

## ***TOXOPLASMA GONDII* GENE EXPRESSION IS UNDER THE CONTROL OF REGULATORY PATHWAYS ACTING THROUGH CHROMATIN STRUCTURE**

BOUGDOUR A.\*\*\*, SAUTEL C.F.\*\*\*, CANNELLA D.\*\*\*, BRAUN L.\*\*\* & HAKIMI M.A.\*\*\*

### **Summary:**

The activity state of a gene is determined by a complex regulatory network of co-acting factors affecting the structure of the chromatin into which the gene is embedded. While significant changes of the transcriptome occur during cell differentiation in apicomplexan parasites, basic mechanisms controlling gene expression are still unknown. Recent studies support and expand the concept of the chromatin environment being key factor for the control of transcriptional activity in these lower eukaryotes organisms. Here, we review recent advances in the field of epigenetic gene regulation in *Toxoplasma gondii*, the model apicomplexan.

**KEY WORDS :** *Apicomplexa*, *Toxoplasma gondii*, chromatin, transcription, epigenetics.

**A** *apicomplexa* is a phylum of unicellular parasites that includes important human pathogens like *Plasmodium* species, the causative agents of the dreadful malaria, and *Toxoplasma gondii*, the common cause of congenital toxoplasmosis and severe encephalitis in immunocompromised individuals. Apicomplexan parasites undergo differentiation into several physiologically and morphologically distinct forms. Upon ingestion of *T. gondii* cysts, the tachyzoites (fast replicative form) go through several rounds of host cell invasion and replication. In response to the host immune response, slow growing bradyzoites form cysts in the tissues of the infected host. This form can revert to the rapidly dividing tachyzoite in immunocompromised individuals. Although, the molecular basis for this developmental switch is still unknown, the process of stage conversion is associated with dramatic changes of the pattern of gene expression (Radke *et al.*, 2005). In addition, transcriptomic and proteomic analysis allowed the identification of many stage specific genes (Radke *et al.*, 2005; TOXODB, proteomics data from J.M. Wastling). Therefore, the transcriptional activity of the genome must be regulated differentially by an undefined mechanism.

\* Laboratoire Adaptation et Pathogénie des Micro-organismes, Université Joseph Fourier Grenoble 1, BP 170, F-38042 Grenoble cedex 9, France.

\*\* CNRS UMR 5163 – ATIP+ group.

Correspondence: Mohamed-Ali Hakimi.

Tel.: 33 (0)4 76 63 74 69 – Fax: 33 (0)4 76 63 74 97.

E-mail: Mohamed-Ali.Hakimi@ujf-grenoble.fr

The control of gene expression is regulated in a highly organized fashion to ensure that specific genes are expressed at the appropriate times and levels to maintain cell homeostasis. In all eukaryotes, gene expression is controlled at multiple levels. First, promoter regions are packaged into chromatin, which regulates promoter accessibility. Second, sequence-specific DNA binding proteins (*i.e.* Transcriptional Factors, TFs) orchestrate the remodeling of chromatin and the recruitment of the transcription machinery. Third, general (or basal) transcription initiation factors (GTFs) such as TFIIA, -B, -D, -E, -F, -H and RNA polymerase II (pol II) assemble into a transcription pre-initiation complex (PIC). Fourth, pol II and associated elongation factors produce an RNA transcript. Apicomplexan genomes reveal a primitive eukaryotic transcription apparatus composed of a limited subset of the general transcription factors found in higher eukaryotes (Callebaut *et al.*, 2005; Meissner & Soldati, 2005; Iyer *et al.*, 2007). Intriguingly, no specific transcription factors were found so far in *Apicomplexa* and none of the well-characterized transcription regulators operating in other eukaryotes can be found by sequence comparison (Sullivan & Hakimi, 2006; Iyer *et al.*, 2007). However, even though we cannot rule out the existence of unconventional TFs (Balaji *et al.*, 2005), it is becoming apparent that an abundance of chromatin remodeling machineries are present (Hakimi & Deitsch, 2007; Iyer *et al.*, 2007). Thus, we hypothesized that *Apicomplexa* rely predominantly on epigenetic mechanisms to modulate chromatin structure and consequently its activity (Sullivan & Hakimi, 2006).

