Biometals
DOI 10.1007/s10534-017-0053-7

TvZNF1 is a C_2H_2 zinc finger protein of *Trichomonas vaginalis*

José Luis Villalpando · Rodrigo Arreola · Jonathan Puente-Rivera · Elisa Azuara-Liceaga · Jesús Valdés · Lilia López-Canovas · Alma Villalobos-Osnaya · Maria Elizbeth Alvarez-Sánchez

Received: 8 September 2017 / Accepted: 19 September 2017

Abstract The zinc fingers proteins (ZNF) are the largest family of DNA binding proteins and can act as transcriptional factors in eukaryotes. ZNF are implicated in activation in response to environmental stimulus by biometals such as Zn^{2+}. Many of these proteins have the classical C_2H_2 zinc finger motifs (C_2H_2-ZNFm) of approximately 30 amino acids, where a Zn^{2+} ion is coordinated by two cysteine and two histidine residues. *Trichomonas vaginalis* is a protozoan parasite than responds to environmental changes including Zn^{2+}. Until now has not been described any ZNF that could be involved in the regulation of genic expression of *T. vaginalis*. Here, we characterized in silico and experimentally an annoted ZNF (TvZNF1) from *T. vaginalis* and isolated the gene, *tvznf1* encoding it. TvZNF1 have eight C_2H_2-ZNFm with residues that maybe involved in the structural stability of DNA binding motifs. In this work we confirmed the Zn^{2+} upregulation expression of *tvznf1* gene. Recombinant TvZNF1 was able to bind to specific DNA sequences according to EMSA assay. Additionally, we demonstrated that recombinant TvZNF1 bind to MRE signature in vitro, which strongly suggests its role in transcriptional regulation, similar to the one observed for mammalian MTF-1. This result suggested a conserved mechanism of genic regulation mediated by ZNFs in *T. vaginalis*.

Keywords *Trichomonas vaginalis* · C_2H_2 Zinc finger protein · DNA binding protein · TvZNF1

Abbreviations

ZNF	Zinc finger protein
ZNFm domain	Zinc finger protein domain(s)
C_2H_2-ZNFm	C_2H_2 zinc finger motif(s)
TvZNF1	Zinc finger protein of *Trichomonas vaginalis* TvZNFm1-8, specific domain of TvZNF1
rTvZNF1	Recombinant TvZNF1

Electronic supplementary material The online version of this article (doi:10.1007/s10534-017-0053-7) contains supplementary material, which is available to authorized users.

J. L. Villalpando · J. Puente-Rivera · E. Azuara-Liceaga · L. López-Canovas · A. Villalobos-Osnaya · M. E. Alvarez-Sánchez (✉)
Posgrado en Ciencias Genómicas, Universidad Autónoma de la Ciudad de México (UACM), San Lorenzo # 290, Col. Del Valle, 03100 Mexico City, Mexico
e-mail: maria.alvarez@uacm.edu.mx

R. Arreola
Psychiatric Genetics Department, Clinical Research Branch, National Institute of Psychiatry, "Ramón de la Fuente", Calzada México-Xochimilco 101, Colonia San Lorenzo Huipulco, Tlalpan, 14370 México City, DF, Mexico

J. Valdés
Department of Biochemistry, Centro de Investigación y de Estudios Avanzados del Instituto Politécnico Nacional, México City, Mexico

Published online: 09 October 2017

Introduction

The family of transcription factors called zinc finger proteins (ZNF) are found in the most eukaryote genomes (Ding et al. 2008). The class of zinc finger motifs (ZNFm) named C_2H_2-ZNFm consist in 20–30 amino acid residues coordinating an ion zinc with two cysteine and two histidine residues. C_2H_2-ZNFm is considered the classic ZNF domain. It contain the signature $CX_{2\text{-}4}CX_{12}HX_{2\text{-}6}H$ C_2H_2 which is involved in the structural stability and function of ZNF and could be present in tandem repeats (Krishna et al. 2003). The super secondary structure of a C_2H_2-ZNFm has two β strands in the N-terminal and one α-helix in the C-terminal sequence (Iuchi 2001). Around the 0.7 and 0.8% of *Arabidopsis thaliana* and *Saccharomyces cerevisiae* proteins have C_2H_2-ZNFm, respectively (Böhm et al. 1997; Englbrecht et al. 2004). Meanwhile, in the human genome more than 3% of total genes encodes for ZNF (Klug 2010). In contrast, proteins with C_2H_2-ZNFm have been annotated but not yet characterized as zinc finger putative in the genome of some protozoan parasites including apicomplexans, *Entamoeba histolytica*, *Trichomonas vaginalis*, phytophthora and ciliates (Harb and Roos 2015). Different functions have been described for the C_2H_2-ZNFm proteins, for instance, Ros, a ZNF from bacteria *Agrobacterium tumefaciens,* is a transcriptional regulator of virulence genes (Esposito et al. 2006).

On the other hand, the metal regulatory transcription factor 1 (MTF-1) and the zinc-responsive transcriptional regulator (Zap1) are transcriptional C_2H_2-ZNFm proteins factors able to sense and maintain the homeostasis of metals as Zn^{2+} in several organisms (Hardyman et al. 2016; Rutherford and Bird 2004). The first role described for the C_2H_2-ZNFm proteins was DNA binding in transcriptional process in response to environmental stimulus such as free radical, temperature and high concentrations of metals such as Zn^{2+}, Cd^{2+} and Cu^{2+}. These proteins also interact with RNA or proteins and may act as splicing factors (Iuchi 2001; Laity et al. 2001). Mostly, C_2H_2-ZNFm proteins are related to transcriptional, splicing factors or DNA damage repair proteins and bind DNA using their ZNFm domains. The first C_2H_2-ZNFm protein domain discovered was in the transcriptional factor III A (TFIIIA) from *Xenopus laevis* and is the most common structural domain found in eukaryotes (Theunissen et al. 1992).

T. vaginalis, the causal agent of trichomoniasis in humans, is adapted to microenvironmental conditions in male and female urogenital tract (Figueroa-Angulo et al. 2012). Meanwhile the basal transcriptional process is mediated by the transcription factor IBP39 (Lau et al. 2006; Liston et al. 2001) the genes encoding the Myb family proteins are modulated by Fe^{2+} concentrations (Horváthová et al. 2012; Lau et al. 2006; Schumacher et al. 2003). In addition, *T. vaginalis* is also able to survive in high Zn^{2+} concentrations by proteomic regulation (Quintas-Granados et al. 2013; Vazquez Carrillo et al. 2011) and may be the response is achieved mediated by ZNF. The aim of this work was to analyze and characterize the structural and functional properties of annotated ZNF in *T. vaginalis* and provide evidences about their possible function as a DNA-binding protein by bioinformatic and experimental methods.

