experimental techniques and phenotypic observations were used to characterize the function of DOT1L in *O. moubata.*

The obtained results indicate high similarities between DOT1L from *O. moubata* and other species in the primary, secondary and tertiary structure. Quantitative real time PCR on *dot1l* mRNA suggests an important role in molting of the ticks. To further investigate this, a larval immersion test (LIT) was conducted using an inhibitor of DOT1L. The results further suggest an essential role of DOT1L in the molting of *O. mouabata.*

This work was done to investigate DOT1L in a soft tick and will pave the way for future studies on the tick epigenetics with implications for vector and possibly pathogen control.

2. Introduction

2.1. Generalities and taxonomy of ticks

Ticks can be found worldwide and there are about 900 different species (Estrada-Peña, 2015b). Ticks cause direct damage to their hosts and are therefore of enormous medical and veterinary relevance. Ticks are second to mosquitoes as vectors of human pathogens and are the most important vectors affecting cattle worldwide (Peter *et al.*, 2005; Manzano-Romn *et al.*, 2012). Most of the ticks can be assigned to one of the two main families, the Argasidae, also known as soft ticks and the Ixodidae, also known as hard ticks. There also exists the family Nuttalliellidae, but this one is monotypic and only contains one species *Nuttalliella namaqua* (Guglielmone *et al.*, 2010). 707 species of hard ticks have been accepted until May 2017 and their taxonomic situation has been studied in depth leading to almost complete agreement of the taxonomic position of the families and genera (Estrada-Peña, 2015b). In contrast to Ixodidae, the family Argasidae , which consists of about 190 species, show uncertainty for genus and species taxonomy levels (Estrada-Peña, 2015b).

Members of the families Argasidae and Ixodidae differ in their life cycles, morphology and physiology (Manzano-Romn *et al.*, 2012; Estrada-Peña, 2015b). Generally, argasids do not possess a dorsal shield or scutum, but their integument is very tough and of rough texture. Their capitulum is less prominent and ventrally located. They have small spiracular plates and their coxae do not have spurs (Manzano-Romn *et al.*, 2012). Morphological differences between soft and hard ticks are shown Figure 1.

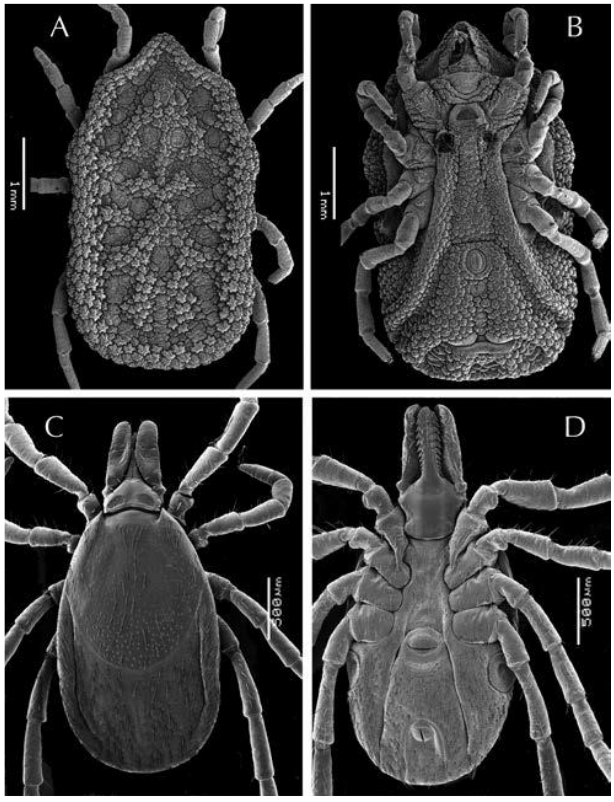

Figure 1: General morphology of ticks of the families Argasidae and Ixodidae; A and B show dorsal and ventral views of a female *Ornithodoros puertoricenis* (soft tick); C and D show dorsal and Ventral views of female *Ixodes ricinus*. Figure taken from (Estrada-Peña, 2015)

2.2. Ticks life cycles

Besides the egg, ticks undergo three different life stages: namely larva, nymph and adult (male or female). Both, soft and hard ticks undergo those life stages, but when they evolve is different. While Ixodidae feed for several days to weeks to ingest more than 100 times their body mass in blood, Argasidae can rapidly expand and engorge up to ten times their body weight within a few hours or even minutes (Estrada-Peña, 2015b). Ixodid ticks ingest blood three times in their life. After the first two blood meals they molt to the next developmental stage. Adult female Ixodidae then ingest an enormous blood meal and mate once. After dropping to the ground, they start laying eggs. They lay thousands of eggs in one batch and die afterwards (Alan R. . Walker 1994).

Soft ticks differ in the fact, that they have more than three developmental stages (Figure 2). The ticks usually have two to eight nymph stages. The actual number varies according to its species, future adult gender and state of nutrition. Females do not die after laying eggs but can feed and reproduce repeatedly (Sonenshine, 1992; Manzano-Romn *et al.*, 2012).

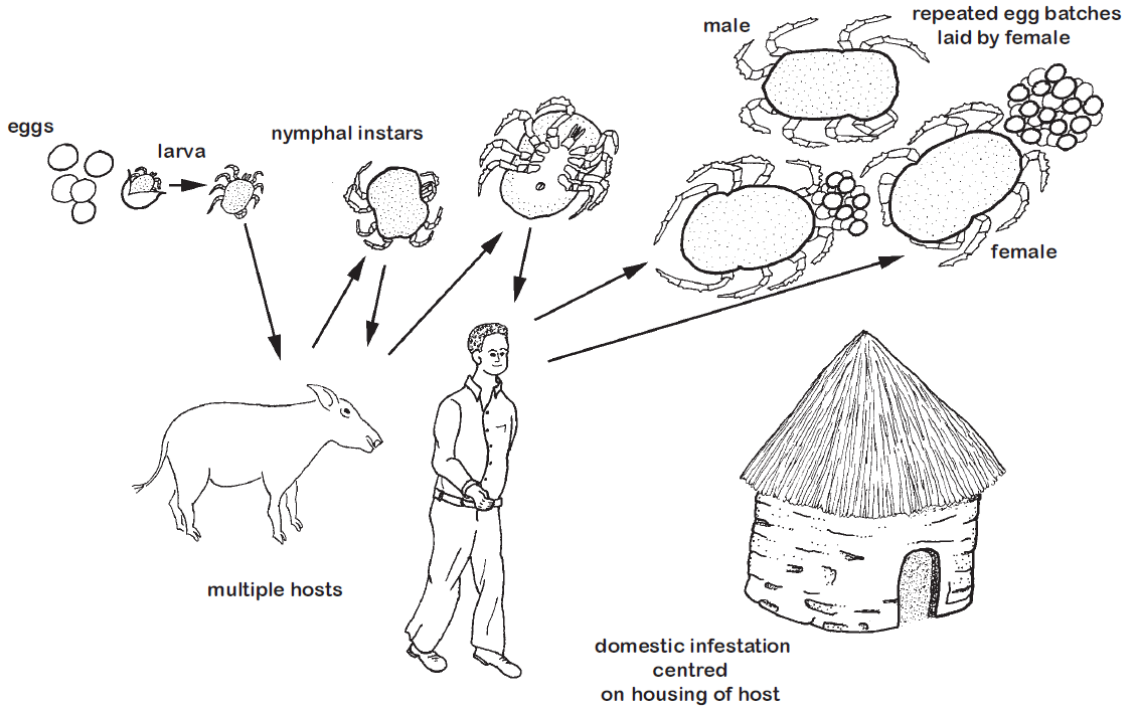

Figure 2: Typical soft tick life cycle, involving several nymph stages and repeated egg laying (taken from: Ticks of Domestic Animals in Africa; Figure taken from (Walker et al. 2014)

2.3. *Ornithodoros*

Ornithodoros is a genus of the family Argasidae. Nymphs and adults show typical soft tick features and its outline from above is oval. They don't have a lateral surture and their integument has numerous small bumps and fine setae. Some even have small eyes, but *O. moubata* don't. Female and male *Ornithodoros* look very similar except for a small difference in the genital pore (Figure 6). They are found in all zoogeographical regions, but prefer dry habitats (Alan R. . Walker 1994). The species *O. moubata* is

located in Central and South Africa. Its hosts include humans as well as domestic and sylvatic pigs (Parola and Raoult, 2001; Ntiamoa-Baidu *et al.*, 2004; Manzano-Romn *et al.*, 2012). Our model organism *Ornithodoros moubata* constitutes one of the most important disease vectors in Africa, including viruses responsible for the African swine fever (Denyer and Wilkinson, 1998). Generally speaking, *Ornithodoros* ticks have a long life span sometimes even up to 15-20 years for some species in adult stage. Therefore, African Swine Fewer Virus (ASFV)-infected soft tick populations can maintain the virus for years (Manzano-Romn *et al.*, 2012).

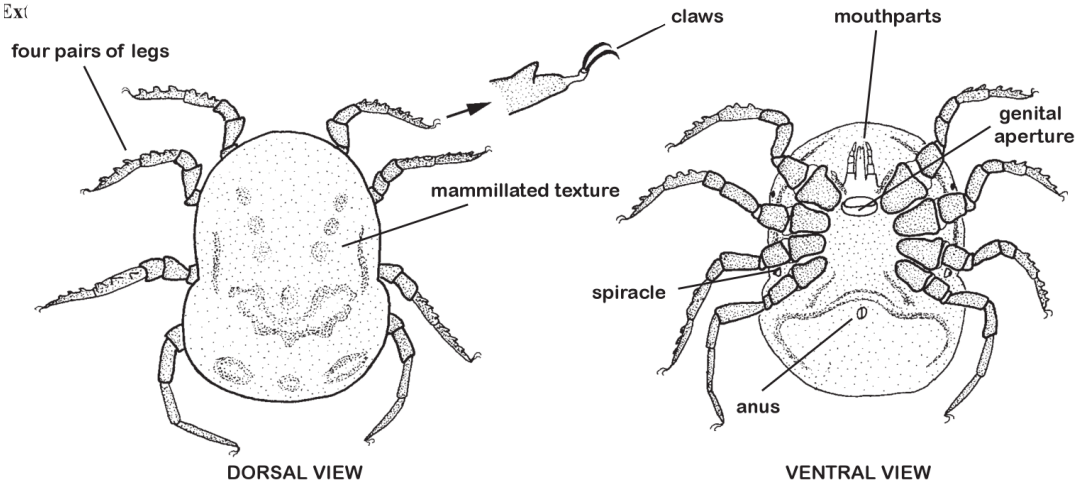

Figure 3: External structure of an *Ornithodoros* tick (Adapted from: Ticks of Domestic Animals in Africa; Figure taken from (Walker et al. 2014)

2.4. Epigenetics

The word "epigenetics" is derived from the Greek word "Epi", which can be translated as "on top of". Therefore, the word "epigenetics" literally means "on top of genes", or in other words, in addition to changes in the gene sequences (Weinhold, 2006). More precisely epigenetics refer to 'stimuli-triggered changes in gene expression due to processes that arise independent of changes in the underlying DNA sequence' (Gómez-Díaz *et al.*, 2012). There exist three major types of epigenetic modifications, namely DNA methylation (Suzuki and Bird, 2008), histone modifications (Bannister and Kouzarides, 2011) and non-coding RNAs (Storz, 2002; Gómez-Díaz *et al.*, 2012; Walter and Hümpel, 2017)

DNA methylation is a chemically very stable covalent modification of mostly cytosine bases. It can even be directly copied and passed on during cell division. Methylation of DNA either leads to repression or activation of gene transcription, but in large parts of the genome it acts as a signal to inactivate repetitive DNA structures and transposons (Walter and Hümpel, 2017).

Histone modifications are performed by so-called histone modifying enzymes (HMEs) and there are several different types of HMEs that perform different modifications on different histones. Methylation, acetlylation, ubiquitylation, phosphorylation and sumoylation comprise the main types of modifications (Weinhold, 2006). Mostly the N-termini of histone tails are targets of HMEs, but studies show that the histone cores can be modified as well, even if they are less accessible for binding partners (Kouzarides,

2007). Histone modifications either lead to gene repression or activation, depending on the type of modifications that are combined.

Small regulatory RNAs can silence genes or transcripts. These epigenetic regulatory mechanisms were first discovered in plants, but subsequently they were also discovered in almost all organisms, including humans. Small RNAs, like pirRNAs (in gametes), miRNAs and siRNAs (in all cells) and long non-coding RNAs (lncRNAs) are important to establish and implement epigenetic processes. RNAs can have a direct epigenetic effect, an indirect intermediary effect or have an implementation function. For example, they can directly silence genes, or they can lead histone-modifying and DNA-modifying enzymes to particular target regions to control the formation of heterochromatin. Therefore, there exists a close interplay between small RNAs and epigenetic modifications (Walter and Hümpel, 2017).