In eukaryotes, the DNA molecule is packed with different proteins especially by histone proteins to form nucleosome. The packaging of the DNA into nucleosomes affect all stages of transcription from PIC formation to elongation by restricting physically the accessibility of the DNA sequence to regulatory proteins or transcription apparatus. This restriction is dynamic and changes during differentiation or in response to exogenous cues. The nucleosome is a dynamic unit tightly regulated by various protein complexes. They are displaced at promoters during gene activation in a process that involves ATP-dependent nucleosome remo-

deling complexes, histone variants and a variety of histone modifications. Indeed, the relatively unstructured histone “tail” domains are evolutionarily conserved, exposed on the nucleosomal surface and subjected to a dizzying array of post-translational modifications, such as serine and threonine phosphorylation, lysine acetylation, lysine and arginine methylation, lysine ubiquitination and sumoylation, and ADP-ribosylation (for review Strahl & Allis, 2000; Kouzarides, 2007). Covalent modifications of histones, both in the tails and in the globular cores, regulate the packaging of genomic DNA into chromatin, gene activity and other functions of the genome. Emerging evidences suggest that a histone modification work alone or in combinatorial fashion on the same tail (cis) or on distinct tails (trans) to contribute to the so-called “histone code” (Strahl & Allis, 2000). Regardless of position or function, these covalent marks have been found to be dynamic and controlled by the opposing actions of various enzymes for their addition and removal (for recent review, see Kouzarides, 2007).

Considerable recent attention has focused on the molecular links between changes in parasite chromatin structure and the subsequent gene regulation. Most of our understanding about the role played by histone modifications in gene expression in *Apicomplexa* comes from studies on *P. falciparum* and *T. gondii*. The variety of histone modifications and the corresponding enzymatic systems that create or maintain them have been reviewed extensively in recent literature (Hakimi & Deitsch, 2007; Sullivan & Hakimi, 2006). The bulk of this review focuses on the influence of histone modifications on transcription in *T. gondii*, but it must be kept in mind that ATP-dependant chromatin remodeling activities (Sullivan & Hakimi, 2006) and histone variants (Sullivan *et al.*, 2006) also play critical roles in regulation of parasite gene expression.

## A NEW MAP FOR NAVIGATING THE *TOXOPLASMA* EPIGENOME

What did we learn so far in *T. gondii*? Like other eukaryotes, the parasite chromosomes display a typical nucleosomal organization involving both canonical and variant histones (Sullivan *et al.*, 2006). Many of the histone residues covalently modified in human or yeast cells are remarkably well conserved in *T. gondii* and are subject to similar modifications (Fig. 1). Besides, some of these marks on histone have been shown to have evolutionarily conserved roles in transcription regulation (Saksouk *et al.*, 2005; Sullivan & Hakimi, 2006; Sullivan *et al.*, 2006; Gissot *et al.*, 2007).

Physical maps describing the exact position of each modified histone on the genome provides an invaluable

tool to analyze its complex regulatory circuits involved in the gene expression. Such information is determined by the chromatin immunoprecipitation (ChIP) technology (Robyr *et al.*, 2002). Theoretically, the use of appropriate antibodies, allows the localization on the genome of either non-histone or histone proteins carrying or not specific covalent modifications. In *T. gondii*, ChIP assay was used for the first time to study histone acetylation, one of the most prominent modifications in the *T. gondii* gene expression (Saksouk *et al.*, 2005; Bhatti *et al.*, 2006). The pattern of histone acetylation correlates precisely with gene activity of the promoters analyzed in tachyzoite and bradyzoite cells. In *T. gondii*, as in other eukaryotic organisms, acetylation is a mark of gene activation (Saksouk *et al.*, 2005, Fig. 1B). Moreover, long-range histone acetylation patterns spanning many kilobases contribute to coordinate transcription of stage-specific tandemly arrayed multi-gene clusters (Saksouk *et al.*, 2005; Gissot *et al.*, 2007). In heterochromatin, the transcription of genes is usually repressed, whereas genes in euchromatin are usually active. Interestingly, in the same fashion as in other studied organisms, *T. gondii* heterochromatin regions are surrounded by short regions (hundreds of base pairs) of chromatin with high-level of histone acetylation (Sautel *et al.*, 2007). The biological meaning of this observation is still unknown but it was found that deacetylated histone-containing sites are associated with a reduced level of expression of the neighboring genes, suggesting a polar effect of these epigenetic modifications on the surrounding genes (*e.g.* bradyzoite-specific BSR4-related genes, Saksouk *et al.*, 2005; Gissot *et al.*, 2007). Thus, several coexpressed genes are clustered into eu- or heterochromatin domains, which suggest that chromatin, can be further divided into separate domains. Possibly, histone deacetylation sites may serve as partition borders for chromatin domains and affect gene expression. This mechanism may facilitate the coregulation of large subset of genes during differentiation or other cell processes.