Materials and methods

Database searching for zinc finger protein in *T. vaginalis*

A gene and protein sequences of *T. vaginalis* G3 (taxid: 412133) annotated as a putative "zinc finger protein" with accession number TVAG_458980 was found in TrichDB database (Aurrecoechea et al. 2008). The amino acids sequences of zinc finger proteins MTF-1 (UNIPROT: Q14872) and SP-1 (UNIPROT: P08047) from *Homo sapiens*, Zap1 (UNIPROT: P47043) from *S. cerevisiae* and Zif268 (UNIPROT: P08046) from *Mus musculus*, were used as query for searching sequence homology in "Non-redundant protein sequences (nr)" database of *T. vaginalis* G3 with an expected threshold parameter of 0.001 and using BLASTp with algorithm Psi-BLAST from NCBI online service server and an expected threshold value of 0.001 (Altschul et al. 1997). Analysis of the TvZNF1 identified zinc finger protein TvZNF1.

TvZNF1 protein sequence was used as query for BLASTp with Psi-BLAST algorithm in NCBI against Protein Data Bank database (Bernstein et al. 1977). The Conserved Domains database coupled to Psi-BLAST from NCBI server (Altschul et al. 1997) was

used to identify the class or family C_2H_2-ZNFm in TvZNF1 sequence (Marchler-Bauer et al. 2014). The super secondary structural motifs predictions were performed using the integrated InterPro Domains database from the TrichBD database after BLAST of TvZNF1 (Mulder et al. 2005).

Secondary structure prediction and structure modeling of TvZNF1

The prediction of secondary structure of TvZNF1 was done using the programs PCOILS (Alva et al. 2016), HHPRED (Hildebrand et al. 2009), GOR4 (Combet et al. 2000), CFSSP (Chou and Fasman 1974) and APSSP (http://imtech.res.in/raghava/apssp/). For protein structure modeling and function prediction I-TASSER program server was used (Lau et al. 2006). The prediction of protein structure of TvZNF1 was done dividing the protein sequence in three blocks named A, B and C. The block A was built from residues 1 to 101 corresponding to ZNFm1 to ZNFm3 using the multidomain ZNF PRDM9 (PRD1-BF1-RIZ1 homologous domain-containing 9) a protein structure (PDB: 5EGB. PRDM9 having four ZNFm in complex with DNA as template; the block B was built from residues 80 to 166 corresponding to ZNFm3 to ZNFm5 using Gfi-1 ZNF complexed with DNA by three ZNFm (3′–5′ orientation) (PDB: 2KMK); and block C from residues 140 to 276 corresponding to ZNFm5 to ZNFm8 using as template RADR, a Zif-268 variant structure with three ZNFm in complex with GCAC DNA sequence (PDB 1A1I) as templates. The models of TvZNF1 were built without Zn^{2+} and DNA to determine the secondary structures stability of each ZNFm. The secondary structure predictions of each block were done with Stride program (Frishman and Argos 1995) and the best secondary structure prediction respect to each model was selected.

Alignments of TvZNF1 and MTF-1

T. vaginalis TvZNF1 and human MTF-1 proteins were manually aligned following the next considerations: (1) TvZNF1 and human MTF-1 bind to MRE (Metal Response Element) 5′-TGCrCnC-3′ core DNA sequences (Langmade et al. 2000) and to consensus signature with a length of seven to nine nucleotides (nt) 5′-TGCRCNC[G/C][G/C]-3′ (Lichtlen et al. 2001). We selected the MRE signature sequence with a length of 10 nucleotides (5′-TGCRCNC[G/C][G/C]C-3′) based on the triple zinc fingers model of DNA binding previously described (Iuchi 2001). In this model, each ZNFm bind to 3–4 nucleotides; (2) Three ZNFm domains from C- to N-terminal orientation bind to DNA signature in 5′ to 3′ orientation because the second and third ZNFm bind to the MRE core sequence 5′-TGCRCNC-3′; (3) Human MTF-1 protein binds to MRE signature through the ZNFm 2, 3 and 4 that constitute the core of DNA-binding domain, whereas ZNFm 5 and 6 are apparently not necessary for the binding to DNA signature in vitro (Bittel et al. 2000; Chen et al. 1999); (4) We considered for the alignment the determinants residues for specific DNA binding determined by the position in the α-helix (residues positions -1, 2, 3 and 6) of each ZNFm (Iuchi 2001). This region includes four residues in the N-terminal of the α-helix and three residues in the C-terminal of the α-helix and next loop of the motif when short loops are present between two ZNFm (Iuchi 2001; Razin et al. 2012) and finally, (5) with these considerations, the ZNFm 2, 3 and 4 of MTF-1 were aligned with six combinations of three ZNFm from TvZNF1 to identify conserved residues involved in the recognition of the MRE signature.

Parasites culture

T. vaginalis isolate HGMN01 were grown until mid-logarithmic phase in Diamond's trypticase-yeast extract-maltose (TYM) medium pH 6.2 supplemented with 10% (v/v) of heat-inactivated horse serum (Gibco BRL, Grand Island, NY, USA). The culture medium was supplemented with 1.6 mM of $ZnCl_2$ (Sigma-Aldrich, Co., St. Louis, MO, USA).

RT-PCR assays

Total RNA from 2×10^7 parasites was extracted by TRIzol method following the manufacturer's recommendations (Invitrogen, Life Technologies, Carlsbad, CA, USA) To obtain cDNA, 5 µg of total RNA was treated with superscript II reverse transcriptase kit and oligo dT (dT18) primer (10 pmol/µl) (Invitrogen). The *tvznf1* gene was amplified using specific internal primers 5′-TGTGGTGCCGAATTTGACCT-3′ (forward), and 5′-TTCCACAACCTTCCACAGGG-3′ (reverse) by semi-quantitative RT-PCR assay. A 112 bp fragment of β-tubulin of *T. vaginalis* was

amplified as control (León-Sicairos et al. 2004). Two independent assays were performed by triplicate.

Expression and purification of recombinant TvZNF1

The *tvznf1* ORF was amplified from genomic DNA using the primers 5′-CCGGATCCATGGGAGACACAGAAGAG-3′ (forward) and 5′-CCAAGCTTTATTTATTTTGCAAATCCTTCTG-3′ (reverse) with the restriction sites *Bam*HI and *Hind*III (underlined) respectively, using the following PCR conditions: 94 °C for 5 min, 94 °C for 1 min, 52 °C for 1 min, 72 °C for 1 min by 30 cycles and a final stage of 72 °C for 7 min. The 831 pb amplicon was cloned into pJET1.2/blunt with CloneJET PCR cloning Kit (Thermo Fisher Scientific, Waltham, MA, USA), subcloned into pQE-80L vector (Qiagen, Hilden, Germany) and transformed into *E. coli* M15 competent cells. Positive clones were sequenced in 3130 Applied Biosystems Genetic Analyzer (PCG-UACM, Mexico) and analyzed using Bioedit software (http://www.mbio.ncsu.edu/BioEdit/bioedit.html). TvZNF1 recombinant (rTvZNF1) expression with a histidine tag was induced with 1 mM IPTG (Sigma) for 16 h at 37 °C. The protein purification was performed with Ni–NTA, IEX and HIC chromatography according to the manufacturer's instructions under native conditions. The sample separation and degree of purification was evaluated by electrophoresis in 10% tricine-SDS-PAGE staining with Coomassie brilliant blue and corroborated by anti-histidine WB recognition.