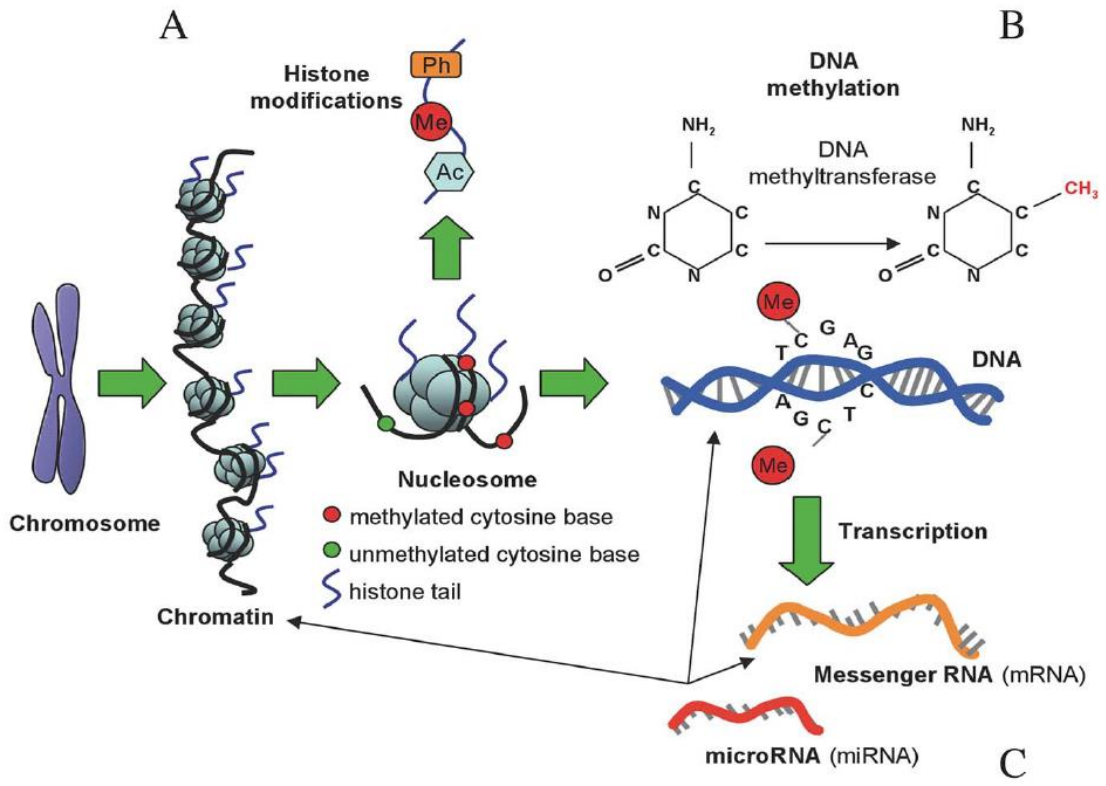

Figure 4: Types of epigenetic modifications. (A) Histones can undergo several different modifications, namely: phosphorylation (Ph), methylation (Me), acetylation (Ac) and more. Certain modifications lead either to gene activation or silencing. (B) DNA molecules are methylated on the cytosine bases by DNA methyltransferase. (C) small RNAs can repress the translation of mRNA into proteins
Figure taken from (Gómez-Díaz et al. 2012)

2.5. Tick epigenetics

In recent years several papers were published addressing HMEs in ticks. The model organism of ticks provides interesting features since first, they are easy to keep and reproduce and second, and more compelling, ticks constitute the most important vector for animal disease and second important vector for human disease worldwide, transmitting pathogens (de la Fuente *et al.*, 2008). A growing number of papers is published revealing that tick-host-pathogen interactions constitute a conflict and cooperation system (de la Fuente, Villar, *et al.*, 2016).

Investigated in the view of evolutionary processes, this system could be the result of coevolution that includes long lasting interactions between organisms (Wade, 2007). These interactions include different mechanisms from tick, host and pathogen. For example, pathogens manipulate host and tick to facilitate infection, multiplication and transmission. While host and tick try to limit pathogen infection by different mechanisms. Those mechanisms, the pathogen profits from, include also the manipulation and control of host cell epigenetics, like in the case of *Anaplasma phagocytophilum* (de la Fuente, Estrada-Peña, *et al.*, 2016). Papers are published investigating epigenetic changes of organisms upon bacterial infection. One paper from Cabezas-Cruz et al. (2016) reported the presence and primary structure of histones and HMEs in the tick *Ixodes scapularis* and further investigated direct changes in the levels of tick HMEs in response to *A. phagocytophilum* infection. The study suggests that there exists a compensatory mechanism from *A. phagocytophilum* to manipulate tick HMEs to regulate transcription and apoptosis for facilitating infection, but on the other hand keep the ticks' fitness to assure survival of the pathogen and the tick. Those studies are important for addressing the interactions of ticks, host and pathogens to investigate new targets for the control of tick borne diseases. Our model organism *O. moubata* constitutes one of the most important disease vectors in Africa, including viruses responsible for the African swine fever (Denyer and Wilkinson, 1998). Therefore, it is also important to consider possible epigenetic changes of the tick upon pathogen infection to be able to address possible targets for control of vector-borne diseases.

2.6. Histone modifying enzymes (HME)

There exists evidence that histone modifications play a fundamental role in multiple biological processes (Bannister and Kouzarides, 2011). Different specific amino acid residues on the histones are target to modifications and each of those modifications is performed by different enzymes. A large variety of those different enzymes is known up to day. These modifications do not regulate gene transcription alone, but also recruit proteins and complexes with specific enzymatic activities.(Bannister and Kouzarides, 2011) (Figure 5) All those different modifications construct a vast network of information inscription. Together they can be called the histone code, after David Allis. Who came up with the hypothesis that different combinations of histone modifications lead to unique cellular responses (Jenuwein, 2001). Although recent research indicates that some histone modifications are closely correlating with each other, which diminishes the number of potential combinations (Schubeler, 2004), the histone code is extending the information potential of the genetic code by forcing chromatin condensation from euchromatin to heterochromatin and vice versa, depending on the constellation of histone modifications (Jenuwein, 2001). Our target enzyme DOT1L is responsible for mono-, di- an trimethylation of Lysine 79 on Histone 3 (K79H3). Depending on the type of methylation, genes can either be activated or repressed (Zhang, 2001). Monomethylation of H3K79 is associated with gene activation, while trimethylation leads to gene repression. Dimethylation does not correlate with the transcriptional statues (Barski *et al.*, 2007; Nguyen and Zhang, 2011).

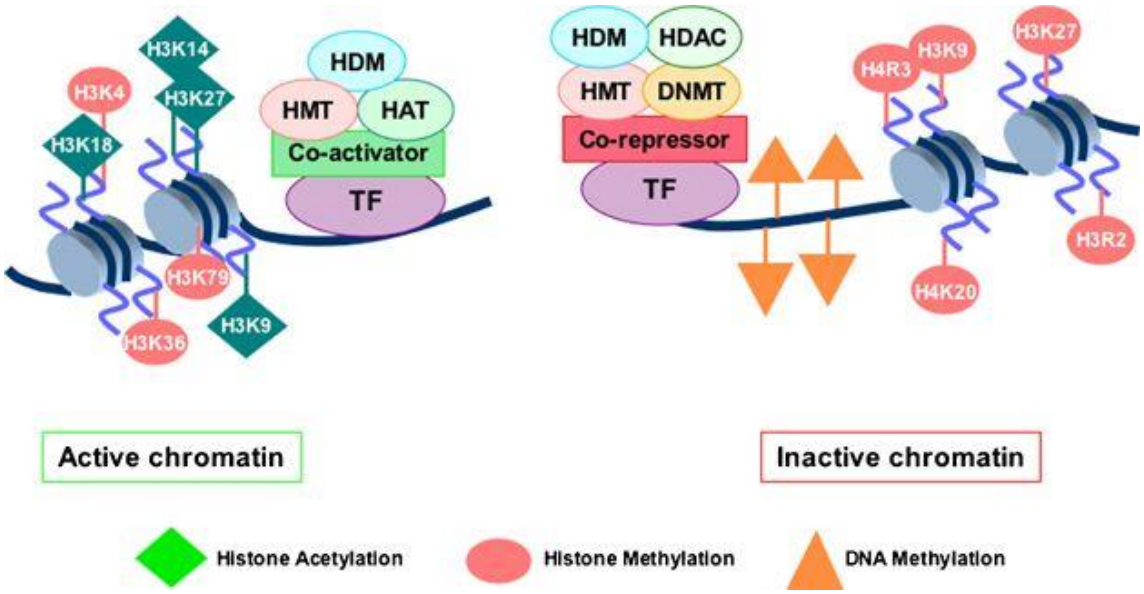

Figure 5: Specific transcription factors (TF) recruit DNA and histone-modifying enzymes to target gene promoters. Hyperacetylation and methylation of H3K4, H3K79 and H3K36 in promotor regions leads to active transcription (left). DNA methylation, hypoacetylations and methylation of H3K9, H3K27 and H4K20 residues leads to gene repression. (right) Those modifications are done by DNA methyltransferases (DNMT), histone acetyltransferases (HAT), histone deacetylases (HDAC), histone methyltransferases (HMT) and histone demethylases (HDM). Figure taken from: (K. L. Rice, Hormaeche, and Licht 2007)

2.7. DOT1L and its role in embryonic development

Histone methylation is catalyzed by different enzymes. The enzymes can be divided into two classes. The first class contains a SET domain, that is evolutionary widely conserved (Jenuwein *et al.*, 1998). The second class does not contain a SET domain, but a SAM binding pocket that is similar to arginine histone methyl transferases (HMTs) (Min *et al.*, 2003; Cheng, Collins and Zhang, 2005). The only enzyme belonging to this class is DOT1 (disruptor of telomeric silencing 1) (Singer *et al.*, 1998), and its homologues in other organisms (Feng *et al.*, 2002). DOT1 can either mono-, di- or trimethylate H3K79 (Zhang, 2001). DOT1 is thought to be the only enzyme responsible of H3K79 methylation, since knockout in yeast, flies and mice lead to complete deletion of H3K79 methylation (van Leeuwen, Gafken and Gottschling, 2002; Shanower, 2004; Jones *et al.*, 2008).

DOT1L methylation is target to an evolutionary conserved crosstalk between ubiquitin on the C-terminus of histone H2B and methylation on H3K79 (Bannister and Kouzarides, 2011). Since histone methylation does not alter the charge of the histones, like acetylation and phosphorylation (Bannister and Kouzarides, 2011), it is possible that it serves as a marker for further modifications to alter the accessibility of the underlying DNA.

Although there exist several studies showing that H3K79 methylation is subject to dynamic regulation (Ooga *et al.*, 2008; Schulze *et al.*, 2009), there was no enzyme characterized yet responsible for the demethylation of this specific lysine residue (Wong, Polly and Liu, 2015).

For a long time, studies on DOT1 were mainly conducted on yeast. But nowadays there exists evidence that DOT1L also plays a critical role in mammalian development (Jones *et al.*, 2008). In one instance mice deficient for DOT1L were generated. The deficiency of the *dot1l* gene was confirmed using the expression of *lacZ* and X-gal staining. The *dot1l* deficient embryos show a variety of developmental abnormalities and died between 9.5 and 10.5 days post coitum (Jones *et al.*, 2008).

In another instance DOT1L was knocked down in the amphibian *Xenopus tropicalis* by a specific transcription activator-like nuclease. When expressed in fertilized embryos the mutation efficiency was high, resulting in DOT1L knockout embryos with little H3K79 methylation. In this case DOT1L knockdown had no apparent effect on embryogenesis of the tadpoles but led to lethality prior metamorphosis (Wen *et al.*, 2015).

Those studies suggest that DOT1L plays a critical role in embryonic development. This makes it an interesting target for investigation in the embryonic development of Arachnids as well.

3. Hypothesis

Dot1l is conserved in *Ornithodoros moubata* and shows similar molecular signatures as homologs from other organisms

4. Objectives

I. To amplify and sequence *dot1l* from the soft tick *Ornithodoros moubata* (OmDot1L).
II. To make a molecular and phylogenetic characterization of OmDOT1L.
III. To explore the potential function of OmDOT1L in *Ornithodoros moubata*.

5. Materials and Methods

5.1. Identification and sequencing

The first goal of our project was the identification and sequencing of the coding sequence of the enzyme DOT1L. Therefore, various methods were used. Namely:

- Primer design
- Isolation of total RNA
- cDNA synthesis
- Gene amplification with polymerase chain reaction (PCR)
- Cloning and transformation in *Escherichia coli*
- Plasmid purification
- Sequencing followed by sequence analysis

5.1.1. Primer design

For the design of primer pairs to amplify *dot1l* from cDNA of *O. moubata*, sequences of *O. rostratus* (Genbank accession GCJJ01002031) and *O. turicata* (Genbank accession GDIE01116948) were aligned using MAFFT 7 alignment server with default settings (Kuraku *et al.*, 2013; Katoh, Rozewicki and Yamada, 2017). *O. rostratus* and *O. turicata* sequences were retrieved from Transcriptome Shotgun Assembly Sequence Database by performing tblastn on the amino acid sequence of DOT1L from *I. scapularis* (Genbank accession ISCW021226).

The resulting alignment was then used to identify sequences of highest similarity. Those highly conserved regions were used as templates to design PCR primers for the amplification of the N-terminus of OmDOT1L. Final primers are shown in Table 1**.**

Table 1: Primers used for the amplification of *dot1l* from *O. moubata*

Primer	Sequence 5’ to 3’
OmDot1LF	GCTCTCGACATCGGACAAA
OmDot1LR	AGGTTCCTGTCTGTGATGCG

5.1.2. Isolation of total RNA from adult *Ornithodoros moubata*

Total RNA isolation of adult, unfed *O. moubata* ticks was performed using RNeasy®minikit (QIAGEN GmbH, Hilden, Germany).

For the isolation of 50µL total RNA, two adult ticks of unknown sex were used. The isolation was done at room temperature. First, the ticks needed to be washed with ethanol to get rid of possible contaminations. This was done by placing the ticks into a 1.5mL centrifuge tube and addition of 1mL 70% ethanol. To dry the ticks, they were placed on a filter paper and then put back again into a dry 1.5mL centrifuge tube.

On top of the ticks, 700µL of buffer RLT were added and subsequently the ticks were grinded. This was done until the lysate turned into a homogeneous solution with a minimum of solid tick residues.

This lysate was then centrifuged at 10000 revolutions per minute (rpm) for three minutes to separate the solid tick residues from the supernatant. Next, the supernatant was transferred to a clean 1.5mL centrifuge tube and 700µL of 70% ethanol solution

were added. The solution was mixed by pipetting up and down. A RNeasy spin column was placed into a collection tube and 700µL of the suspension were added on top of the spin column. Then the collection tube containing the spin column was centrifuged at 10000 rpm for 15 seconds and the flow through was discarded. This was done until the lysate/ethanol mixture was used up, to achieve a good RNA concentration.

Next, 700µL buffer RW1 were added on top of the spin column membrane, followed by centrifugation at 10000 rpm for 15 seconds. The flow-through was discarded again and 500µL of buffer RPE were added. Another centrifugation step was done, this time at 10000 rpm for two minutes, with subsequent discarding of the flow-through.

This was followed by elution of the RNA. Therefore, the spin column was placed in a clean 1.5mL centrifuge tube, 50µL RNase free water were added directly onto the spin column membrane, incubated for one minute at room temperature and then centrifuged at 13000rpm for one minute. Total RNA concentration was measured using NanoDrop 3300 (Thermo Fisher Scientific, Carlsbad, USA). The RNA was stored at -20°C until further use.