In addition to histone acetylation, H3R17me2 and H3K4me3 methylations have been attributed to parasite gene activation (Saksouk *et al.*, 2005; Gissot *et al.*, 2007; Fig. 1B). Recent data has pointed out a central role of mono-, di- and trimethylated H4K20 and H3K9 in heterochromatin-mediated gene silencing (Sautel *et al.*, 2007). Also, the phosphorylation of serine 10 on histone H3 – a conserved epigenetic mark of mitosis – and H4K20me1 correlate perfectly with chromosome condensation during *T. gondii* tachyzoite division (Sautel *et al.*, 2007). These marks could signify the presence of epigenetic memory during the *T. gondii* life-cycle. More recent advances in ChIP technology have made possible to analyze genome-wide chromatin characteristics in eukaryotic cells. Global patterns of histone modifications are observed by combining the specifi-

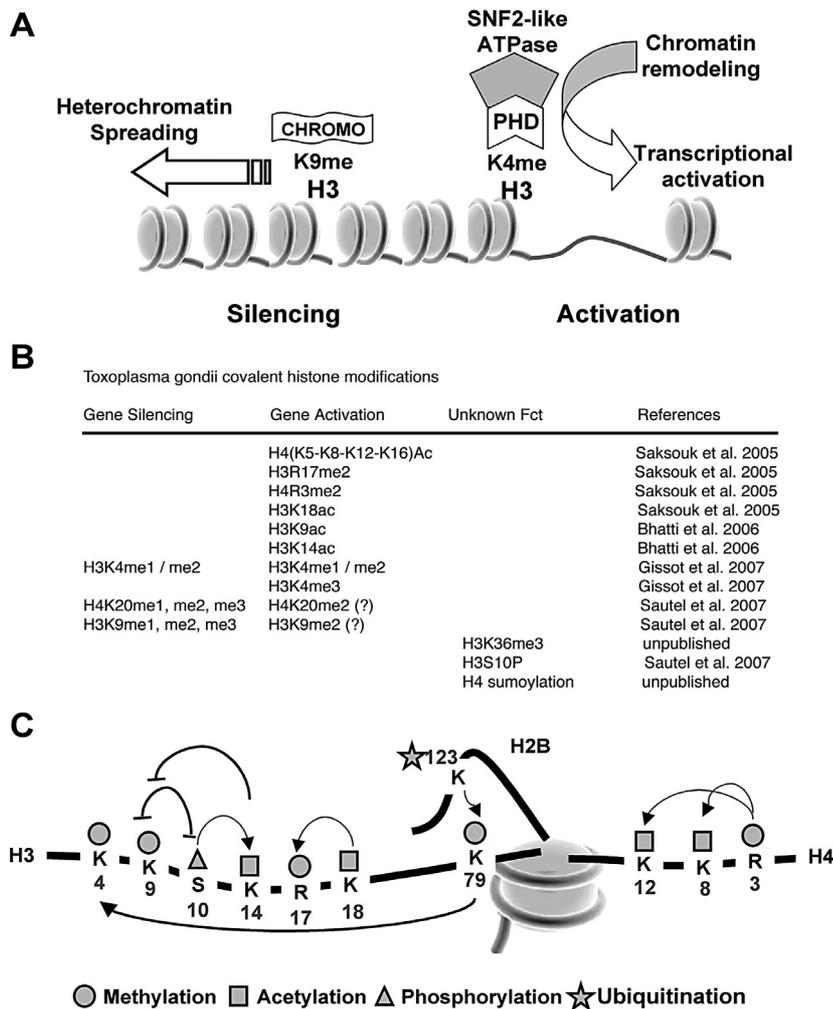


Fig. 1. – Chromatin modifications and their functions in *T. gondii*. (A) Gene expression alterations are brought by the recruitment of co-repressor (left side) or co-activator (right side) complexes, which contain specific “reading” domains recognizing repressive and active marks, respectively. (B) Overview of different covalent histone modifications identified in *T. gondii* so far. (C) Crosstalk between histone modifications. The positive or the negative influence of histone modifications are shown by an arrow or by a dish line, respectively. Adapted from Kouzarides T. (2007).