Electrophoretic mobility shift assay (EMSA)

EMSA assay was performed with the probes MRE2 (MRE2f 5′-CGAGACACACATGCACACGCACACAGG-3′) and MRE2r (5′-TCGACCTGTGTGCGTGTGCATGTGTGTCTCGAGCT-3′) which are specific for the human metal response element as previously reported (Lichtlen et al. 2001). The non-related element probe NRE (NREf 5′-TCTAGAAATAAAGTTATT-3′, NREr 5′-TCTAGAAATAACTTTATT-3′) was used as negative control. The DNA probes were radiolabeled at 5′ end with α-^{32} P using DNA Polymerase I, Large (Klenow) fragment (New England Biolabs) and purified with QIAquick Nucleotide Removal Kit (Qiagen) according to the manufacturer's instructions. The EMSA binding reaction was done mixing binding buffer 1 × (12 mM Hepes–KOH (pH 7.9), 0.5 mM EDTA,10 μM $ZnCl_2$, 4 mM $MgCl_2$, 4 mM spermidine, 1 mM KCl, 1 mM DTT and 10% glycerol) with 44 μM of rTvZNF1 and 1 μg of polydI-dC (Invitrogen) and incubated for 20 min to 4 °C. The radiolabeled probe (20,000 c.p.m) was added and incubated for 25 min at 4 °C. The DNA–protein complexes formation was checked running the reaction in 6% polyacrylamide gel at 100 V for 6 h. Dried gel was exposed to X-ray film (Fujifilm) to document the results. Three independent assays were performed.

Results

Identification of a ZNF in *T. vaginalis*

A putative sequence annotated as "zinc finger protein" was identified in *T.* vaginalis G3 isolate with an ORF of 831 bp was found in TrichDB with the access number TVAG_458980 (Gene Bank Accession: XP_001324049; UNIPROT: A2E6A3). It encodes a protein of 276 amino acids which has a theoretical molecular weight of 32.199 kDa and a *p*I of 8.06. The gene and encoded protein were named here *tvznf1* and TvZNF1, respectively. The location of *tvznf1* was mapped in *T. vaginalis* G3 in the position 98138-98968 (+) genomic scaffold. It is flanked by TVAG_458970 and TVAG_458990 genes, which encodes a conserved hypothetical protein and a golgi IMH1 putative protein, respectively (Fig. 1a). To identify homologous ZNF, the protein sequences of MTF-1 and SP1 of *H. sapiens* (Q14872 and P08047, respectively), Zap1 of *S. cerevisiae* (P47043) and Zif 268 of *M. musculus* (P08046), were used for searching by BLAST in TrichDB database. The sequence with the access number TVAG_458980 corresponding to TvZNF1 of *T. vaginalis* was the only found in the database.

The TvZNF1 sequences searching in databases did not revealed any possible orthologous or paralogous protein sequences. However, a further BLASTp analysis identified two possible orthologous proteins to TvZNF1; they had identities of 38–41%, respectively and significant E-values with hypothetical peptides of a protein of *Tritrichomonas foetus* (Gene Bank accession: OHT14463.1) and other corresponding to a protein of *Takifugu rubripes* (GB accession:

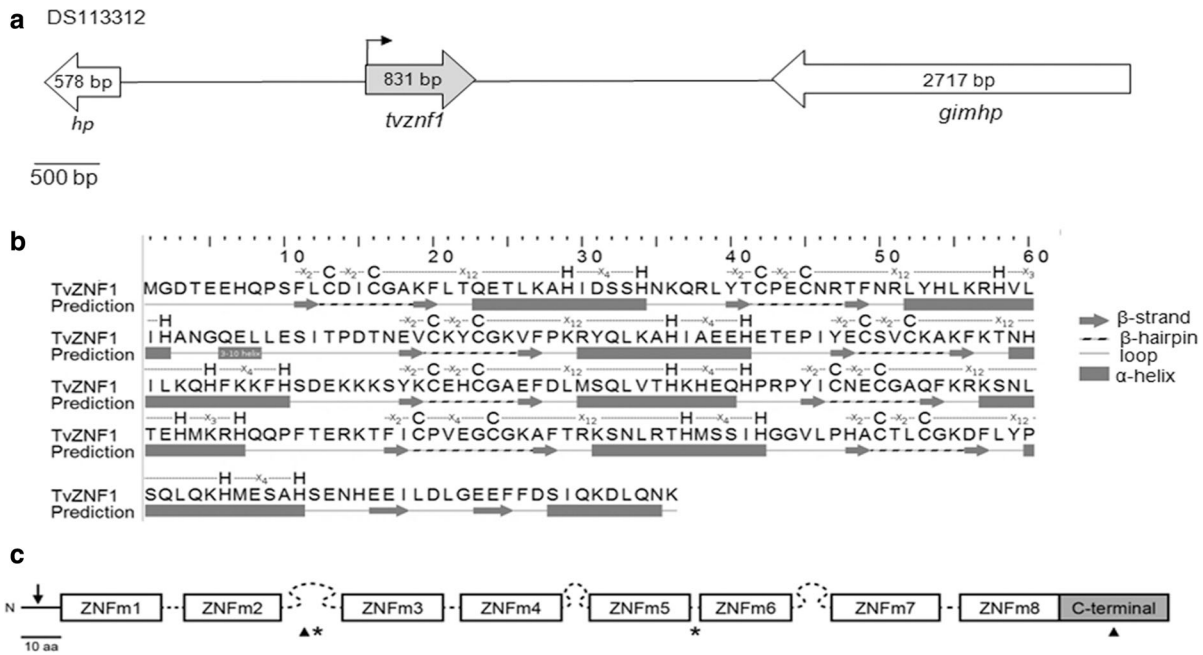

Fig. 1 Identification of TvZNF1 in *T. vaginalis*. **a** The *T. vaginalis* genome sequence TVAG_458980 is present in the contig +98,138-98,908 (grey arrow). *hp* and *gimh1p* genes codify the hypothetical protein and Golgin IMH1 putative, respectively. Bar represents 100 bp. **b** TvZNF1 sequence and prediction of C_2H_2-ZNFm. β-strands are represented as arrows and α-helix as rectangles and the loops containing Cys are represented as dashed lines between the β-strands. **c** Predicted structural topology of the eight ZNF motifs (ZNFm) in TvZNF1. Linkers amino acids are represented as dotted lines; loops 1, 2 and 3 (L1-3) as dotted loops. Linkers and loops without Arg and Lys are marked with asterisks. The amino (N) terminal is marked with an arrow and the carboxyl (C)-terminal subdomain (black box) and the L1 with hypothetical protein–protein interaction region is marked with black triangles. Bar represents 10 aa

XP_011617401.1). Additional, others proteins with identities of around 30% with TvZNF1 were found, they have eight or more ZNFm and belong to a family predicted as gastrula ZNF (data not show).

Predictions of the motifs, domains and the general organization of TvZNF1

TvZNF1 amino acid sequence was analyzed to predict its domains using InterPro domains database and Conserved Domains program. These analyses predicted eight C_2H_2-ZNFm in TvZNF1 residues at positions 11–34 (ZNFm1), 40–62 (ZNFm2), 78–101 (ZNFm3), 107–130 (ZNFm4), 138–161 (ZNFm5), 165–187 (ZNFm6), 197–222 (ZNFm7) and 228–252 (ZNFm8). All C_2H_2-ZNFm presented the canonical signature motif X_2-C-$X_{2\text{-}4}$-C-X_{12}-H-$X_{2\text{-}5}$-H for ZNF (Krishna et al. 2003) and four short loops as linkers of 3–5 residues between them and three loops with length of 7–15 residues (Fig. 1b).