5.1.3. First-strand cDNA synthesis

First- strand cDNA synthesis was done using Oligo(dt) primers and SuperScript III Reverse Transcriptase (Invitrogen, Carlsbad, CA, USA). For a reaction volume of 20µL, following components were added to a nuclease-free microcentrifugation tube:

- 1µL of 50µM oligo$(dT)_{20}$
- 500ng total RNA, which was equal to 1µL
- 1µL of 10mM dNTP mix
- 10µL H_2O – sterile and distilled

The solution was mixed by pipetting up and down, incubated at 65°C for five minutes and then placed on ice for one minute.

The tube was briefly centrifuged to collect the contents and following components were added:

- 4µL 5X First-Strand Buffer
- 1µL 0.1M DTT
- 1µL RNaseOUT^{TM} Recombinant RNase Inhibitor
- 1µL SuperScript^{TM} III reverse transcriptase

The solution was mixed by pipetting gently up and down, followed by incubation at 50°C for 60 minutes. The reaction was then inactivated by incubation for 15 minutes at 70°C. The cDNA was then stored at -20°C until use.

5.1.4. Polymerase Chain Reaction (PCR)

The gene coding for DOT1L from the organism *O. moubata* was amplified by PCR using the primer set in table 1.

The PCR reaction was set up with OneTaq® 2X Master Mix (New England Biolabs, Massachusetts, United States), with a primer concentration of 0.0002mM at a final volume of 50µL. The master mix contains the reaction buffer, Taq polymerase and dNTPs at a final reaction concentration of 0.2mM. For amplification 10ng of previously synthesized cDNA were used. The PCR program settings are shown in Table 2 and 40 cycles were used.

Table 2: PCR settings used for amplification of *dot1l* from *O. moubata*

PCR settings			
PCR step	T/°C	Hold time	Cycles
Denaturation	95	2 min	1
Denaturation in cycle	95	30 s	40
Annealing in cycle	57	30 s	
Elongation in cycle	72	1 min	
Final elongation	72	15 min	1
Hold	4	∞	1

5.1.5. Gel electrophoresis

To visualize the amplified PCR products, as well as the products from the restriction digest, gel electrophoresis was performed. For preparation of a 1.2% agarose gel, 1.2g agarose were dissolved in 100mL 1x TAE buffer by heating in a microwave. When the agarose was completely dissolved, the liquid gel was poured into a form, a comb of appropriate size was inserted, and the gel was left to cool down and solidify. When the gel had completely solidified, the comb was pulled out and the gel was put into an electrophoresis tank, which was filled up with 1x TAE buffer.

Each 5µL of the PCR sample were mixed with DNA loading dye (Thermo Fisher Scientific, Carlsbad, USA)**,** which was previously infused with SYBR safe (Invitrogen, Carlsbad, CA, USA**)** for DNA staining**.** This mixture was then loaded into the wells of the gel and Mupid®-One electrophoresis system was used to run the gel. For visualization of the DNA sample sizes, 1kb plus DNA ladder (Invitrogen, Carlsbad, CA, USA**)** was used. After running the gel at 100V, the DNA was visualized under UV light and a picture was taken using alliance Uvitec gel documentation system.

5.1.6. Ligation

The PCR product of desired size was ligated into the pCR™2.1-TOPO® vector. For this purpose, the Invitrogen TOPO TA Cloning® Kit (Thermo Fisher Scientific, Carlsbad, USA) was used following the manufacturers protocol. First, the ligation reaction was prepared on ice as can be seen in Table 3.

Table 3: ligation reaction mixture for PCR product into TOPO TA vector

Reagent	Volume/µL
PCR product	4
Salt Solution	1
TOPO® vector	1
Final volume	6

After mixing the components by gently pipetting up and down, the reaction was incubated for 5 minutes at room temperature. Then the reaction mixture was kept on ice until it was proceeded with transformation of *E. coli* cells.

5.1.7. Cloning

2μL of the TOPO® cloning reaction were added carefully into a vial of One Shot® Chemically Competent *E. coli* (New England Biolabs, Massachusetts, United States). The reaction was then incubated on ice for 30 minutes. Next, the cells were heat-shocked for 30 seconds at 42°C in a water bath without shaking and immediately transferred to ice afterwards. Then, 250μL S.O.C outgrowth medium, (New England Biolabs, Massachusetts, United States), at room temperature, were added and the tube was shaken horizontally for one hour at 37°C.

In the meantime, LBA (lysogenic broth agar) plates containing carbenicillin (Sigma-Aldrich, Prague, Czech Republic) were treated with 40μL X-Gal (Thermo Fisher Scientific, Prague, Czech Republic) and kept in the incubator at 37°C until use.

After shaking the cells for an hour, they were spread on LBA plates with a sterilized inoculation loop. Different volumes of the cell suspension were used. One plate was treated with 70μL cell suspension and another with 140μL. The plates were kept in an incubator at 37°C overnight.

5.1.8. Inoculation of bacterial cultures

Due to the addition of X-Gal to the plates, the bacterial colonies appeared in blue and white colour. Only white colonies were chosen for further multiplication, because only in those, the ligation was successful.

6mL of room temperature LB media was aliquoted into 10 tubes of a 15mL volume. To each tube 3μL ampicillin (100μg/mL) (Sigma-Aldrich, Prague, Czech Republic) were added to achieve a final concentration of 50ng/mL. White colonies were picked with a 10μL pipet tip and the whole tip was put into the media. The tubes were then kept in the horizontal shaker at 37°C and 200rpm overnight.

5.1.9. Plasmid extraction

Plasmid extraction was done using the silica-binding/spin column method of extraction with the NucleoSpin® plasmid Mini prepkit (Macherey-Nagel, Düren, Germany), following manufacturers protocol. All following centrifugation steps were carried out at room temperature. *E. coli* were harvested by centrifuging the 15mL tubes containing the bacterial cultures at 11000 x g for one minute. Then the supernatant was removed, and the cell pellets were resuspended completely in 250μL of Buffer A1 (resuspension buffer) by vortexing. The resuspended cell solution was then transferred to 1.5mL microcentrifuge tubes. 250μL Buffer A2 (Lysis buffer) were added and the solution was gently mixed by inverting the tubes 6-8 times. Then the samples were incubated at room temperature for 5 minutes. 300μL Buffer A3 (Neutralization buffer) were added and mixed thoroughly by again inverting the tubes 6-8 times, until the samples turned colourless. This was followed by centrifugation at 11000 x g. Next, 750μL of the clear supernatant were pipetted onto the NucleoSpin® Plasmid/Plasmid (NoLid) Column, which was placed in a Collection Tube and again centrifuged at 11000 x g, but for one minute. The flow through was discarded and the column placed back in the collection tube. The silica membrane was then washed by addition of 600μL Buffer A4 (Wash buffer) with subsequent centrifugation for one minute at 11000 x g. The flow through was discarded and the columns placed back into the empty collection tubes. They were then centrifuged for two minutes at 11000 x g to dry the membranes. The columns were

placed into 1.5mL microcentrifuge tubes, while the collection tubes were discarded. 50µL Buffer AE (Elution buffer) were added onto the centre of each column, followed by incubation at room temperature for one minute. The procedure was finished by centrifugation for one minute at 11000 x g, thereby collecting the plasmid DNA in the microcentrifuge tube. The plasmid DNA was directly used for restriction enzyme digestion or stored at -20°C for further use.

5.1.10. Restriction Enzyme Digestion

For verification of successful ligation of the *dot1l* gene into the pCR®2.1-TOPO®

Vector, restriction digestion was performed using EcoRI (New England Biolabs, Massachusetts, United States). The reaction mixture was prepared as described in Table 4. Followed by incubation at 37°C for an hour and inactivation of the enzyme at 80°C for 20 minutes.

The products of the digestion were visualized by gel electrophoresis as previously described. Together with the products of the digestion, an undigested plasmid was run on the gel as a control.

Table 4: EcoRI digestion reaction mixture components

Restriction digestion with EcoRI	
Reagent	Volume/µL
EcoRI	0.5
EcoRI buffer	1
H_2O	6.5
Plasmid DNA	2
Final volume	10

5.1.11. Sample preparation for sequencing

In the case a band showed up in electrophoresis of the restriction digest samples, which had the expected size of the amplified *dot1l* gene, the sample was sent for sequencing. Before that, it needed to be prepared. For this purpose, 5µL of the plasmid DNA were pipetted into a 1.5mL microcentrifuge tube. M13 reverse primer, from the TOPO TA Cloning® Kit, was diluted and 5µL of the diluted primer were added to the plasmid DNA to achieve a final concentration of 5pmol μL^{-1}. The samples were then submitted to GATC Biotech Lightrun sequencing service.

5.1.12. Sequence analysis

When retrieving the sequences from GATC Biotech Lightrun sequencing service, we needed confirmation that they are sequences for the *dot1l* gene. Therefore, first the sequences were trimmed to only contain the gene of interest and not the plasmid sequence. Then, the trimmed sequence was aligned to the *dot1l* sequences of *O. rostratus* and *O. turicata* to check for sequence identity. The nucleotide sequence was translated to an amino acid sequence using ExPASy translate tool (Gasteiger *et al.*, 2003).

5.2. Phylogenetic, molecular and structural characterization

A phylogenetic tree was constructed to investigate how conserved the sequence of DOT1L is in *O. moubata*. Furthermore, to get a molecular and structural characterization, secondary structure prediction and protein modelling was done, and the results compared to the well characterized structure of DOT1L in *H. sapiens.*

5.2.1. Phylogenetic tree

To facilitate the construction of a phylogenetic tree, first homologues amino acid sequences to DOT1L from *O. moubata* needed to be collected. This was done by the aid of Blastp tool from BLAST (Boratyn *et al.*, 2012). The organisms involved in the analysis belong to different clades in the domain of *Eukaryota*, involving representative species from different clades, covering taxa from *Vertebrates* to *Arthropoda* as well and *Fungi* and *Amoebozoa* (Table 8). MAFFT 7 alignment server (Kuraku *et al.*, 2013; Katoh, Rozewicki and Yamada, 2017). was used with default settings to align all 32 amino acid sequences to HsDOT1L to confirm nucleotide identity (Figure 10). A phylogenetic tree was then computed by Maximum Likelihood method based on Le Gascuel 2008 model using MEGA V6.0. Initial trees for heuristic search were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using a JTT model. A discrete gamma distribution was applied to model evolutionary rate differences among sites. All positions containing gaps and missing data were eliminated (Le and Gascuel, 2008; Tamura *et al.*, 2013).

5.2.2. Secondary structure prediction

The amino acid sequence of OmDOT1L was submitted to PSIPRED V3.3 (http://bioinf.cs.ucl.ac.uk/psipred/) for secondary structure prediction (Jones, 1999). The resulting secondary structure was then further compared with the secondary structure of HsDOT1L (Min *et al.*, 2003).

5.2.3. Protein modelling

In order to determine the tertiary structure of OmDOT1L, the amino acid sequence was submitted to multiple servers which were I-tasser (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) (Zhang, 2008; Roy, Kucukural and Zhang, 2010; Yang *et al.*, 2015), Phyre 2 (www.sbg.bio.ic.ac.uk/~phyre/) (Kelley *et al.*, 2015), Robetta (robetta.bakerlab.org) (Raman *et al.*, 2009) and Swiss Model (https://swissmodel.expasy.org/) (Arnold *et al.*, 2006; Kopp, 2006; Bordoli *et al.*, 2009; Guex, Peitsch and Schwede, 2009; Kiefer *et al.*, 2009; Biasini *et al.*, 2014; Bienert *et al.*, 2017) In parallel, the target protein sequence was submitted to BLASTp (Boratyn *et al.*, 2012) to perform to PSI- to find closer homologues, using a non-redundant database and five iterations. Then the position specific scoring matrix (PSSM) was obtained and submitted to the Protein databank (https://www.rcsb.org/) to identify homologues with known crystal structures as template for modeling. The “maximum matches in a query range” was set to 10. The closest homologue found was used to perform homologues protein modelling using UCSF Chimera (https://www.cgl.ucsf.edu/chimera/) software program (Pettersen *et al.*, 2004).

The obtained models were then submitted to structure evaluation tools, namely: Qmean (https://swissmodel.expasy.org/qmean/) (Benkert, Tosatto and Schomburg, 2008; Studer, Biasini and Schwede, 2014), RESPROX (https://omictools.com/resprox-tool) (Berjanskii *et al.*, 2012) and ModFOLD 4 (http://www.reading.ac.uk/bioinf/ModFOLD/) (Maghrabi and McGuffin, 2017; McGuffin *et al.*, 2018). The obtained scores were compared between all the models and finally the most probable protein structure was selected for further analysis.

The obtained 3D protein structure was then superimposed with the crystal structure of HsDOT1L in complex with SGC0946 (PDB entry 4ER6)(Yu *et al.*, 2012) using UCSF Chimeras "MatchMaker" option. This allowed comparison of the amino acids and residues that are most likely involved in binding to the inhibitor.

5.3. Functional characterization

To obtain insights of the function of OmDOT1L several experiments were done. First, adult *O. moubata* ticks needed to be separated according to sex. *In vitro* feeding was then conducted on female *O. moubata* ticks. In the case of the inhibition assay the blood that was given to the test group contained an inhibitor for DOT1L,SGC0946. This inhibitor is an analogue of EPZ004777 and shows higher potency and solubility (Yu *et al.*, 2012). Both inhibitors work based on the same principle. They are chemical derivatives of DOT1Ls cofactor S-adenosyl methionine (SAM) and bind to DOT1Ls cofactor binding site, therefore making the enzyme unable to work (Yu *et al.*, 2012). The egg lying behavior as well as the offspring from the ticks were monitored. Then, quantitative real time PCR (qPCR) was conducted to obtain insights about the expression levels of *dot1l* in various life stages of the tick. All these techniques are detailed below.

5.3.1. Sexual dimorphism and gender distinction

Ornithodoros ticks have a sexual dimorphism that can be used to separate them according to sex (Estrada-Peña, 2015a)In particular, the appearance of the genital pore is different in male and females (Figure 6)(Walker, 1994) This was done before on other *Ornithodoros* species, but not on *O. moubata.* On the hand of an electron raster microscope picture we got from Estrada-Peña (Estrada-Peña, 2015a) it was possible to also make a sexual distinction of *O. moubata*. For this purpose, the living adult ticks were placed under a stereoscope to be able to see their ventral surfaces. They were then separated according to their gender based on the pattern of their genital pore.