city of chromatin immunoprecipitation with the unbiased, high-throughput capabilities of microarrays (ChIP-on-chip; Robyr *et al.*, 2002). Applied to *T. gondii*, ChIP-on-chip analyses [genome-wide (Sautel *et al.*, 2007) or on a tiling arrays encompassing a 0.65-Mb contiguous portion of the chromosome 1b (Gissot *et al.*, 2007)] reveal that several modifications co-localize at precise loci and correlate either with significant gene expression or gene silencing. In the same manner as mentioned above, we observed at certain loci, that contiguous sets of genes were coregulated. The underlying mechanism of such regulation is currently unknown, but possibly, the presence of enhancer regions in the vicinity of the acetylated histones is required. At other loci, the effect of neighboring modifications may be reduced by histone deacetylation preventing the distal enhancers from accessing the promoter (unpublished data). The combination of data from genome-wide histone modifications and mRNA expression profiling may provide an integrated picture of the genome and epigenome of *T. gondii*. The identified binding sites may be used as a basis for annotating functional elements in genomes, including promoters, enhancers, silencing elements, insu-

lators, and boundary elements. ChIP-on-chip data may be used as a basis for re-annotating the predicted genome. Gissot *et al.* (2005) discovered transcripts in regions where genes are not predicted by looking for the enrichment of active marks. So far, such genome-wide data were generated using tachyzoite chromatin for immunoprecipitation and not bradyzoite or other specific conditions. It is therefore impossible to draw a clear-cut interpretation about the physiological functions of the histone modifications throughout the parasite life cycle. Understanding the overall mechanisms by which histone modifications control gene expression in *Apicomplexa* should contribute to identify subset of genes directly involved in the control of stage conversion. Analysis of the entire genome with high-resolution maps of modified histones combined with transcriptional data should help to understand the physiological state of cells at specific stages of the life cycle and predict its commitments. Transcription of multiple genes may be achieved by a single developmental event triggering a cascade of gene expressions. Still the robustness and buffering of determined epigenetic networks need to be tested. This could be also done to some extent by interfering with

the enzymes in charge of the 'code' writing. In other words, can we target pharmacologically chromatin-remodeling activity to treat Toxoplasmosis?

## CHROMATIN MODIFYING ENZYMES: THE *T. GONDII* ACHILLES' HEEL?

There is an extensive but still largely unexplored repertoire of proteins predicted to modify histones in the *T. gondii* genome (Sullivan & Hakimi, 2006). The marks can be placed and removed by a variety of site-specific enzymes such as histone acetyltransferases (HATs) and histone deacetylases (HDACs), histone methyltransferases (HMTs) and histone demethylases, and so on.

The histone deacetylases play a major role in keeping the balance between the acetylated and deacetylated states of chromatin, thereby controlling gene expression. Inhibitors of HDACs (HDACi) are under considerable exploration, in part because of their potential roles in reversing the silenced genes in transformed tumor cells by modulating transcriptional processes (Xu *et al.*, 2007). HDACi are now emerging as an interesting new class of antiprotozoal agent. Indeed, apicidin, a fungal metabolite, has broad-spectrum activity against the *apicomplexan* parasites, presumably via inhibition of protozoan histone deacetylases, and demonstrates efficacy against *Plasmodium berghei* malaria in mice (Darkin-Rattray *et al.*, 1996). The use of different class of HDACi is valuable to study the cause-and-effect relationship between histone acetylation and gene activity in *T. gondii*. Theoretically, each HDACi having a noticeable effect on global levels of histone acetylation could be used to determine the impact of the different histone modifications on chromatin activity during stage-conversion. A prerequisite for such studies is to determine whether each parasite HDACs identified are direct targets for such inhibitors.