Secondary structure prediction of TvZNF1

The initial peptide sequence analysis of TvZNF1 produced a non-optimal prediction of the C_2H_2-ZNFm secondary structure. For this reason, the secondary structure prediction of TvZNFm(s) was performed analyzing the conformation of 11 structured individuals models generated in three blocks. As a result, the canonical C_2H_2-ZNFm and the β-harping-β-α-helix secondary structures were predicted with stable secondary structures except for the TvZNFm3 and TvZNFm6 models where the β-strand conformations were lost. The 11 structured individuals models had a marginal variation in the α-helix length, with an average of 12 residues (Fig. 1b, grey boxes). The disulphide bridges between Cys in TvZNF1 sequence were not identified, and a relaxation in the His residues involved in Zn^{2+} interaction were identified in all structures of the motifs. This result was expected due to cation absence in the generated models.

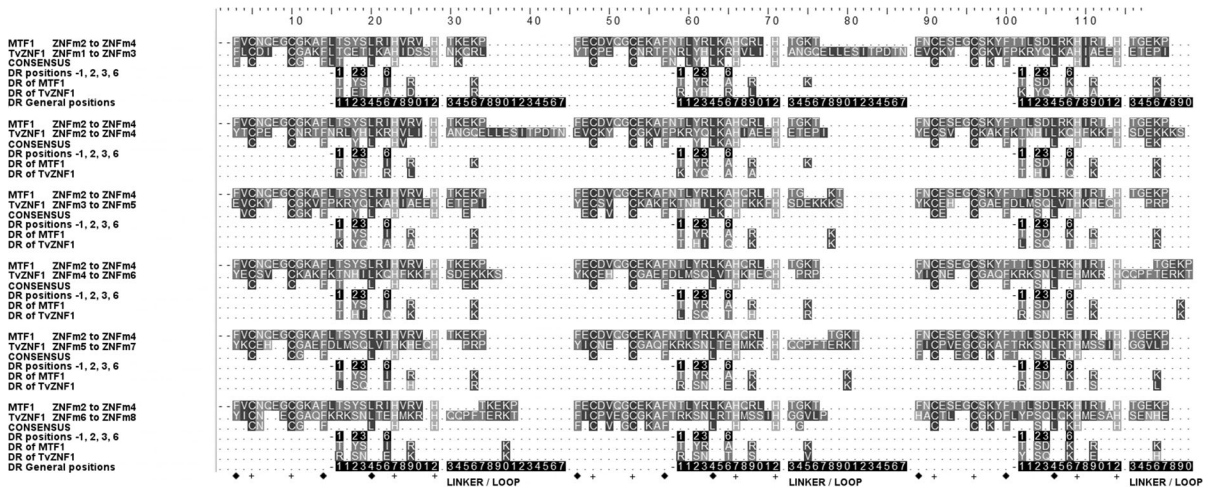

Fig. 2 Comparison between MTF-1 and TvZNF1 sequences. DNA-binding core C_2H_2- ZNFm (ZNFm2 to ZNFm4) of MTF-1 was compared with all possible combinations of three TvZNFm from TvZNF1. The numeration represents the relevant positions of the determinant residues for the specific binding to DNA (−1, 2, 3, 6). The additional position (8 to 27) are cover residues from the first His to posterior linkers or loops that could be interacting with the phosphate backbone of DNA. The conserved structural residues are marked with filled diamond and the Cys and His residues that coordinate the Zn^{2+} ion are represented by plus

Additionally, in the 25 residues of C-terminal region a potential β-harping-β-α-helix-like motif with two β-strands in the positions 256-EEI-258, 263-EEF-265 and one short α-helix with eight residues in the positions 268–275 with the sequence SIQKDLQN were also predicted. Therefore, this potential motif did not has the canonical C_2H_2-ZNFm, because the positions of Cys residues were replaced by Asp (260) and Glu (263). This region also might have a coiled-coil type subdomain. However, if this C-terminal region sequence of TvZNF1 is a real functional motif is not known yet.

Prediction of residues in TvZNF1 involved in DNA binding

The transcriptional factor MTF-1 has six ZNFm (Uniprot access number: Q14872) and is able to bind to metal response element (MRE) signature 5′-TGCRCNC-3′ of target gene (Langmade et al. 2000) with the ZNFm 2, 3 and 4. These motifs constitute the amino acidic core of DNA-binding domain.

To identify similar residues with the binding capacity to DNA in TvZNF1 sequence, the ZNFm of TvZNF1 were aligned and analyzed with MTF-1 guided by the mentioned characteristics of MTF-1. We found that TvZNF1 and MTF-1 are able to bind to the same DNA signature. We selected the sequences of ZNFm 2 to 4 of MTF-1 for comparisons with the equivalent segments of TvZNFm1 to 3, TvZNFm2 to 4, TvZNFm3 to 5, TvZNFm4 to 6, TvZNFm5 to 7 and TvZNFm6 to 8 of TvZNF1 (Fig. 2). TvZNF1 and MTF-1 proteins conserved residues in all the ZNFm analyzed, which correspond to the canonical class C_2H_2-ZNFm. We also found additional conserved positions related to structural functions (Fig. 2, black rhombuses) of the specific type of canonical C_2H_2-ZNFm (Fig. 2, plus symbols).

The predicted structural function signature for C_2H_2-ZNFm was [**F/Y/H**]-X-C-$X_{2\text{-}4}$-C-X_3-[**F/Y**]-X_5-**L**-X_2-H-$X_{2\text{-}5}$-H, where the residues in bold have structural functions in the correct folding on the C_2H_2-ZNFm. The positions of determinant residues considered as relevant in the MRE signature recognition in DNA were −1, 2, 3 and 6 (Fig. 2, putative DNA binding residues) but these particular residues between each ZNFm had low or no conservation.

A second strategy to elucidate which of the triple ZNFm of TvZNF1 was responsible for the binding to MRE was the searching of the sequence pattern in the ZNFm2, 3 and 4 of MTF-1. We found that the first and the second motif of MTF-1 shared similar residues able to bind to DNA, but the third motif showed differences in this function. TvZNF1 had a similar pattern in the domain of the ZNFm6, 7 and 8, but without a high conservation of residues in relation to

Table 1 Possible different set of determinant residues of MTF-1 and TvZNF1 to bind to DNA subsites

First Model of binding to MRE signature	
ZNFm.....8..7..6	
.....1234567890	
MRE consensus .....**TGCrCnC****	Core signature
AFP MRE2s 5-cgagacacaca-**TGCACACGC**acacagg-3	Used EMSA Probe
Subsite 1a .....TGC.....	
Subsite 2a .....ACA.....	
Subsite 3a .....CGC....	

Second Model of binding to MRE signature
ZNFm.....8...7...6
.....1234567890123
Subsite 2a .....TGCn.....
Subsite 2b .....CACn.....
Subsite 2c .....CAC..