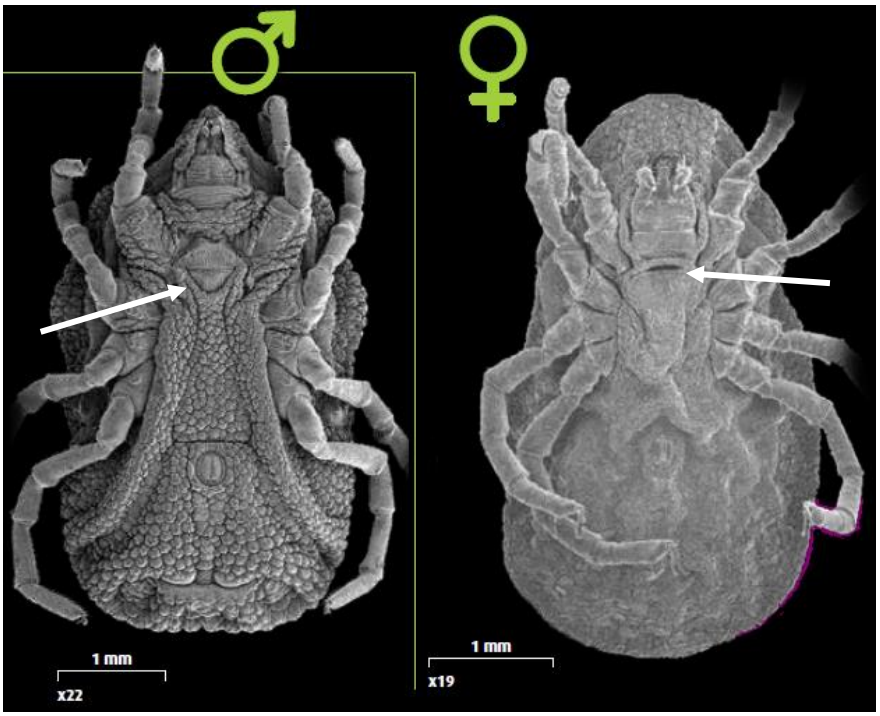

Figure 6: Electron raster microscopy of male and female soft tick respectively;
The genital aperture is marked with an arrow; Figure adapted from: (Estrada-Peña, 2015a)

5.3.2. *In vitro* feeding

In vitro feeding was done to dose adult *O. moubata* with the inhibitor SGC0946, but also to obtain fully fed *O. moubata* (without inhibitor) of all life stages, as samples for qPCR.

The wells of a tissue culture plate were filled with bovine blood, which was infused with gentamicin at a concentration of 5µg mL^{-1} and ATP at a concentration of 1mM, as feeding stimulant. In the case of the inhibition experiment, 50% of the blood samples were infused with SGC0946 (dissolved in DMSO) at a concentration of 20nM. Only DMSO was added to the other 50% as a control.

Then feeding units were placed on top of the wells. The feeding units consisted of plastic cylinders with a feeding membrane at the bottom. The membrane consisted either out of two layered parafilm or out of silicone, while parafilm worked better for the feeding. The plates were then equilibrated in a regular water bath to a temperature of 37°C. Then the ticks were placed in the feeding units until fully fed (Figure 7). The fully fed ticks were then kept in 50mL falcon tubes in a thermo room, at 27.5 °C and 80% humidity for further use.

Figure 7: *in vitro* feeding system, (1) feeding units on top of blood; (2) tissue culture plate; (3) feeding membrane; (4) fully fed, adult *O. moubata*; (5) non fed, adult *O. moubata*

5.3.3. DOT1L inhibition assay on adult, female *O. moubata*

Adult, female *O. moubata* ticks were fed as described above. The ticks were then placed in a thermo room at 27.5°C and 80% humidity**.** They were observed every day to determine the time until oviposition and time until hatching of larvae. By counting the number of eggs that did not hatch as well as living and dead larvae, we determined the percentage of eggs that hatched and the mortality per group. ANOVA statistical test was performed to determine significant differences between the test and control groups (Figure 18). Three biological replicates were accomplished with N = 3 per replicate. Differences were considered significant when $p<0.05$.

5.3.4. Larval immersion test (LIT)

The well establish method for testing arachnicides, the LIT (Klafke *et al.*, 2006; Santos *et al.*, 2013; Webster *et al.*, 2018), was adapted on testing the DOT1L inhibitor SGC0946 on *O. moubata.* Therefore, *O. moubata* larvae were collected and separated into groups of about 60 larvae each. A dilution series of the inhibitor was prepared by dissolving the inhibitor in DMSO. The concentrated solutions were then further diluted in water. 2µL of each inhibitor solution, as well as a control solution which consisted only out of DMSO, were then diluted in 1000µL H_2O, in a 1.5mL microcentrifuge tube, to result in the concentrations shown in Table 5. The larvae groups were immersed in the test and control solutions and incubated at room temperature for 15 minutes, while slightly shaking. After incubation, the larvae groups were taken out and dried on a filter paper. They were then put into 50mL tubes and kept in a thermo room at 27.5°C and

80% humidity for one week. Next, about 50% of the larvae from each group, were fed on mice. When fully fed, the larvae were again kept in the thermo room until completely molted. By counting the number of dead and living nymph the % mortality was determined, for certain concentrations of the inhibitor. One biological replicate was done.

Table 5: concentrations of SGC0946 solutions used for the LIT

Stock solution concentration/mM	final solution concentration/µM
5	10
0.5	1
0.05	0.1

5.3.5. Quantitative real time PCR and statistical analysis

The expression of OmDOT1L was characterized using the total RNA extracted from *O. moubata* eggs, fed and unfed larvae, fed and unfed nymphs and adult males and females. The ticks were fed according to the artificial feeding system described above. For the adult RNA extract each two ticks were crushed. For the nymph RNA extract 6 ticks each were crushed and for the larvae extract about 30. The procedure described in 4.1.2. 'Isolation of total RNA from adult Ornithodoros moubata' was followed. Equal amounts of total RNA were taken for each group of first strand cDNA synthesis, which was done as described before in 4.1.3. First-strand cDNA synthesis**.** Specific primers for RT-qPCR were designed using GenScript Real-time PCR Primer Designing software (https://www.genscript.com/tools/real-time-pcr-tagman-primer-design-tool) (Table 6). Amplification efficiency was verified using serially diluted cDNA samples. Previously reported, ribosomal protein S4 gene was selected as the internal control (Koci, Simo and Park, 2013). The RT-qPCR amplifications conditions can be found in Table 7.

Table 6: Primer pairs used for qPCR

Primer set	Forward primer	Forward primer sequence	Reverse primer	Reverse primer sequence
1	RealAF	TGGAGTTCAGC TCCGACGAC	RealAR	GCAGCCACTTGC AGCACAA
2	Dot1q1F	ATCCCGAACGG CTCAACT	Dot1q1R	AGCCACTTGCAG CACAACCT
3	Dot1q1F	CGATCCCGAAC GGCTCAACT	Dot1q3R	GTCACACAGGGT CTTCATGCTTTC
Reference	rpS4F	GGCCACTGGCA AGATTGACG	rpS4R	GTGTGGCCCTGG GAATCCTT

FastStart Universal SYBR Green Master (Rox) (Sigma-Aldrich, Prague, Czech Republic) and Roche Light Cycler 480 were used to perform the quantitative assay. The preparation of the PCR mix (20 µl) was done according to manufacturers protocol and contained 0.3µM of the primers. Further, melt curve analysis and gel electrophoresis were performed to verify no amplification of spurious or undesirable amplicons. The quantification of mRNA levels of *dot1l* was estimated using Relative Standard Curve

method. Output data were analyzed using the $2^{-\Delta\Delta Ct}$ ratio method (Pfaffl, 2001). Further, ANOVA statistical test was performed to determine significant differences ($p< 0.05$) in the gene expression level between the samples (Figure 19). Gene expression was normalized against *rpS4* gene as previously reported (Koci, Simo and Park, 2013). Three technical replicates were done. Each replicate with the same cDNA sample, but with the different primer sets shown in Table 6 (1 to 3).

Table 7: qPCR reaction conditions used for all primer pairs

qPCR step	Analysis Mode	T/°C	Hold time	Cycles
Activation of FastStart Taq DNA Polymerase	None	95	10 min	1
Amplification and real-time analysis	None	95	15 s	40
	Quantification	60	1 min	

6. Results

6.1. Identification and sequencing of *dot1l*

To amplify *dot1l* from *O. moubata* the first approach we used, was to design degenerate primers. Unfortunately, those primers did not amplify *dot1l.* Therefore, new sets of primers (OmDot1LF and OmDot1LR1, Table 1) were designed, which lead to specific amplification of the gene (Figure 8). The expected product size, according to the template sequences was 759 base pairs. In Figure 7 there is a clear band at approximately this size. Therefore, both PCR product were directly ligated into the pCR®2.1-TOPO® vector (Thermo Fisher Scientific, Carlsbad, USA).

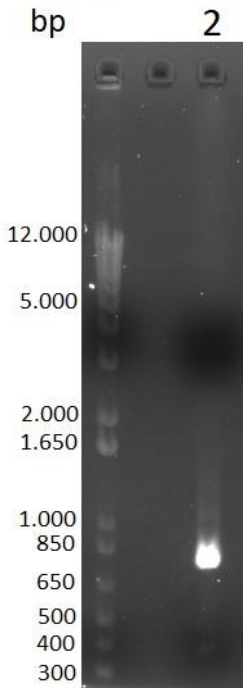

Figure 8: 1.2% Agarose gel; (L) 1Kb plus DNA ladder; (1) amplified *dot1l* sequences from *O. moubata* using primer pair from Table 1

After ligation the plasmid was inserted into competent *E. coli* cells according to '4.1.7. Cloning', for blue white screening. After incubation overnight blue and white colonies appeared. The white colonies were selected and further grown in liquid LB media. The cell cultures were again incubated overnight and on the next day plasmid extraction was performed using mini preps. The extracted plasmid DNA was then further restriction digested using EcoRI restriction enzyme. The desired products of the digestion are the *dot1l* gene and the open-cut TOPO vector. Gel electrophoresis was performed to visualize the presence of the inserted gene and its size (Figure 9).

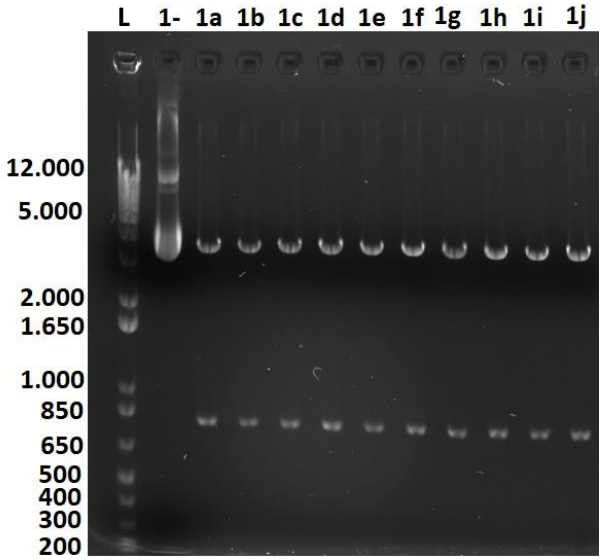

Figure 9: 1.2% Agarose gel showing the products of the restriction digest. 1 indicates for the vectors into which the PCR product amplified with primer pair I was inserted, 1- is the undigested TOPO vectors, (L) 1Kb plus DNA ladder on the very left

The smaller sized bands that can be seen in Figure 9 appear in the same range as the bands visible in Figure 8 which confirms successful ligation and cloning.

Samples which indicated the presence of a recombinant clones were sent for sequencing. The results were compared by the help of BLAST in the NCBI database and it was proven that the amplified product was OmDOT1L.

After *dot1l* was identified in the organism *O. moubata*, its sequence was compared to available *dot1l* sequences from other organisms. The comparison shows a really high identity on the nucleotide level: 90.99% between OmDot1L and OrDot1L; and 87.90% between OmDot1L and OtDot1L (Stothard, 2000). Unfortunately, we were not able to obtain the full length sequence, but only a fragment. Therefore, the score was calculated using trimmed nucleotide sequences of same length. Alignments from OmDot1l with HsDot1l show an identity of 65.90% (Stothard, 2000), which indicates that *dot1l* is widely conserved over a big range of taxa. Another alignment using the nucleotide sequences of *dot1l* from *O. moubata*, *O.rostratus*, *O.turicata*, *Ixodes scapularis* and *Homo sapiens* shows again really high similarities. For the alignment see Figure 10.