The class of methyltransferase enzymes are other key players in the regulation of chromatin modelling and activity, thereby expanding the variety of regulatory mechanisms (Kouzarides, 2007). The monomethyl mark on lysine 20 of histone H4 correlates with intracellular levels of the methyltransferase TgSet8 throughout cell cycle, suggesting that this enzyme controls and is regulated through the developmental cycle of *T. gondii* (Sautel *et al.*, 2007). Also, H3R17 dimethylation was detected as a mark of gene activation in *Toxoplasma*, during both tachyzoite and bradyzoite stages. Attempts to generate a knockout clone of TgCARM1, which dimethylated H3R17, have been unsuccessful. Pretreatment of parasites with an inhibitor of TgCARM1 prior to infection of host cells increased the frequency of bradyzoite conversion, suggesting that the H3R17me2

mark is involved in the control of the tachyzoite to bradyzoite differentiation process (Saksouk *et al.*, 2005). Altogether, these data indicate that methyltransferase family of proteins has essential biological functions in *Apicomplexa*. Despite the increasing number of evidences demonstrating the important role played by histone modifying enzymes, very little is known about the mechanisms controlling their target specificities and functions in shaping the epigenome and orchestrating the development program of the parasite. At term, interfering with the 'writing' would end up with unraveling how active domains in chromatin are generated and altered during *T. gondii* stage transition.

## THE "HISTONE CODE" READERS

Histone modifications are thought to constitute a "histone code" that is used by the cell to encrypt various chromatin conformations and thereby, gene expression states. A clear correlation between certain modification of histones and DNA packaging in either active or inactive chromatin is conserved in *T. gondii*. In strong support to the histone code hypothesis, cross talk between marks on one or multiple histone tails was reported in *T. gondii*. Methylation of H3R17 was concurrent with acetylation of H3K18, suggesting a possible dual signature for gene activation (Saksouk *et al.*, 2005; Fig. 1C). In contrast, the selective enrichment of distinct H3K9 and H4K20 methylations with genes whose transcription is repressed support the presence of a H4K20 and H3K9 trans-tail histone code in the parasite (Sautel *et al.*, 2007). Fundamentally, still remains to be determined the biological significance of each covalent modification in *Apicomplexa*. Accumulating evidences suggest that these modifications act as selective binding platforms for recruiting specific regulatory proteins at the proper place and specific time; the histone code readers. Thus, numerous transcriptional regulators have domains that are capable of recognizing specific histone modifications. Among such readers are proteins with bromodomains (BROMO) that recognizes acetylated lysines, proteins with chromodomains (CHROMO) that recognized methylated residues, proteins with plant homeodomains (PHD) finger and tudor domains that bind methylated histones (Kouzarides, 2007; Fig. 1A). Sequence comparison analysis revealed that *T. gondii* has several uncharacterized genes encoding for BROMO-, CHROMO- and PHD-containing proteins (Hakimi & Deitsch, 2007; Iyer *et al.*, 2007). As in other organisms, many of these domains are found within parasite chromatin-modifying enzymes themselves, such as HAT, HMT and ATP-dependent remodeling enzymes. Still needs to be addressed how each domain reads the parasite dialect and transduces the epigenetic information into a cellular response.

## FUTURES DIRECTIONS

Epigenetic information contained in chromatin is a fascinating feature used by cells to modulate and extend the genetic code. As mentioned above, we have attempted to lay out the current state of knowledge on chromatin-dependent gene expression in *T. gondii*. In a not-so-distant future, we will decipher the complete transcriptional regulatory codes for *T. gondii* and construct blueprints for the initiation and maintenance of complex cellular processes, including parasite proliferation and differentiation.

Attention needs to be focused on the specific targeting of the chromatin remodeling enzymes to gene loci which has been shown to be mediated in part by DNA-binding transcription factors in other organisms. The relative paucity of such factors in the *Apicomplexa* phylum must imply an alternative targeting mechanism. How are the modifying enzymes recruited to specific DNA sequences? Noncoding RNAs participate by a still unknown mechanism in the control of epigenetic regulation and chromosomal dynamics such as gene silencing by RNA interference (RNAi), dosage compensation or imprinting (Zaratiegui *et al.*, 2007). It is therefore possible that non-coding RNAs provide the missing link in understanding the specific targeting of chromatin modifying and/or remodeling complexes to the proper chromosomal locations. While only a few conventional DNA-binding factors are found in *Apicomplexa*, a recent genome wide search suggests that a significant number of noncoding RNAs are present (Radke *et al.*, 2005; Hakimi & Deitsch, 2007). Small RNAs could target differentially modified histone proteins to specific loci, which could organize in return the genome into stretches of active and silent transcriptional regions.

## ACKNOWLEDGEMENTS

Dr Hakimi MA is supported by grants from ANR (MIME Program), CNRS (ATIP+ Program) and INSERM (Contrat Interface).