Third Model of binding to MRE signature
ZNFm.....8..7.....6
.....1234567890123
Subsite 3a .....TGC.....
Subsite 3b .....ACAnnn....
Subsite 3c .....ACA.

Fourth Model of binding to MRE signature
ZNFm.....8...7....6
.....1234567890123
Subsite 3a .....TGC.....
Subsite 3b .....nCACnn....
Subsite 3c .....ACA.

MTF-1. The ZNFm4 of MTF-1 and the ZNFm8 of TVZNF1 only showed a 50% of identity in the conserved positions 2 (S) and 6 (L) (Fig. 2, MTF-1 DNA binding residues). Probably these are the binding residues of the subsite MRE2 s in the 5′-TGC-3′ region. The ZNFm2 and 3 of MTF-1 apparently have their own signature of residues for DNA binding, as well as TvZNFm6 and 7 of TvZNF1 that apparently have also their own signature due to the possible recognition and bind to 3′- region of the MRE2s (5′-ACACGC-3′).

The comparison between ZNFm(s) of TvZNF1 and MTF-1 suggested that TvZNF1 could bind to MRE signature recognizing DNA subsites in different ways to respect to MTF-1. It should be note the MRE2 probe analysis shown different combination of possible subsites in the sequence (Lichtlen et al. 2001) (Table 1).

The BLASTp analyses in PDB databases also revealed a large collection of ZNF structures but any of them were useful for predicting the binding domain of DNA signature in TvZNF1 because these structures had not identity in the determinant residues of the specific binding of DNA region. Then, they should not be associated with the binding to a DNA signature sequence. Only, TvZNFm2, 6 and 7 showed identity with PDB structures with a significant E-value, the others ZNFm appeared to have not significant E-values and had fragmented or small regions with any identity and similarity (Fig. 3).

Expression of *tvznf1* in high Zn^{2+} concentrations

In order, to analyze the expression profile of *tvznf1* under normal and high Zn^{2+} concentrations several RT-PCR assays of the gene in *T. vaginalis* HGMN01 isolate were performed. They showed an increase in the expression of *tvznf1* gene as the response to the stimulation with $ZnCl_2$ (Fig. 4a). We demonstrated that in presence of Zn^{2+} in the culture medium the

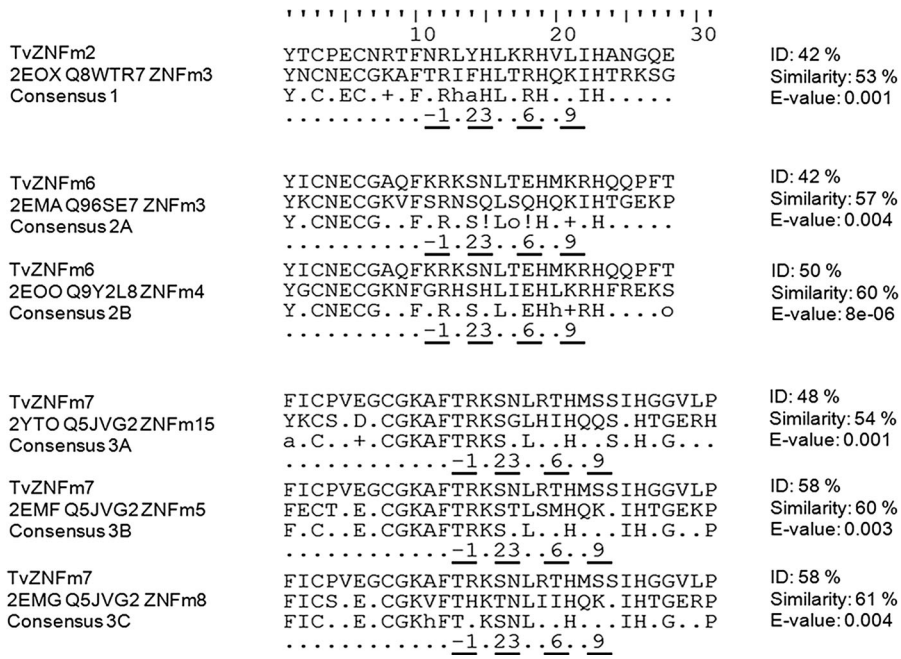

Fig. 3 The most similar PDB structure sequences to C_2H_2-ZNFm from TvZNF1. TvZNFm7 show the higher identity with the same three ZNFm of "zinc finger protein 484" (UNIPROT: Q5JVG2. PDB: 2YTO, 2EMF and 2EMG). Not conservation was observed in the determinant residues of the specific binding to DNA (positions -1, 2, 3 and 6). Nomenclature: + positive residue, – Negative residue, *a* Aromatic residue, *h* hydrophobic residue, *o* Ser/Thr, ! neutral residue. E Value, identities and similarities from BLASTp result

amount of PCR product after the amplification of *tvznf1* was 1.4 higher than the one obtained from parasites cultured in the absence of the cation (Fig. 4b).

Recombinant TvZNF1 (rTvZNF1) recognizes MRE DNA sequence

For demostrating, that rTvZNF1 is capable to bind to DNA, first, the recombinant protein was purified by affinity chromatography (~ 33 kDa, 90% purity) (Fig.S1) and used for EMSA assay with the DNA probes MRE2 and NRE as negative control. MRE2 is recognized by MTF-1, a transcription factor regulated by Zn^{2+} in *H. sapiens* (Hardyman et al. 2016) and NRE is a non-related element corresponding to a fragment of the promoter of EHI_169240 gene of *E. histolytica*. The EMSA assay showed a strong band by the formation of an rTvZNF1-MRE2 complex under Zn^{2+} presence (Fig. 5, lane 1). It confirmed the binding to a specific DNA signature associated with the formation of a stable TvZNFm where the Zn^{2+} is coordinated by Cys and His residues. However, rTvZNF1 in presence of NRE did not produce the formation of any complex (Fig. 5, lane 2). MRE2 in presence of unrelated protein (rGST) or NRE did not promote the formation of a DNA–protein complex (Fig. 5 lane 3 and lane 4, respectively). This result demonstrated the specificity of the protein towards MRE signature.

Discussion

T. vaginalis have one of the biggest genome with 160 Mb and contains a large number of genes. Around of *T. vaginalis* 60, 000 genes are predicted as protein-coding genes. Some of them codify Zn^{2+}-dependent proteins which have been widely found in the parasite including a putative ATP-dependent Zn^{2+} protease (TVAG_ 212500), and ZNF (Carlton et al. 2007). In this work, we characterized a previously annoted ZNF from *T. vaginalis*, which was named here as TvZNF1.