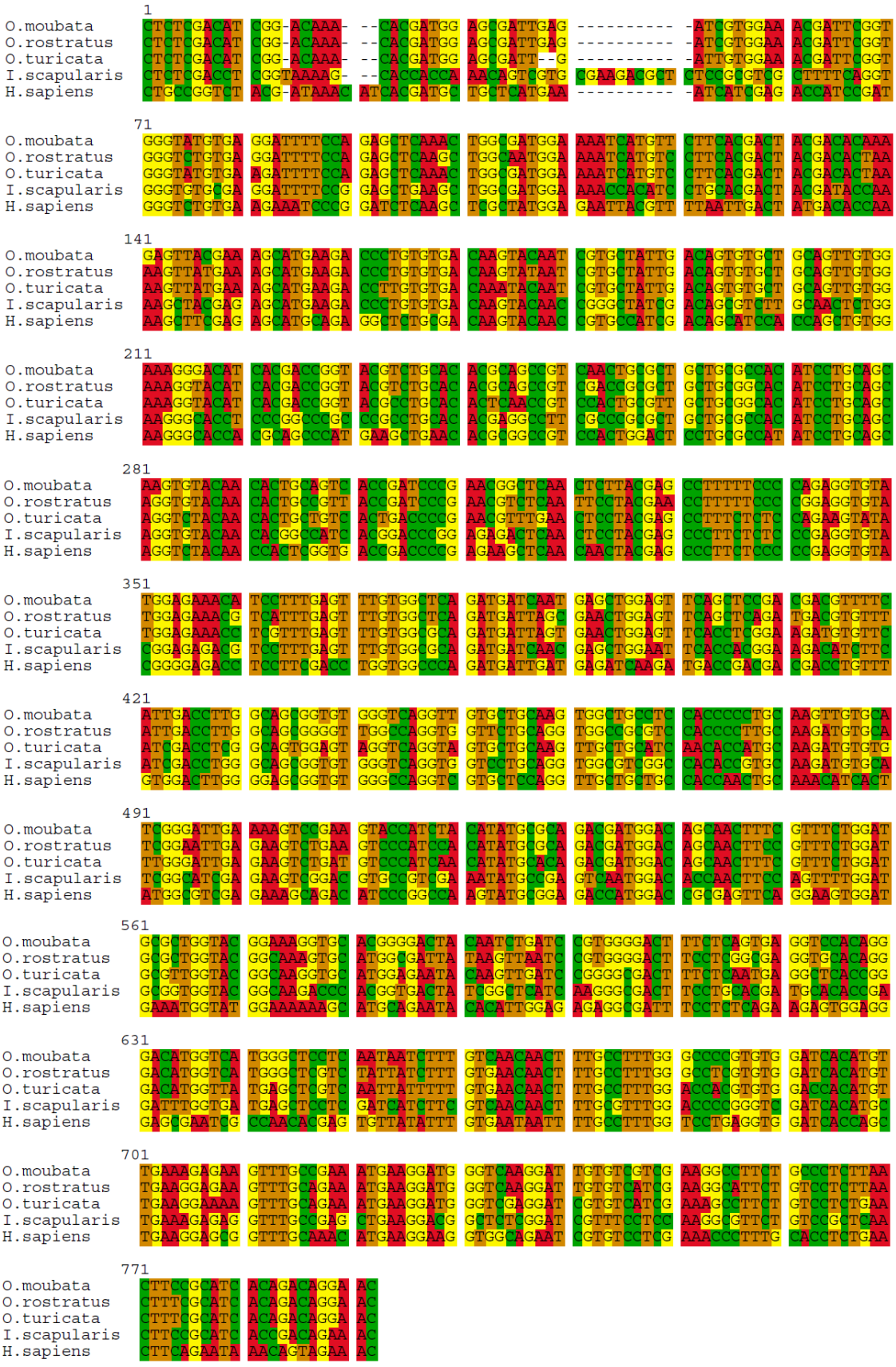

Figure 10: Nucleotide alignment of *dot1l* sequences from the organisms *O. rostratus* (Accession number GCJJ01002031.1), *O. turicata* (GDIE01116948), *Ixodes scapularis* (XM_002403920.1) and *H. sapiens* (XM_005259660.3) to the identified sequence of *O. moubata* (MF431592).

Upon translating the obtained nucleotide sequence to an amino acid sequence (Rice, Longden and Bleasby, 2000; Goujon *et al.*, 2010) and subsequently aligning it to other DOT1L amino acid sequences (Kuraku *et al.*, 2013; Katoh, Rozewicki and Yamada, 2017). One interesting feature can be observed. All amino acid residues that act upon binding to the cofactor S-adenosine-methionine (SAM) are conserved throughout various sequences, including Arthropoda, mammals as well as fungi and bacteria (Figure 11, SAM binding amino acids marked in blue).

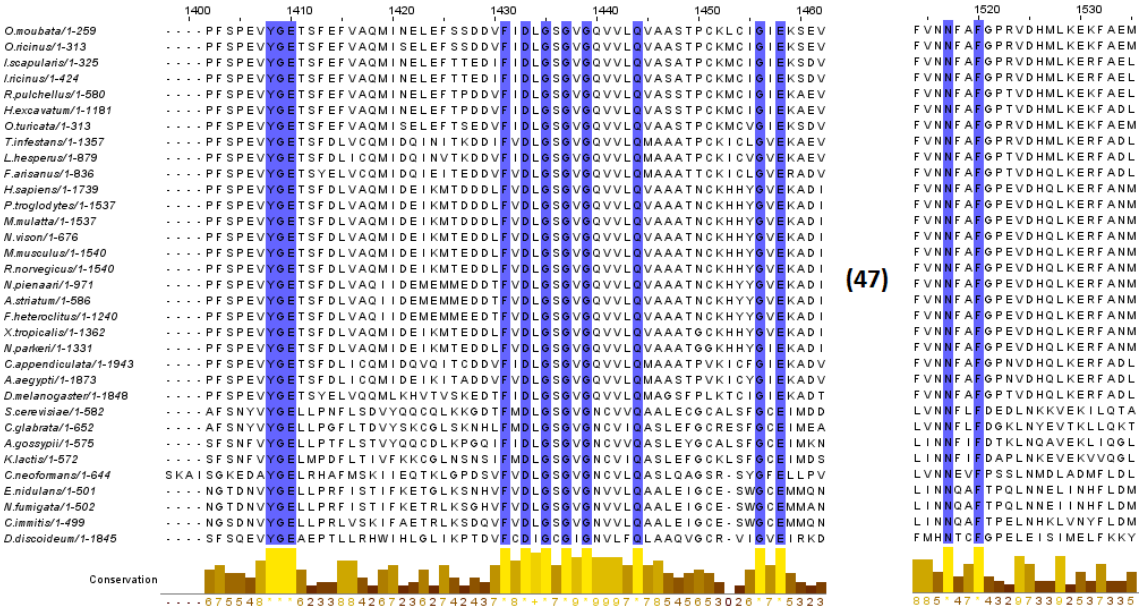

Figure 11: Sequence alignment of DOT1L amino acid sequences from the organisms *O.moubata* (MF431592), *O. rostratus* (GCJJ01002031), *I. scapularis* (XP_002403964), *I. ricinus* (JAB74321), *Rhipicephalus pulchellus* (JAA55304), *Hyalomma excavatum* (JAP65612), *O. turicata* (GDIE01116948), *Triatoma infestans* (JAC15148), *Lygus hesperus* (JAQ01757), *Fopius arisanus* (JAG79228), *H. sapiens* (Q8TEK3), *Pan troglodytes* (JAA03328), *Macaca mulatta* (AFE80476), *Neovison vison* (CCP76587), *Mus Musculus* (NP_955354), *Rattus norvegicus* (XP_006241023), *Nothobranchius pienaari* (SBR38912), *Aphyosemion striatum* (SBP29748), *Fundulus heteroclitus* (JAQ25843), *Xenopus tropicalis* (XP_017952486), *Nanorana parkeri* (XP_018428784), *Corethrella appendiculata* (JAB56398), *Aedes aegypti* (JAN95542), *Drosophila melanogaster* (Q8INR6), *Saccharomyces cerevisiae* (Q04089), *Candida glabrata* (Q6FNM5), *Ashbya gossypii* (Q756E1), *Kluyveromyces lactis* (Q6CWV1), *Cryptococcus neoformans* (P0CN14), *Emericella nidulans* (Q5BH89), *Neosartorya fumigata* (Q4WVH4), *Coccidioides immitis* (Q1DKD8), *Dictyostelium discoideum* (Q55AX2); SAM binding amino acid residues are highlighted in blue; '(47)' indicates the number of amino acids per sequence that are not shown in this alignment

6.2. Phylogenetic, molecular and structural characterization of OmDOT1L

6.2.1. Phylogenetic tree

Based on the high evolutionary similarity of OmDot1L to *dot1l* in other organisms a phylogenetic tree was calculated on the basis of the amino acid sequences from several DOT1L sequence from different species. The sequences were obtained using BLAST

search based on IsDOT1L amino acid sequence. The following phylogenetic tree was obtained (Figure 12). Species from the phylogenetic tree can be found in Figure 11. The tree does not indicate anything unexpected since the sequences cluster in a fashion that would be expected, due to their evolutionary relationships. On the other hand, this clustering could indicate that the enzyme DOT1L evolved parallel in different species.

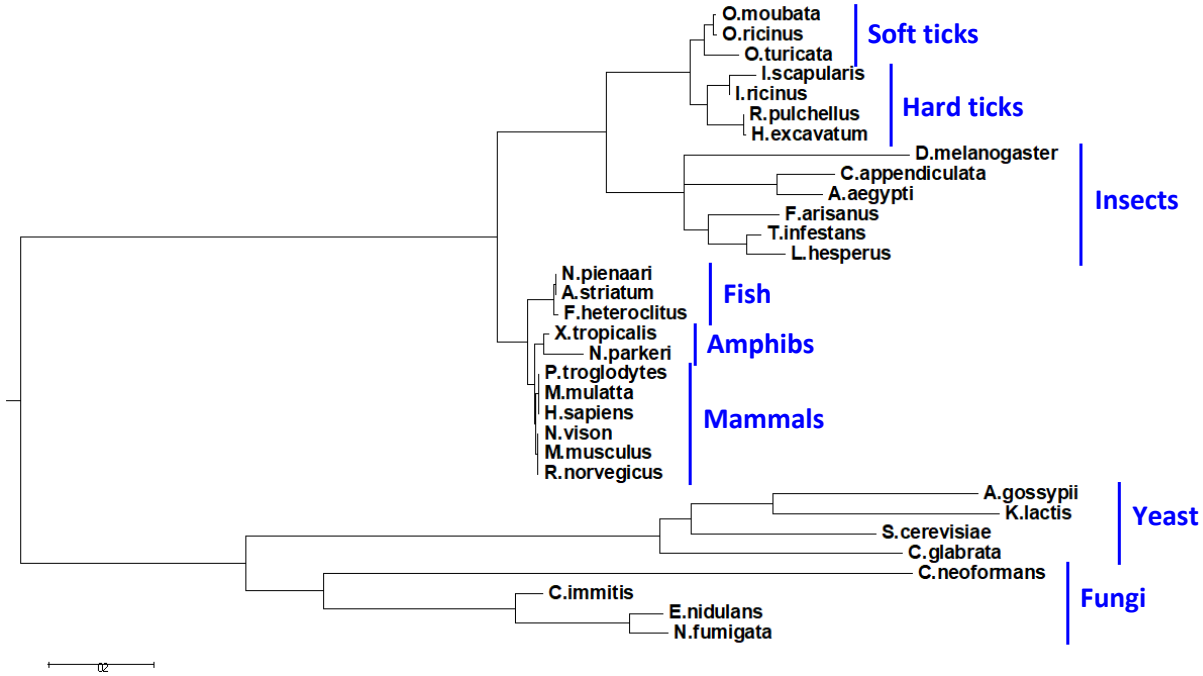

Figure 12: Phylogenetic analysis of 32 DOT1L amino acids sequences (for full species name and Accession number see Figure 11); Computed by Maximum Likelihood method based on Le Gascuel 2008 model using MEGA V6.0; initial trees were obtained automatically by applying Neighbour-Join and BioNJ algorithms to a matrix of pairwise distances estimated using JTT model; a discrete gamma distribution was applied to model evolutionary rate differences among sites; all positions containing gaps and missing data were eliminated

6.2.2. Secondary structure prediction

The amino acid sequence of OmDOT1L was further analyzed concerning its secondary structure. It is known from other organisms that the enzyme possesses conserved secondary structure motifs like the alternating α-helix β-pleated sheet structure (Min *et al.*, 2003). Upon comparison to the secondary structure of HsDOT1L, that was investigated by Jinrong Min et. al. (Min *et al.*, 2003), it can be assumed that also the secondary structure is conserved from human to tick. In Figure 13 the predicted secondary structure from OmDOT1L can be seen (Jones, 1999). Starting from the α-helical structure "C" all secondary structure motifs seem to be conserved in *O. moubata.* First the three α-helical structures C, D and E whereas E is much shorter than the others. Followed by a repeating α-helix β-pleated sheet motif (Figure 13).

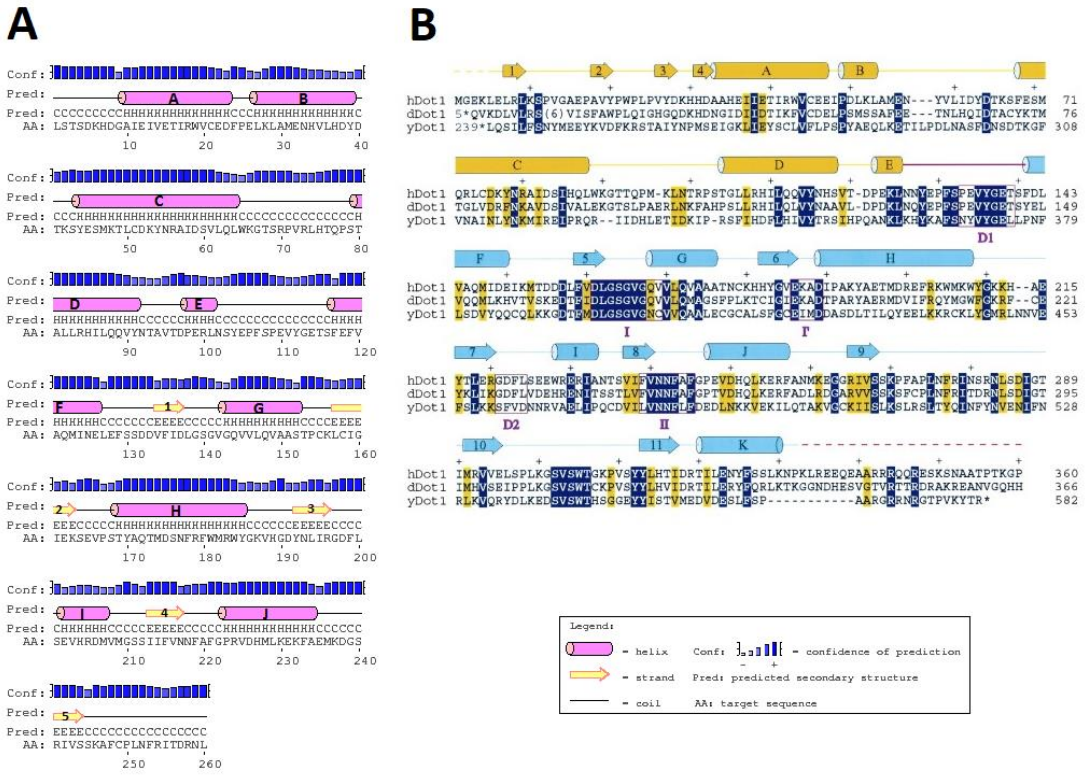

Figure 13: (A) Secondary structure prediction of the amino acid sequence of the catalytic domain of OmDOT1L; (B) Catalytic domain of dot1(L) proteins; Approximately 360 amino acids at the N terminus of HsDot1L show sequence homology with the C-terminal region of yeast dot1 and the N-terminal region of *Drosophila* homolog. Identical amino acids are shown in white letters over blue background, similar residues are highlighted in yellow. Purple rectangles show elements that interact with SAM; secondary structure elements are shown above the sequence; Figure adapted from (Min et al. 2003)

6.2.3. Protein homology modeling of OmDOT1L

Based on the high similarities of HsDot1 and OmDOT1L further investigations were done on the level of tertiary structures. Therefore, several tools were used to gain a 3D model of DOT1L from *O. moubata* (4.2.3. Protein modelling). The obtained 3D structure was then further superimposed to the crystal structure of DOT1 of *H. sapiens* in complex with the inhibitor SGC0946 (PDB entry 4ER6) using UCSF chimera (Pettersen *et al.*, 2004). Therefore, first the structure of *H. sapiens* was cropped to have the same length as the one of *O. moubata.* We used the crystal structure of DOT1 in complex with this inhibitor, because SGC0946 was also used in further inhibition assays. The amino acid residues from HsDot1 that are responsible for binding, according to a 2D diagram showing the possible interactions between HsDot1 and SGC0946 (Figure 14), were highlighted and compared to those in OmDOT1L (Figure 16). Our model suggests that all except for one amino acid residue are conserved, indicating again high similarity, not only based on the primary and secondary structures but also on tertiary structures. Therefore, functional conservation could be hypothesized.