## REFERENCES

- BALAJI S., BABU M.M., IYER L.M. & ARAVIND L. Discovery of the principal specific transcription factors of Apicomplexa and their implication for the evolution of the AP2-integrase DNA binding domains. *Nucleic Acids Res.*, 2005, 33 (13), 3994-4006.
- BHATTI M.M., LIVINGSTON M., MULLAPUDI N. & SULLIVAN W.J. Jr. Pair of unusual GCN5 histone acetyltransferases and ADA2 homologues in the protozoan parasite *Toxoplasma gondii*. *Eukaryotic Cell*, 2006, 5 (1), 62-76.
- CALLEBAUT I., PRAT K., MEURICE E., MORNON J.P. & TOMAVO S. Prediction of the general transcription factors associated with RNA polymerase II in *Plasmodium falciparum*: conserved features and differences relative to other eukaryotes. *BMC Genomics*, 2005, 23, 6-100.
- DARKIN-RATTRAY S.J. *et al.* Apicidin: a novel antiprotozoal agent that inhibits parasite histone deacetylase. *Proc. Natl. Acad. Sci. USA*, 1996, 93 (23), 13143-13147.
- GISSOT M., KELLY K.A., AJIOKA J.W., GREALLY J.M. & KIM K. Epigenomic modifications predict active promoters and gene structure in *Toxoplasma gondii*. *PLoS Pathog.*, 2007, 3 (6), 77.
- HAKIMI M.A. & DEITSCH K.W. Epigenetics in Apicomplexa: control of gene expression during cell cycle progression, differentiation and antigenic variation. *Curr. Opin. Microbiol.*, 2007, 10 (4), 357-362.
- IYER L.M., ANANTHARAMAN V., WOLF M.Y. & ARAVIND L. Comparative genomics of transcription factors and chromatin proteins in parasitic protists and other eukaryotes. *Int. J. Parasitol.*, 2007, Sep. 15 [Epub ahead of print].
- KOUZARIDES T. Chromatin modifications and their function. *Cell*, 2007, 128 (4), 693-705.
- MEISSNER M. & SOLDATI D. The transcription machinery and the molecular toolbox to control gene expression in *Toxoplasma gondii* and other protozoan parasites. *Microbes Infect.*, 2005, 7 (13), 1376-1384.
- RADKE J.R., BEHNKE M.S., MACKAY A.J., RADKE J.B., ROOS D.S. & WHITE M.W. The transcriptome of *Toxoplasma gondii*. *BMC Biol.*, 2005, 2, 3-26.
- ROBYR D., SUKA Y., XENARIOS I., KURDISTANI S.K., WANG A., SUKA N. & GRUNSTEIN M. Microarray deacetylation maps determine genome-wide functions for yeast histone deacetylases. *Cell*, 2002, 109 (4), 437-446.
- SAKSOUK N., BHATTI M.M., KIEFFER S., SMITH A.T., MUSSET K., GARIN J., SULLIVAN W.J. JR, CESBRON-DELAUW M.F. & HAKIMI M.A. Histone-modifying complexes regulate gene expression pertinent to the differentiation of the protozoan parasite *Toxoplasma gondii*. *Mol. Cell Biol.*, 2005, 25 (23), 10301-10314.
- SAUTEL C.F., CANNELLA D., BASTIEN O., KIEFFER S., ALDEBERT D., GARIN J., TARDIEUX I., BELRHALI H., & HAKIMI M.A. SET8-mediated methylations of histone H4 lysine 20 mark silent heterochromatic domains in apicomplexan genomes. *Mol. Cell Biol.*, 2007, 27 (16), 5711-5724.
- STRAHL B.D. & ALLIS C.D. The language of covalent histone modifications. *Nature*, 2000, 403 (6765), 41-45.
- SULLIVAN W.J. Jr & HAKIMI M.A. Histone mediated gene activation in *Toxoplasma gondii*. *Mol. Biochem. Parasitol.*, 2006, 148 (2), 109-116.
- SULLIVAN W.J. Jr, NAGULESWARAN A. & ANGEL S.O. Histones and histone modifications in protozoan parasites. *Cell. Microbiol.*, 2006, 8 (12), 1850-1861.
- XU W.S., PARMIGIANI R.B. & MARKS P.A. Histone deacetylase inhibitors: molecular mechanisms of action. *Oncogene*, 2007, 26, 5541-5452.
- ZARATIEGUI M., IRVINE D.V. & MARTIENSSSEN R.A. Noncoding RNAs and gene silencing. *Cell*, 2007, 128 (4), 763-776.