The general organization of TvZNF1 deduced by sequence analysis evidenced a probable cluster of 4 paired C_2H_2-ZNFm or a kind of a multiple-adjacent C_2H_2-ZNFm. The C_2H_2-ZNFm proteins are known to bind to DNA/RNA signatures and have several rearrangements determinate by the number of conserved ZNFm able to interact with DNA (Garton et al. 2015). One type of domain with a triple C_2H_2-ZNFm, a second type of multiple-adjacent C_2H_2-ZNFm, and a third type with two or three separated-paired-C_2H_2-ZNFm have been reported. This type of ZNF contains

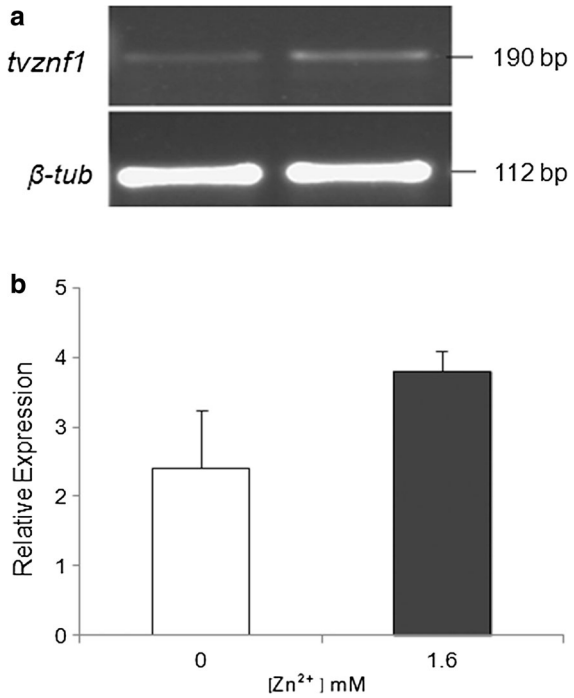

Fig. 4 Expression analysis of *tvznf1* in control and Zn^{2+} conditions. **a** RT-PCR assay of *tvznf1* gene from *T. vaginalis* HGMN01 isolate grown in control (0 mM) and high concentrations of Zn^{+2} (1. 6 mM) showed the expression of the gene (190 bp) and its upregulation by the cation. β-tubulin gene fragment amplification was used as control. **b** Densitometric analysis of relative expression profile of *tvznf1*. Two independent assays were performed by triplicate with similar results

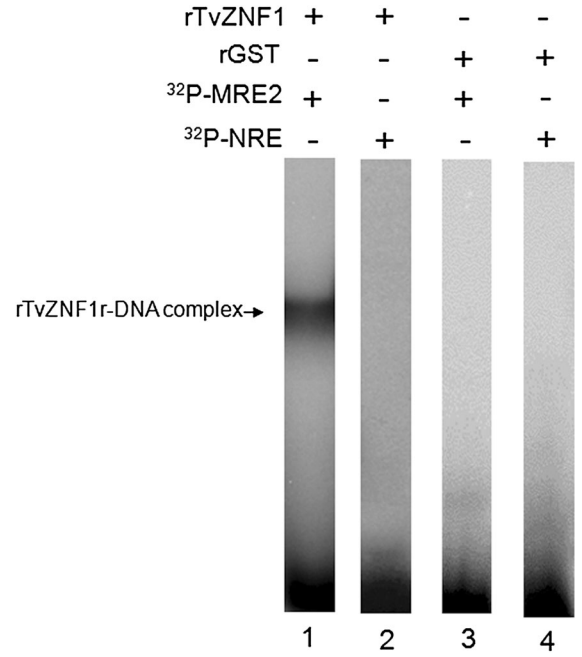

Fig. 5 rTvZNF1 bind to DNA. EMSA assay using MRE2 from *H. sapiens* as probe (lane 1) and ^{32}P MRE2 incubated with purified rTvZNF1 (lane 2). As negative control, we used rGST as unrelated protein (lane 4) and as unrelated probe, we used ^{32}P NRE DNA probe from *E. histolytica* (lane 5). The only complex formed in the assays was observed with MRE2 probe incubated with rTvZNF1

C_2H_2-ZNFm at similar intervals as, MTF-1, TFIIIA, WTI and p43 (Iuchi 2001). This C_2H_2-ZNFm contains Lys and Arg residues in the linkers for contacting the backbone of DNA. Interestingly, TvZNF1 has four linkers and three loops, but the linkers two, four and the loop one (L1) lack of Lys and Arg. The loop L1 separates the TvZNFm2 and 3 suggesting a protein–protein interaction subdomain. The loops L2 and L3 are also linkers but with additional residues. The variation of these residues could condition the properties of binding by modifying the signature recognition processes. TvZNF1 has in the last 25 residues of the C-terminal, a possible small subdomain with unknown function, but it could be is a second protein–protein interaction subdomain.

TvZNF1 protein conserves eight C_2H_2-ZNFm with the canonical signature of consensus pattern $X_2CX_{2\text{-}4}CX_{12}HX_{3\text{-}5}H$. C_2H_2-ZNFm has important cellular functions as splicing factors, DNA-damage repair proteins (Ohyama et al. 2013), or transcriptional factor like the"zinc finger protein 473/Zfp-100″ (PDB: 2EOX) involved in transcriptional regulation (Dominski et al. 2002; Wagner and Marzluff 2006; Wagner et al. 2006) or also in response to microenvironmental stimulus (Hogstrand et al. 2008; Zhao et al. 1998). Interestingly, TvZNF1 also presents these conserved residues in their structure and we could build a secondary signature as [**F/Y/H**]-X-C-$X_{2\text{-}4}$-C-X_3-[**F/Y**]-X_5-**L**-X_2-H-$X_{2\text{-}5}$-H with the aromatic residues F, Y and H that represent a specific scaffold-subset of classical C_2H_2-ZNFm as transcriptional factors. The C_2H_2-ZNFm shared a high variability in their residues, but aromatic residues (F/Y/H) and L were conserved in three specific positions of the motifs. The structural–functional conservation of these internal positions of the C_2H_2-ZNFm of TvZNF1 might be associated with the stability of the C_2H_2-ZNFm as it was reported for structures with the same consensus pattern (Berg 1988; Miller et al. 1985; Payre and Vincent 1988).

The binding of C_2H_2-ZNFm to DNA/RNA could play an important role as sensing system and

transcriptional modulator in multiple physiological processes. For example, in similar manner to other ZNFs such as the PR domain zinc finger protein 9 (PRDM9) which is involved in the localization of recombination hotspots in mammals (Patel et al. 2016) and growth factor independent 1 transcriptional repressor, Gfi-1, a zinc finger protein repressor of cell proliferation and differentiation in hematopoietic stem cells (Lee et al. 2010). TvZNF1 conserved C_2H_2-ZNFm in tandem as poly-ZNF similar to MTF-1 and Zap-1, (Emerson and Thomas 2009), but with additional large domains on the N- or C-terminal with non-integrated signatures. DNA-binding proteins and core promoter elements in some genes has been also reported in parasitic protozoan, including *T. vaginalis* (Horváthová et al. 2012; Tangsombatvichit et al. 2015). However, there is not ZNF with DNA-binding function or as transcriptional factor identified yet in this parasite (Esposito et al. 2006).

The comparison of C_2H_2-ZNFm(s) sequences between TvZNF1 and MTF-1 suggested that the binding of MRE is achieved by a different DNA-recognition mechanism of the same nucleotides sequences by the use of a different set of determinant residues for the specific binding to DNA. MTF-1 and TvZNF1 could bind to different subsites in the same DNA signature, as a convergence of functionality to modulate the response to Zn^{2+}. The size of the loops between TvZNFm in TvZNF1 also suggests a different docking over the DNA signature and different distances that could be favored to bind to the MRE respect to MTF-1. The TvZNFm6 and 7 of TvZNF1 contained a particular combination of determinant residues and probably also recognized different subsites of the same MRE signature.