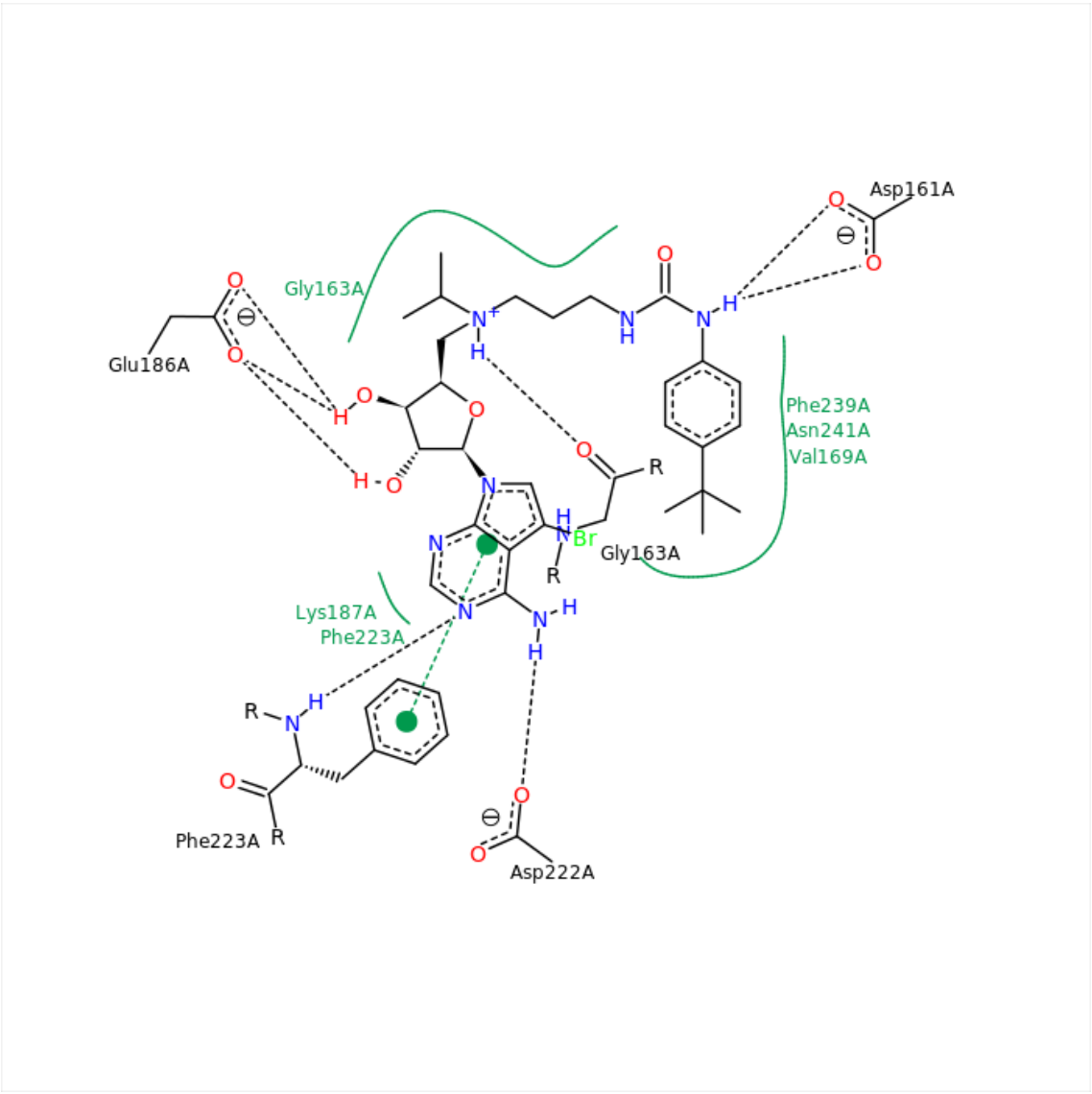

Figure 14: The amino acid residues of HsDot1 that interact with SGC0946 (5-bromo-7-{5-[(3-{[(4-tert-butylphenyl)carbamoyl]amino}propyl)(propan-2-yl)amino]-5-deoxy-beta-D-ribofuranosyl}-7H-pyrrolo[2,3-d]pyrimidin-4-amine) are shown in a 2D diagram; black dashed lines indicate hydrogen bonds, salt bridges and metal interactions; green solid lines show hydrophobic interactions; green dashed lines show π-π and π-cation interactions; Figure taken from (Yu et al. 2012)

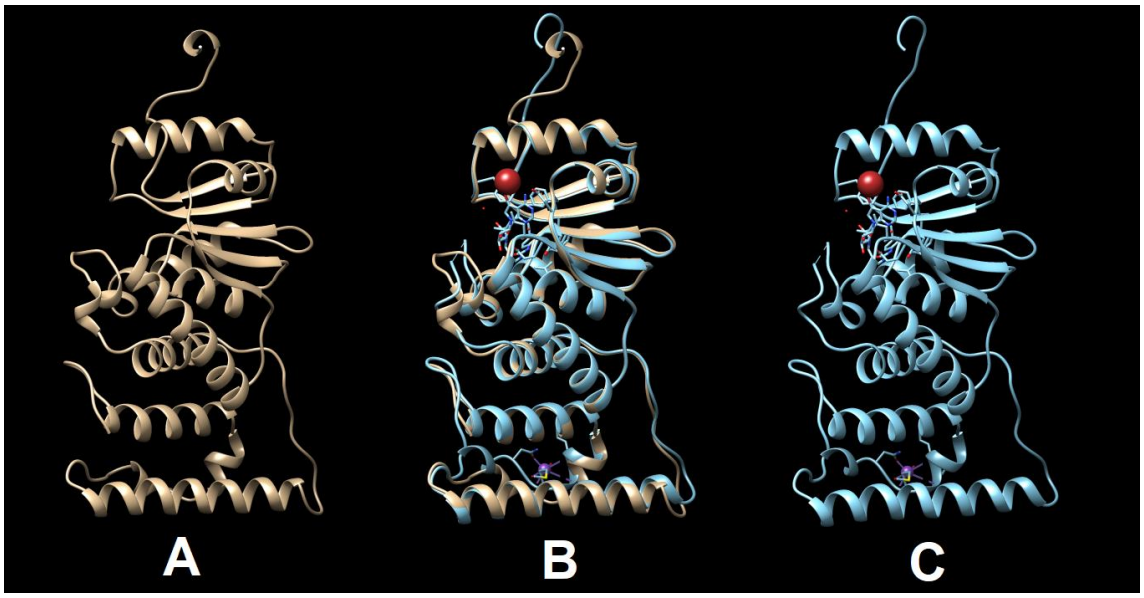

Figure 15: (A) Predicted 3D protein structure of active DOT1L domain of *O. moubata* (beige); (B) Superimposed structures of OmDOT1L (beige) and HsDOT1 (blue) (PDB entry: 4ER6) in complex with SGC0946; (C) Crystal structure of HsDot1 in complex with its inhibitor SGC0946 (PDB entry: 4ER6); the ribbon of HsDOT1L was cut to have the same length as OmDOT1L (this applies to B and C)

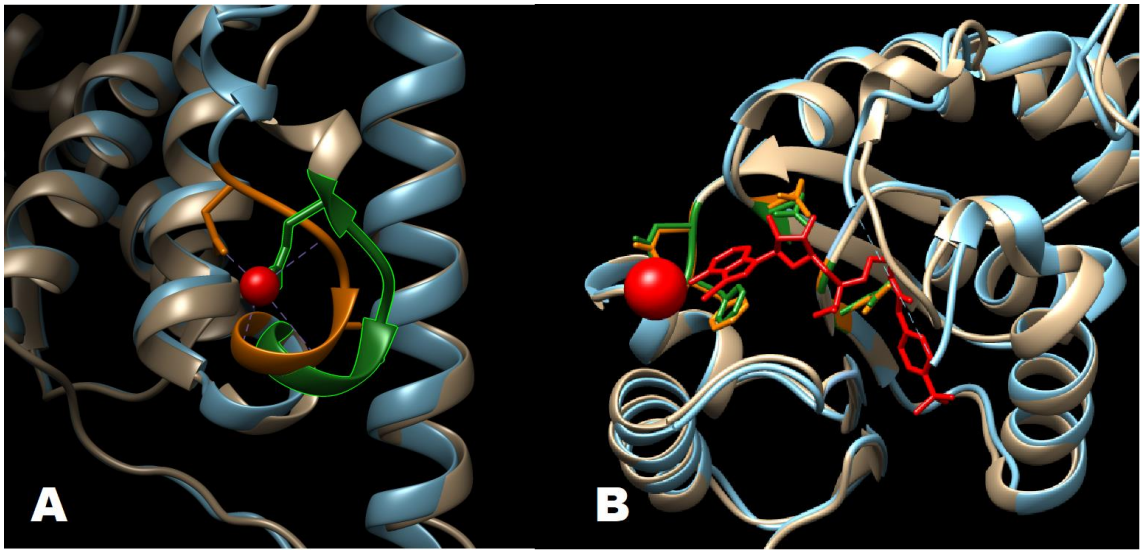

Figure 16: Superimposed HsDot1 (blue) in complex with SGC0946 (red) and OmDOT1L (beige); amino acids from OmDOT1L involved in binding to Na^+ ion (green); amino acids from HsDot1 involved in binding to Na^+ ion (orange)

6.3. Functional characterization of OmDOT1L

6.3.1. Sexual dimorphism and sex distinction in *O. moubata*

Adult *O. moubata* were separated according to their gender. The genital aperture (Figure 3) was investigated using a stereoscope and compared to Figure 6. Thereby we were able to separate the ticks depending on their sex. As it can be seen in Figure 17 the main difference between male and female *O. moubata* remains to be the genital pore. In the case of a female tick the lower part of the genital pore seems to grow over the upper

part forming a little hill. In contrast to females, the genital pore of the male ticks forms something that can be described as a pocket, where the upper part kind of folds into the lower one.

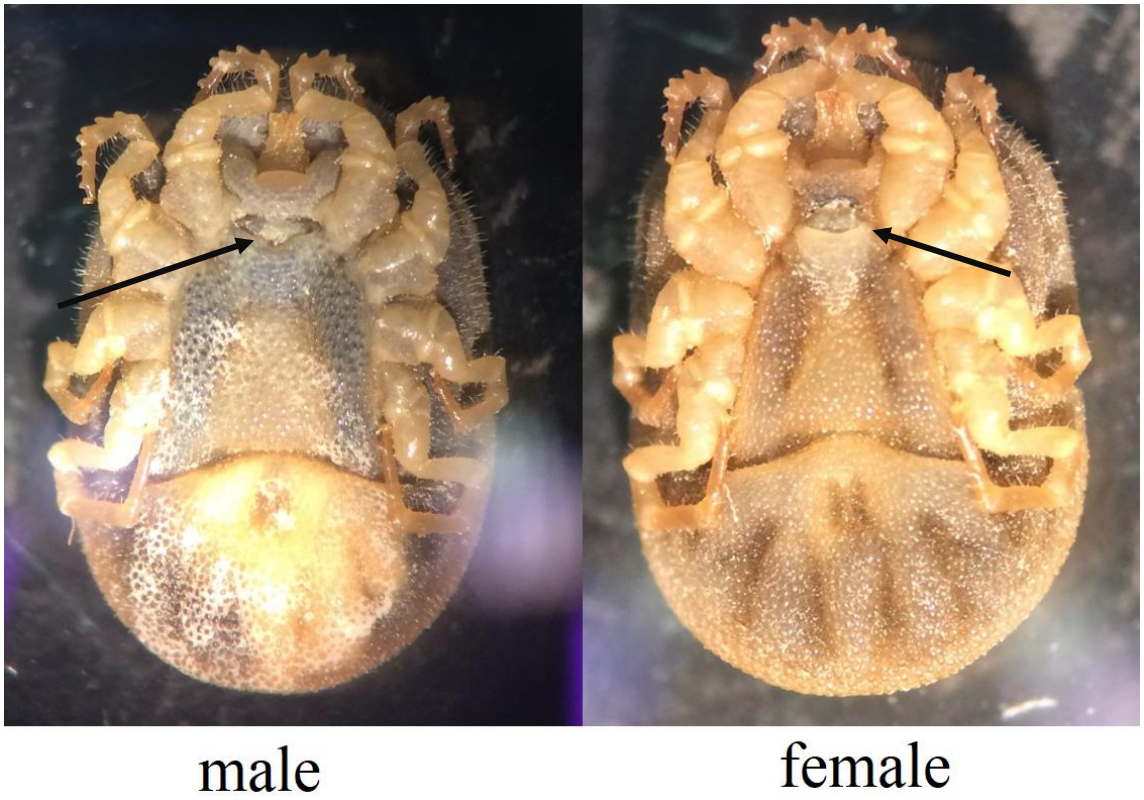

Figure 17: Picture taken through stereoscope, showing ventral surfaces of male *O. moubata* (left) and Ventral surface of female *O. moubata* (right); arrows are marking the genital pores

6.3.2. DOT1L inhibition assay on adult, female *O. moubata*

Adult female *O. moubata* were artificially fed (according to 4.3.2. *In vitro* feeding). After fully fed, the ticks were kept in a thermo room and observed every day for egg laying and hatching of the eggs. In the first biological replicate there was a significant difference between the number of eggs laid and vital larvae from a female treated with the inhibitor versus a female that was not treated with the inhibitor. Unfortunately, those results could not be reproduced in the two other biological replicates. The statistical analysis, from all three biological replicates, of the treated vs the untreated group can be found in Figure 18. These results do not indicate any effect of DOT1L on the embryonic development of the ticks, but further studies and repetitions of the experiment should be conducted to assure the results. The inhibitor we used is known to have a high efficiency but nevertheless it is highly sensitive (Yu *et al.*, 2012). Therefore it is possible that degradation of the inhibitor lead to the fact that the experiment could not be reproduced.

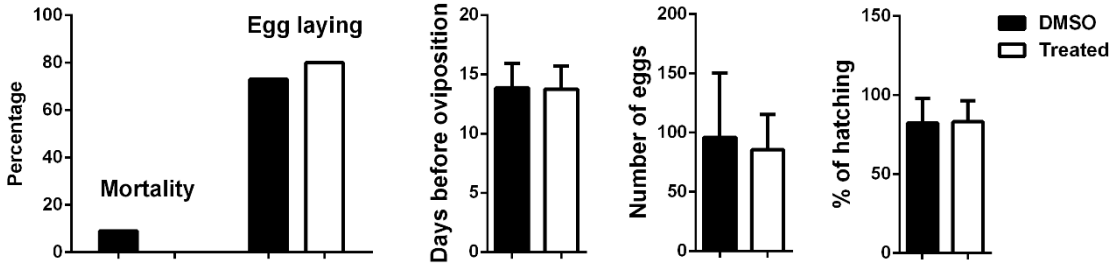

Figure 18: Statistical analysis of the inhibition assay on adult female *O. moubata*; "treated" indicates that the blood, the ticks were fed on, was infused with the inhibitor of DOT1L SGC0946 at a concentration of 20nM; The results indicate no significant difference in any of the parameters that were observed

6.3.4. Quantitative real time PCR

Quantitative real time PCR (qPCR) was conducted to measure the expression levels of *dot1l* in various life stages of fed and unfed *O. moubata* (4.3.5. Quantitative real time PCR and statistical analysis). Unfortunately, at that time we were not able to obtain fed adults, therefore we could only determine the expression of *dot1l* of unfed adults. Nevertheless, the expression levels undergo drastic changes in two cases. The first significant change is upregulation of the mRNA expression after fed larvae molt to nymphs. The expression of *dot1l* mRNA undergoes a second significant change when the nymphs are fully engorged (Figure 19). The upregulation of *dot1l* in unfed nymphs compared to unfed and fed larvae suggests that this enzyme may have a role during tick molting. To test this hypothesis, unfed larvae were treated with the DOT1L inhibitor SGC0946 in LIT assay.

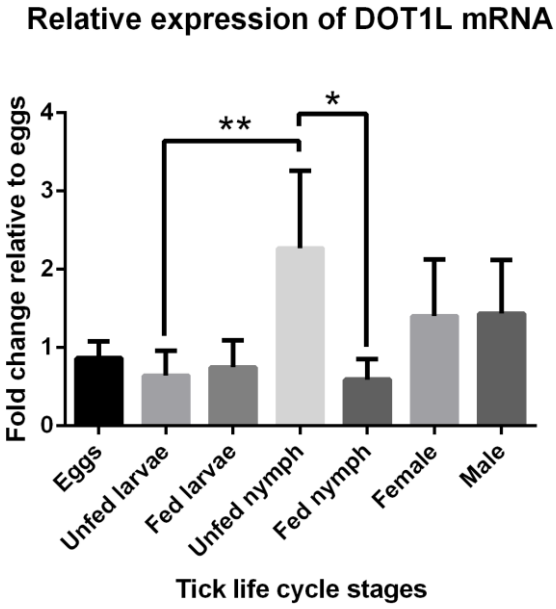

Figure 19: relative mRNA expression of *dot1l*, scaled so that the expression of *dot1l* in eggs equals 1 and expression in other life stages are represented in fold changes relative to *dot1l* mRNA expression in eggs

6.3.3. Larval Immersion Test (LIT)

As the name suggest, this assay was conducted with larvae. The LIT was adapted according to other LIT protocols done with acaricides (Klafke *et al.*, 2006; Santos *et al.*, 2013; Webster *et al.*, 2018). After incubating the larvae in the test solutions they were kept in a thermo room for a week and fed on mice afterwards. Before feeding of the larvae no mortality was observable in any of the test groups. The percentage of mortality was determined after molting to nymphs, and it was found that mortality in the control group, without inhibitor, was about 6.9%. At an inhibitor (SGC0946) concentration of 0.1µM the mortality does not change significantly (7.7%). But at an inhibitor concentration of 1µM and above, there is a significant enhancement in mortality, namely 39.1% mortality at a SGC0946 concentration of 1µM and 26.7% mortality at a concentration of 10µM. Therefore, a lower proportion of larvae molt to nymphs when treated with increasing concentration of the inhibitor. GraphPad program was used to construct Figure 20 and to perform unpaired non-parametric Mann Whitney's test. Differences were considered significant when $p<0.05$. GraphPad program was used to make figure and perform Chi-squared ($\chi 2$) test ($\chi 2 = 50.21$, $p<0.0001$).

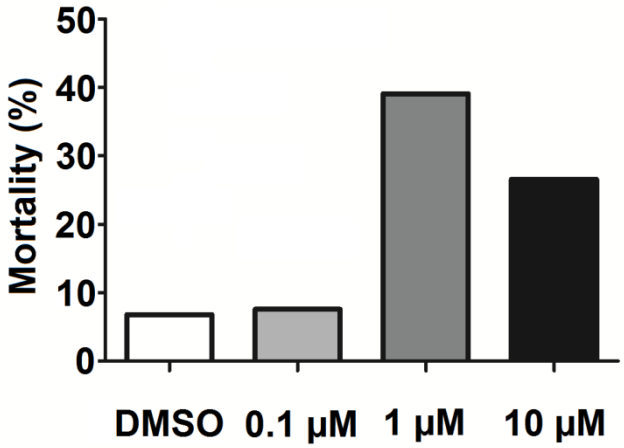

Figure 20: results of the LIT showing percent mortality; larvae were immersed in solutions containing the inhibitor SGC0946 incubated for 15 minutes, then dried and incubated for a week in a thermoroom; after incubation they were fed on mice and let to molt; percent mortality were calculated from the freshly molted nymphs

Nevertheless, only one biological replicate was done so far, so this experimental setup has to be repeated before a premature conclusion is taken.

7. Discussion

Dot1 was initially discovered in *Saccharomyces cerevisiae* as disruptor of telomeric silencing (Singer *et al.*, 1998; Jones *et al.*, 2008). Subsequently, the dot1 homolog gene, dot1-like (*dot1l*) has been found in a broad range of species, spanning from Arthropods to protozoa, and mammals (Janzen *et al.*, 2006; Jones *et al.*, 2008; List *et al.*, 2009; Wong, Polly and Liu, 2015; Cabezas-Cruz *et al.*, 2016). With this study we successfully confirmed the presence of DOT1L in the soft tick species *Ornithodoros moubata.* With the obtainment of its sequence we were able to show the high conservation of this enzyme through various clades. On the nucleotide level *dot1l* shows remarkable 65.9% sequence identity between *H. sapiens* and *O. moubata.* This and the phylogenetic analysis show that *dot1l* is widely conserved on a high level of sequence identity. When focusing just on the amino acid residues that are responsible for the binding to the cofactor SAM, we can conclude from our analysis, for the species that were involved, that all of those residues are conserved. The phylogenetic tree (Figure 12) also shows that amino acid sequences of DOT1L are highly conserved over a broad range of taxa.

When focusing more on structural details, like secondary and tertiary protein structure, one can also conclude high conservation. On the hand of the secondary structure prediction one can see that the pattern, in which α-helix and β-pleated sheet structures alternate, is also conserved from *O. moubata* to *H. sapiens* (Min *et al.*, 2003). Which makes functional conservation more probable than just the sequence identity.

On the hand of our 3D protein structure prediction this hypothesis gets even more compelling. Even though the computed 3D model from OmDOT1L shows deviations to the crystal structure of HsDOT1 (PDB entry: 4ER6) at the protein backbone, the amino acid residues necessary for binding to SGC0946 are conserved. Therefore, not only the residues responsible for the binding of the cofactor SAM are conserved (primary structure), but also the residues responsible for binding to an inhibitor.

On the hand of the LIT experiment and the results from the conducted qPCR, we can hypnotize that the ticks are in need of upregulating *dot1l* to be able to successfully molt to its next life stage. These results and the one from the inhibition assay (5.3.2.) would coincide with the results from Luan Wen et al. (Wen *et al.*, 2015). In his experiment they knocked down endogenous DOT1L in fertilized eggs from *Xenopus tropicalis.* This resulted in successful hatching of tadpoles, but those tadpoles could not further evolve to viable toads. The knockdown leads to a severe retarded growth of the tadpoles and led to lethality prior metamorphosis. This reflects somehow our results. There was no significant difference in the number of eggs laid from female ticks treated with an inhibitor for DOT1L and there was neither a significant difference in the egg hatching. But when larvae were treated with the same inhibitor, fed and let to molt, there was a significant different concerning lethality. This molting step could be somehow synonymous to the metamorphosis step of the tadpoles of *X. tropicalis.*

8. Conclusion

I. The N-terminus of *dot1l* from *O. moubata* was successfully amplified but attempts to get the full sequence failed. The *dot1l* sequence is well-conserved over a broad range of taxa as expected

II. From the conducted structural analysis we can conclude that the N-terminus bears the catalytic domain of DOT1L

III. OmDOT1L probably plays an important role in tick molting

9. References

Arnold, K. *et al.* (2006) 'The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling.', *Bioinformatics (Oxford, England)*, 22(2), pp. 195–201. doi: 10.1093/bioinformatics/bti770.

Bannister, A. J. and Kouzarides, T. (2011) 'Regulation of chromatin by histone modifications', *Cell Research*. Nature Publishing Group, 21(3), pp. 381–395. doi: 10.1038/cr.2011.22.

Barski, A. *et al.* (2007) 'High-resolution profiling of histone methylations in the human genome.', *Cell*. Elsevier, 129(4), pp. 823–37. doi: 10.1016/j.cell.2007.05.009.

Benkert, P., Tosatto, S. C. E. and Schomburg, D. (2008) 'QMEAN: A comprehensive scoring function for model quality assessment.', *Proteins*, 71(1), pp. 261–77. doi: 10.1002/prot.21715.

Berjanskii, M. *et al.* (2012) 'Resolution-by-proxy: a simple measure for assessing and comparing the overall quality of NMR protein structures.', *Journal of biomolecular NMR*, 53(3), pp. 167–80. doi: 10.1007/s10858-012-9637-2.

Biasini, M. *et al.* (2014) 'SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information.', *Nucleic acids research*, 42(Web Server issue), pp. W252-8. doi: 10.1093/nar/gku340.

Bienert, S. *et al.* (2017) 'The SWISS-MODEL Repository—new features and functionality', *Nucleic Acids Research*, 45(D1), pp. D313–D319. doi: 10.1093/nar/gkw1132.

Boratyn, G. M. *et al.* (2012) 'Domain enhanced lookup time accelerated BLAST', *Biology Direct*, 7(1), p. 12. doi: 10.1186/1745-6150-7-12.

Bordoli, L. *et al.* (2009) 'Protein structure homology modeling using SWISS-MODEL workspace.', *Nature protocols*, 4(1), pp. 1–13. doi: 10.1038/nprot.2008.197.

Cabezas-Cruz, A. *et al.* (2016) 'Anaplasma phagocytophilum increases the levels of histone modifying enzymes to inhibit cell apoptosis and facilitate pathogen infection in the tick vector Ixodes scapularis', *Epigenetics*, 11(4), pp. 303–319. doi: 10.1080/15592294.2016.1163460.

Cheng, X., Collins, R. E. and Zhang, X. (2005) 'Structural and Sequence Motifs of Protein (Histone) Methylation Enzymes', *Annual Review of Biophysics and Biomolecular Structure*, 34, pp. 267–294. doi: 10.1146/annurev.biophys.34.040204.144452.

Denyer, M. S. and Wilkinson, P. J. (1998) 'African Swine Fever', in *Encyclopedia of Immunology*. Elsevier, pp. 54–56. doi: 10.1006/rwei.1999.0015.

Estrada-Peña, A. (2015a) *Garrapatas - Morfología, fisiología y ecología*.

Estrada-Peña, A. (2015b) 'Ticks as vectors : taxonomy , biology and ecology', *Revue scientifique et technique (International Office of Epizootics)*, 34(1), pp. 53–65. doi: 10.20506/rst.34.1.2345.

Feng, Q. *et al.* (2002) 'Methylation of H3-lysine 79 is mediated by a new family of HMTases without a SET domain', *Current Biology*, 12(12), pp. 1052–1058. doi: 10.1016/S0960-9822(02)00901-6.

Gasteiger, E. *et al.* (2003) 'ExPASy: The proteomics server for in-depth protein knowledge and analysis.', *Nucleic acids research*, 31(13), pp. 3784–8. Available at: http://www.ncbi.nlm.nih.gov/pubmed/12824418 (Accessed: 13 July 2018).

Gómez-Díaz, E. *et al.* (2012) 'Epigenetics of Host-Pathogen Interactions: The Road Ahead and the Road Behind', *PLoS Pathogens*, 8(11). doi:

10.1371/journal.ppat.1003007.

Goujon, M. *et al.* (2010) 'A new bioinformatics analysis tools framework at EMBL-EBI', *Nucleic Acids Research*, 38(Web Server), pp. W695–W699. doi: 10.1093/nar/gkq313.