We found in *T. vaginalis* the conservation of the DNA signature associated with a transcriptional modulation in response to microenvironmental variations of cations in the urogenital tract of the men (Krieger 1995; Krieger and Rein 1982; Mitteregger et al. 2012). For example, the Zn^{2+} regulates the protein expression of the metalloproteinase, TvMP50 (Quintas-Granados et al. 2013) or the Fe^{2+} regulates the expression of Myb proteins (Horváthová et al. 2012).

MTF-1 from *Danio rerio* recognizing MRE sequences in the 5′-UTR of genes in response to Zn^{2+} (Hogstrand et al. 2008) and Zap1 of *S. cerevisae* are able to bind to a conserved sequence of the zinc response element (ZRE) and the protein increase until 90-fold their expression in low Zn^{2+} concentrations (Bird et al. 2000; Zhao et al. 1998).

Interestingly, TvZNF1 and MTF-1 are not orthologous, but paralogous proteins of the same family recruited with a similar function, which suggest that the mechanism of MRE recognition might be even more ancestral that known proteins involved in this mechanism (Tadepally et al. 2008). MTF-1 sequences can be recognized easily using BLASTp in the Phylum Chordata, but in others Phyla, the MTF-1 homologues identification result with low significant values. It suggests different origin of MTF-1 respect to *T. vaginalis* TvZNF1 protein, which is a highly divergent mitochondrion-related organism in the anaerobic protozoan parasites (Makiuchi and Nozaki 2014). The specific proteins lineage of TvZNF1 apparently is not conserved on other taxonomical Phyla different to Parabasalia. However, proteins with C_2H_2-ZNFm could be present in others protozoans such as *Ciliophora*, *Apicomplexa* and *Mycetozoa* with similar number of C_2H_2-ZNFm. This evidence strongly suggests that TvZNF1 could be a transcriptional factor involved in modulation of the response to microenvironmental stimulus by Zn^{2+}.

TvZNF1 binds to MRE probe using three zinc fingers motifs domain, equivalently as been described for MTF-1, but with a different functional solution of residues. TvZNF1 presented a particular combination of residues to bind specifically to the signature MRE. MTF-1 is involved in Zn^{2+} intracellular homeostasis, recognizing a core promoter element sequence, with the nucleotides consensus core signature 5′-TGCRCNC-3′ (Langmade et al. 2000) conserved, and TvZNF1 could bind to a probe sequence with these nucleotides 5′-TGCACACGC-3′ (Lichtlen et al. 2001).

The RT-PCR assays demonstrated the expression of *tvznf1* gene is upregulated by Zn^{2+} presence. The prediction of TvZNF1 peptide sequences revealed eight C_2H_2-ZNFm with DNA binding functions, demonstrated by the rTvZNF1 recognition of the DNA probe MRE2. It contains the core sequence signature of MRE, which is the target of MTF-1 (Hardyman et al. 2016; Rutherford and Bird 2004). The binding capability of these motifs was demonstrated by the bind of rTvZNF1 to the MRE core sequence signature of the radiolabeled probe MRE2.

This is the first report of one ZNF protein with DNA interaction mediated by Zn^{2+} in *T. vaginalis* that may be involved in transcriptional regulation process.

Acknowledgements This work was supported by UACM and a Grant from CONACYT Mexico (83808) to M.E.A.S., J. L. V was supported by a scholarship from ICyTDF (ICyTDF/SRI/65/2011) and COMECYT (14BTD0104). We appreciate the technical assistance of Laura Vázquez-Carrillo and Alfredo Padilla-Barberi.

Compliance of ethical standards

Conflict of interest The authors declare that they have no conflict of interest regarding the publication of this article.

References

Altschul SF, Madden TL, Schäffer AA, Zhang J, Zhang Z, Miller W, Lipman DJ (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 25:3389–3402

Alva V, Nam S-Z, Söding J, Lupas AN (2016) The MPI bioinformatics Toolkit as an integrative platform for advanced protein sequence and structure analysis. Nucleic Acids Res 44:W410–W415

Aurrecoechea C et al (2008) GiardiaDB and TrichDB: integrated genomic resources for the eukaryotic protist pathogens Giardia lamblia and Trichomonas vaginalis. Nucleic Acids Res 37:D526–D530

Berg JM (1988) Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. Proc Natl Acad Sci USA 85:99–102

Bernstein FC et al (1977) The protein data bank. FEBS J 80:319–324

Bird AJ et al (2000) A dual role for zinc fingers in both DNA binding and zinc sensing by the Zap1 transcriptional activator. EMBO J 19:3704–3713

Bittel DC, Smirnova IV, Andrews GK (2000) Functional heterogeneity in the zinc fingers of metalloregulatory protein metal response element-binding transcription factor-1. J Biol Chem 275:37194–37201

Böhm S, Frishman D, Mewes HW (1997) Variations of the C2H2 zinc finger motif in the yeast genome and classification of yeast zinc finger proteins. Nucleic Acids Res 25:2464–2469

Carlton JM et al (2007) Draft genome sequence of the sexually transmitted pathogen *Trichomonas vaginalis*. Science 315:207–212

Chen X, Chu M, Giedroc DP (1999) MRE-Binding transcription factor-1: weak zinc-binding finger domains 5 and 6 modulate the structure, affinity, and specificity of the metal-response element complex. Biochemistry 38:12915–12925

Chou PY, Fasman GD (1974) Prediction of protein conformation. Biochemistry 13:222–245

Combet C, Blanchet C, Geourjon C, Deleage G (2000) NPS@: network protein sequence analysis. Trends Biochem Sci 25:147–150

Ding G, Lorenz P, Kreutzer M, Li Y, Thiesen H-J (2008) SysZNF: the C_2H_2 zinc finger gene database. Nucleic Acids Res 37:D267–D273

Dominski Z, Erkmann JA, Yang X, Sànchez R, Marzluff WF (2002) A novel zinc finger protein is associated with U7 snRNP and interacts with the stem–loop binding protein in the histone pre-mRNP to stimulate 3′-end processing. Genes Dev 16:58–71

Emerson RO, Thomas JH (2009) Adaptive evolution in zinc finger transcription factors. PLoS Genet 5:e1000325

Englbrecht CC, Schoof H, Böhm S (2004) Conservation, diversification and expansion of C_2H_2 zinc finger proteins in the *Arabidopsis thaliana* genome. BMC Genom 5:39

Esposito S et al (2006) A novel type of zinc finger DNA binding domain in the *Agrobacterium tumefaciens* transcriptional regulator Ros. Biochemistry 45:10394–10405

Figueroa-Angulo EE et al (2012) The effects of environmental factors on the virulence of *Trichomonas vaginalis*. Microb Infect 14:1411–1427

Frishman D, Argos P (1995) Knowledge-based protein secondary structure assignment. Proteins 23:566–579