Guex, N., Peitsch, M. C. and Schwede, T. (2009) 'Automated comparative protein structure modeling with SWISS-MODEL and Swiss-PdbViewer: a historical perspective.', *Electrophoresis*, 30(1), pp. 162–173. doi: 10.1002/elps.200900140.

Guglielmone, A. A. *et al.* (2010) 'The Argasidae , Ixodidae and Nuttalliellidae (Acari : Ixodida): a world list of valid tick names', *Zootaxa*, 2528, pp. 1–28.

Janzen, C. J. *et al.* (2006) 'Selective Di- or Trimethylation of Histone H3 Lysine 76 by Two DOT1 Homologs Is Important for Cell Cycle Regulation in Trypanosoma brucei', *Molecular Cell*, 23(4), pp. 497–507. doi: 10.1016/j.molcel.2006.06.027.

Jenuwein, T. *et al.* (1998) 'SET domain proteins modulate chromatin domains in eu- and heterochromatin.', *Cellular and molecular life sciences : CMLS*, 54(1), pp. 80–93.

Jenuwein, T. (2001) 'Translating the Histone Code', *Science*, 293(5532), pp. 1074–1080. doi: 10.1126/science.1063127.

Jones, B. *et al.* (2008) 'The histone H3K79 methyltransferase Dot1L is essential for mammalian development and heterochromatin structure', *PLoS Genetics*, 4(9). doi: 10.1371/journal.pgen.1000190.

Jones, D. T. (1999) 'Protein secondary structure prediction based on position-specific scoring matrices.', *Journal of molecular biology*, 292(2), pp. 195–202. doi: 10.1006/jmbi.1999.3091.

Katoh, K., Rozewicki, J. and Yamada, K. D. (2017) 'MAFFT online service: multiple sequence alignment, interactive sequence choice and visualization', *Briefings in Bioinformatics*. doi: 10.1093/bib/bbx108.

Kelley, L. A. *et al.* (2015) 'The Phyre2 web portal for protein modeling, prediction and analysis.', *Nature protocols*, 10(6), pp. 845–58. doi: 10.1038/nprot.2015.053.

Kiefer, F. *et al.* (2009) 'The SWISS-MODEL Repository and associated resources', *Nucleic acids research*, 37, pp. 387–392. doi: 10.1093/nar/gkn750.

Klafke, G. M. *et al.* (2006) 'Larval immersion tests with ivermectin in populations of the cattle tick Rhipicephalus (Boophilus) microplus (Acari: Ixodidae) from State of Sao Paulo, Brazil', *Veterinary Parasitology*, 142(3–4), pp. 386–390. doi: 10.1016/j.vetpar.2006.07.001.

Koci, J., Simo, L. and Park, Y. (2013) 'Validation of internal reference genes for real-time quantitative polymerase chain reaction studies in the tick, Ixodes scapularis (Acari: Ixodidae).', *Journal of medical entomology*, 50(1), pp. 79–84. Available at: http://www.ncbi.nlm.nih.gov/pubmed/23427655 (Accessed: 14 July 2018).

Kopp, J. (2006) 'The SWISS-MODEL Repository: new features and functionalities', *Nucleic Acids Research*, 34, pp. 315–318. doi: 10.1093/nar/gkj056.

Kouzarides, T. (2007) 'Chromatin Modifications and Their Function', *Cell*, 128(4), pp. 693–705. doi: 10.1016/j.cell.2007.02.005.

Kuraku, S. *et al.* (2013) 'aLeaves facilitates on-demand exploration of metazoan gene family trees on MAFFT sequence alignment server with enhanced interactivity', *Nucleic Acids Research*, 41(W1), pp. 22–28. doi: 10.1093/nar/gkt389.

de la Fuente, J. *et al.* (2008) 'Overview: Ticks as vectors of pathogens that cause disease in humans and animals.', *Frontiers in bioscience : a journal and virtual library*, 13, pp. 6938–46. Available at: http://www.ncbi.nlm.nih.gov/pubmed/18508706 (Accessed: 7 July 2018).

de la Fuente, J., Estrada-Peña, A., *et al.* (2016) 'Anaplasma phagocytophilum Uses Common Strategies for Infection of Ticks and Vertebrate Hosts', *Trends in Microbiology*, 24(3), pp. 173–180. doi: 10.1016/j.tim.2015.12.001.

de la Fuente, J., Villar, M., *et al.* (2016) 'Tick–Host–Pathogen Interactions: Conflict and Cooperation', *PLoS Pathogens*, 12(4), pp. 1–7. doi: 10.1371/journal.ppat.1005488.

Le, S. Q. and Gascuel, O. (2008) 'An Improved General Amino Acid Replacement Matrix', *Molecular Biology and Evolution*, 25(7), pp. 1307–1320. doi: 10.1093/molbev/msn067.

van Leeuwen, F., Gafken, P. R. and Gottschling, D. E. (2002) 'Dot1p modulates silencing in yeast by methylation of the nucleosome core', *Cell*, 109(6), pp. 745–756. doi: 10.1016/S0092-8674(02)00759-6.

List, O. *et al.* (2009) 'Overexpression of *grappa* encoding a histone methyltransferase enhances stress resistance in *Drosophila*', *Hereditas*, 146(1), pp. 19–28. doi: 10.1111/j.1601-5223.2008.02080.x.

Maghrabi, A. H. A. and McGuffin, L. J. (2017) 'ModFOLD6: an accurate web server for the global and local quality estimation of 3D protein models', *Nucleic Acids Research*, 45(W1), pp. W416–W421. doi: 10.1093/nar/gkx332.

Manzano-Romn, R. *et al.* (2012) 'Soft Ticks as Pathogen Vectors: Distribution, Surveillance and Control', *Parasitology*. doi: 10.5772/32521.

McGuffin, L. J. *et al.* (2018) 'Accurate template-based modeling in CASP12 using the IntFOLD4-TS, ModFOLD6, and ReFOLD methods', *Proteins: Structure, Function, and Bioinformatics*, 86(1), pp. 335–344. doi: 10.1002/prot.25360.

Min, J. *et al.* (2003) 'Structure of the Catalytic Domain of Human DOT1L , a Non-SET Domain Nucleosomal Histone Methyltransferase', *Cell*, 112(5), pp. 711–723.

Nguyen, A. T. and Zhang, Y. (2011) 'The diverse functions of Dot1 and H3K79 methylation.', *Genes & development*. Cold Spring Harbor Laboratory Press, 25(13), pp. 1345–58. doi: 10.1101/gad.2057811.

Ntiamoa-Baidu, Y. *et al.* (2004) 'An updated list of the ticks of Ghana and an assessment of the distribution of the ticks of Ghanaian wild mammals in different vegetation zones', *Bulletin of Entomological Research*, 94(03), pp. 245–60. doi: 10.1079/BER2004302.

Ooga, M. *et al.* (2008) 'Changes in H3K79 Methylation During Preimplantation Development in Mice', *Biology of Reproduction*, 78(3), pp. 413–424. doi: 10.1095/biolreprod.107.063453.

Parola, P. and Raoult, D. (2001) 'Ticks and Tickborne Bacterial Diseases in Humans: An Emerging Infectious Threat', *Clinical Infectious Diseases*, 32(6), pp. 897–928. doi: 10.1086/319347.

Peter, R. J. *et al.* (2005) 'Tick, fly, and mosquito control—Lessons from the past, solutions for the future', *Veterinary Parasitology*. Elsevier, 132(3–4), pp. 205–215. doi: 10.1016/J.VETPAR.2005.07.004.

Pettersen, E. F. *et al.* (2004) 'UCSF Chimera?A visualization system for exploratory research and analysis', *Journal of Computational Chemistry*, 25(13), pp. 1605–1612. doi: 10.1002/jcc.20084.

Pfaffl, M. W. (2001) 'A new mathematical model for relative quantification in real-time RT-PCR.', *Nucleic acids research*, 29(9), p. 45. Available at: http://www.ncbi.nlm.nih.gov/pubmed/11328886 (Accessed: 1 August 2018).

Raman, S. *et al.* (2009) 'Structure prediction for CASP8 with all-atom refinement using Rosetta', *Proteins: Structure, Function, and Bioinformatics*, 77(S9), pp. 89–99. doi:

10.1002/prot.22540.

Rice, K. L., Hormaeche, I. and Licht, J. D. (2007) ‘Epigenetic regulation of normal and malignant hematopoiesis’, *Oncogene*. Nature Publishing Group, 26(47), pp. 6697–6714. doi: 10.1038/sj.onc.1210755.

Rice, P., Longden, I. and Bleasby, A. (2000) ‘EMBOSS: The European Molecular Biology Open Software Suite’, *Trends in Genetics*, 16(6), pp. 276–277. doi: 10.1016/S0168-9525(00)02024-2.

Roy, A., Kucukural, A. and Zhang, Y. (2010) ‘I-TASSER: a unified platform for automated protein structure and function prediction’, *Nature Protocols*, 5(4), pp. 725–738. doi: 10.1038/nprot.2010.5.

Santos, T. R. B. *et al.* (2013) ‘Comparison of three larval bioassays to evaluate susceptibility of Rhipicephalus (Boophilus) microplus to amitraz.’, *Brazilian journal of veterinary parasitology*, 22(4), pp. 495–501. doi: 10.1590/S1984-29612013000400008.

Schubeler, D. (2004) ‘The histone modification pattern of active genes revealed through genome-wide chromatin analysis of a higher eukaryote’, *Genes & Development*, 18(11), pp. 1263–1271. doi: 10.1101/gad.1198204.

Schulze, J. M. *et al.* (2009) ‘Linking Cell Cycle to Histone Modifications: SBF and H2B Monoubiquitination Machinery and Cell-Cycle Regulation of H3K79 Dimethylation’, *Molecular Cell*, 35(5), pp. 626–641. doi: 10.1016/j.molcel.2009.07.017.

Shanower, G. A. (2004) ‘Characterization of the grappa Gene, the Drosophila Histone H3 Lysine 79 Methyltransferase’, *Genetics*, 169(1), pp. 173–184. doi: 10.1534/genetics.104.033191.

Singer, M. S. *et al.* (1998) ‘Identification of high-copy disruptors of telomeric silencing in Saccharomyces cerevisiae.’, *Genetics*, 150(2), pp. 613–32.

Sonenshine, D. E. (1992) *Biology of Ticks*.

Storz, G. (2002) ‘An Expanding Universe of Noncoding RNAs’, *Science*, 296(5571), pp. 1260–1263. doi: 10.1126/science.1072249.

Stothard, P. (2000) ‘The sequence manipulation suite: JavaScript programs for analyzing and formatting protein and DNA sequences.’, *BioTechniques*, 28(6), pp. 1102, 1104.

Studer, G., Biasini, M. and Schwede, T. (2014) ‘Assessing the local structural quality of transmembrane protein models using statistical potentials (QMEANBrane)’, *Bioinformatics*, 30(17), pp. i505–i511. doi: 10.1093/bioinformatics/btu457.

Suzuki, M. M. and Bird, A. (2008) ‘DNA methylation landscapes: provocative insights from epigenomics’, *Nature Reviews Genetics*, 9(6), pp. 465–476. doi: 10.1038/nrg2341.

Tamura, K. *et al.* (2013) ‘MEGA6: Molecular Evolutionary Genetics Analysis Version 6.0’, *Molecular Biology and Evolution*, 30(12), pp. 2725–2729. doi: 10.1093/molbev/mst197.

Wade, M. J. (2007) ‘The co-evolutionary genetics of ecological communities’, *Nature Reviews Genetics*, 8(3), pp. 185–195. doi: 10.1038/nrg2031.

Walker, A. R. (1994) *The arthropods of humans and domestic animals : a guide to preliminary identification*. Chapman & Hall. Available at: https://books.google.at/books/about/Arthropods_of_Humans_and_Domestic_Animal.html?id=WKMeTbKOBuIC&redir_esc=y (Accessed: 7 July 2018).

Walker, A. R. *et al.* (2014) *Ticks of Domestic Animals in Africa: a guide to identification of species*, *Bioscience Reports*. doi: 10.1023/A:1025381712339.

Walter, J. and Hümpel, A. (2017) ‘Introduction to Epigenetics’, in *Epigenetics, Ethical,*

Legal and Social Aspects, pp. 11–29. doi: 10.1007/978-3-658-14460-9.

Webster, A. *et al.* (2018) 'Comparative study between Larval Packet Test and Larval Immersion Test to assess the effect of Metarhizium anisopliae on Rhipicephalus microplus tick larvae', *Experimental and Applied Acarology*. Springer International Publishing, 74(4), pp. 455–461. doi: 10.1007/s10493-018-0235-1.

Weinhold, B. (2006) 'Epigenetics: the science of change.', *Environmental health perspectives*. National Institute of Environmental Health Science, 114(3), pp. 160–167. doi: 10.1289/ehp.114-a160.

Wen, L. *et al.* (2015) 'Histone methyltransferase Dot1L plays a role in postembryonic development in Xenopus tropicalis.', *FASEB journal*. The Federation of American Societies for Experimental Biology, 29(2), pp. 385–93. doi: 10.1096/fj.14-252171.

Wong, M., Polly, P. and Liu, T. (2015) 'The histone methyltransferase DOT1L: Regulatory functions and a cancer therapy target', *American Journal of Cancer Research*. e-Century Publishing Corporation, 5(9), pp. 2823–2837. Available at: http://www.ncbi.nlm.nih.gov/pubmed/26609488 (Accessed: 19 April 2018).

Yang, J. *et al.* (2015) 'The I-TASSER Suite: protein structure and function prediction.', *Nature methods*, 12(1), pp. 7–8. doi: 10.1038/nmeth.3213.

Yu, W. *et al.* (2012) 'Catalytic site remodelling of the DOT1L methyltransferase by selective inhibitors', *Nature Communications*. Nature Publishing Group, 3, pp. 1–11. doi: 10.1038/ncomms2304.

Zhang, Y. (2001) 'Transcription regulation by histone methylation: interplay between different covalent modifications of the core histone tails', *Genes & Development*, 15(18), pp. 2343–2360. doi: 10.1101/gad.927301.

Zhang, Y. (2008) 'I-TASSER server for protein 3D structure prediction', *BMC Bioinformatics*, 9(1), p. 40. doi: 10.1186/1471-2105-9-40.