Garton M, Najafabadi HS, Schmitges FW, Radovani E, Hughes TR, Kim PM (2015) A structural approach reveals how neighbouring C2H2 zinc fingers influence DNA binding specificity. Nucleic Acids Res 43:9147–9157

Harb OS, Roos DS (2015) The eukaryotic pathogen databases: a functional genomic resource integrating data from human and veterinary parasites. Parasit Genom Protoc 120:1–18

Hardyman J et al (2016) Zinc sensing by metal-responsive transcription factor 1 (MTF1) controls metallothionein and ZnT1 expression to buffer the sensitivity of the transcriptome response to zinc. Metallomics 8:337–343

Hildebrand A, Remmert M, Biegert A, Söding J (2009) Fast and accurate automatic structure prediction with HHpred. Proteins 77:128–132

Hogstrand C, Zheng D, Feeney G, Cunningham P, Kille P (2008) Zinc-controlled gene expression by metal-regulatory transcription factor 1 (MTF1) in a model vertebrate, the zebrafish. Portland Press Limited, London

Horváthová L et al (2012) Transcriptomic identification of iron-regulated and iron-independent gene copies within the heavily duplicated *Trichomonas vaginalis* genome. Genom Biol Evol 4:1017–1029

Iuchi S (2001) Three classes of C_2H_2 zinc finger proteins. Cell Mol Life Sci 58:625–635

Klug A (2010) The discovery of zinc fingers and their development for practical applications in gene regulation and genome manipulation. Quart Rev Biophys 43:1–21

Krieger JN (1995) Trichomoniasis in men: old issues and new data. Sex Transm Dis 22:83–96

Krieger JN, Rein MF (1982) Zinc sensitivity of *Trichomonas vaginalis*: in vitro studies and clinical implications. J Infect Dis 146:341–345

Krishna SS, Majumdar I, Grishin NV (2003) Structural classification of zinc fingers: survey and summary. Nucleic Acids Res 31:532–550

Laity JH, Lee BM, Wright PE (2001) Zinc finger proteins: new insights into structural and functional diversity. Curr Opin Struct Biol 11:39–46

Langmade SJ, Ravindra R, Daniels PJ, Andrews GK (2000) The transcription factor MTF-1 mediates metal regulation of the mouse ZnT1 gene. J Biol Chem 275:34803–34809

Lau AO, Smith AJ, Brown MT, Johnson PJ (2006) *Trichomonas vaginalis* initiator binding protein (IBP39) and RNA polymerase II large subunit carboxy terminal domain interaction. Mol Biochem Parasitol 150:56–62

Lee S, Doddapaneni K, Hogue A, McGhee L, Meyers S, Wu Z (2010) Solution structure of Gfi-1 zinc domain bound to consensus DNA. J Mol Biol 397:1055–1066

León-Sicairos CR, Leon-Felix J, Arroyo R (2004) tvcp12: a novel *Trichomonas vaginalis* cathepsin L-like cysteine proteinase-encoding gene. Microbiology 150:1131–1138

Lichtlen P, Wang Y, Belser T, Georgiev O, Certa U, Sack R, Schaffner W (2001) Target gene search for the metal-responsive transcription factor MTF-1. Nucleic Acids Res 29:1514–1523

Liston DR, Lau AO, Ortiz D, Smale ST, Johnson PJ (2001) Initiator recognition in a primitive eukaryote: IBP39, an initiator-binding protein from *Trichomonas vaginalis*. Mol Cell Biol 21:7872–7882

Makiuchi T, Nozaki T (2014) Highly divergent mitochondrion-related organelles in anaerobic parasitic protozoa. Biochimie 100:3–17

Marchler-Bauer A et al (2014) CDD: NCBI's conserved domain database. Nucleic Acids Res 43:D222–D226

Miller J, McLachlan A, Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from *Xenopus* oocytes. EMBO J 4:1609

Mitteregger D, Aberle SW, Makristathis A, Walochnik J, Brozek W, Marberger M, Kramer G (2012) High detection rate of *Trichomonas vaginalis* in benign hyperplastic prostatic tissue. Med Microbiol Immunol 201:113–116

Mulder NJ et al (2005) InterPro, progress and status in 2005. Nucleic Acids Res 33:D201–D205

Ohyama Y et al (2013) Characterization of AtSTOP1 orthologous genes in tobacco and other plant species. Plant Physiol 162:1937–1946

Patel A, Horton JR, Wilson GG, Zhang X, Cheng X (2016) Structural basis for human PRDM9 action at recombination hot spots. Genes Dev 30:257–265

Payre F, Vincent A (1988) Finger proteins and DNA-specific recognition: distinct patterns of conserved amino acids suggest different evolutionary modes. FEBS Lett 234:245–250

Quintas-Granados LI, Villalpando JL, Vázquez-Carrillo LI, Arroyo R, Mendoza-Hernández G, Álvarez-Sánchez ME (2013) TvMP50 is an immunogenic metalloproteinase during male trichomoniasis. Mol Cell Proteomics 12:1953–1964

Razin S, Borunova V, Maksimenko O, Kantidze O (2012) Cys2His2 zinc finger protein family: classification, functions, and major members. Biochemistry 77:217–226

Rutherford JC, Bird AJ (2004) Metal-responsive transcription factors that regulate iron, zinc, and copper homeostasis in eukaryotic cells. Eukaryot Cell 3:1–13

Schumacher MA, Lau AO, Johnson PJ (2003) Structural basis of core promoter recognition in a primitive eukaryote. Cell 115:413–424

Tadepally HD, Burger G, Aubry M (2008) Evolution of C_2H_2-zinc finger genes and subfamilies in mammals: species-specific duplication and loss of clusters, genes and effector domains. BMC Evol Biol 8:176

Tangsombatvichit P, Semkiv MV, Sibirny AA, Jensen LT, Ratanakhanokchai K, Soontorngun N (2015) Zinc cluster protein Znf1, a novel transcription factor of non-fermentative metabolism in *Saccharomyces cerevisiae*. FEMS Yeast Res 15:fou002

Theunissen O, Rudt F, Guddat U, Mentzel H, Pieler T (1992) RNA and DNA binding zinc fingers in *Xenopus* TFIIIA. Cell 71:679–690

Vazquez Carrillo LI, Quintas Granados LI, Arroyo R, Mendoza Hernández G, González Robles A, Carvajal Gamez BI, Alvarez Sánchez ME (2011) The effect of Zn^{2+} on prostatic cell cytotoxicity caused by *Trichomonas vaginalis*. J Integr OMICS 1:198–210

Wagner EJ, Marzluff WF (2006) ZFP100, a component of the active U7 snRNP limiting for histone pre-mRNA processing, is required for entry into S phase. Mol Cell Biol 26:6702–6712

Wagner EJ, Ospina JK, Hu Y, Dundr M, Matera AG, Marzluff WF (2006) Conserved zinc fingers mediate multiple functions of ZFP100, a U7snRNP associated protein. Rna 12:1206–1218

Zhao H, Butler E, Rodgers J, Spizzo T, Duesterhoeft S, Eide D (1998) Regulation of zinc homeostasis in yeast by binding of the ZAP1 transcriptional activator to zinc-responsive promoter elements. J Biol Chem 273:28713–28720