FEBS Letters 587 (2013) 1474–1481

journal homepage: www.FEBSLetters.org

Genetic and epigenetic mutations affect the DNA binding capability of human ZFP57 in transient neonatal diabetes type 1

Ilaria Baglivo [a,1], Sabrina Esposito [a,1], Lucia De Cesare [a], Angela Sparago [a,b], Zahra Anvar [b], Vincenzo Riso [b], Marco Cammisa [b], Roberto Fattorusso [a], Giovanna Grimaldi [b], Andrea Riccio [a,b,*], Paolo V. Pedone [a,*]

[a] Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Seconda Università degli Studi di Napoli, Via Vivaldi 43, 81100 Caserta, Italy
[b] Istituto di Genetica e Biofisica A. Buzzati-Traverso, CNR, Via Pietro Castellino 111, 80131 Napoli, Italy

ARTICLE INFO

Article history:
Received 12 November 2012
Revised 31 January 2013
Accepted 20 February 2013
Available online 13 March 2013

Edited by Ivan Sadowski

Keywords:
ZFP57
Transient neonatal diabetes
Imprinting
Hypomethylation
DNA binding

ABSTRACT

In the mouse, ZFP57 contains three classical Cys_2His_2 zinc finger domains (ZF) and recognizes the methylated $TGC^{met}CGC$ target sequence using the first and the second ZFs. In this study, we demonstrate that the human ZFP57 (hZFP57) containing six Cys_2His_2 ZFs, binds the same methylated sequence through the third and the fourth ZFs, and identify the aminoacids critical for DNA interaction. In addition, we present evidences indicating that hZFP57 mutations and hypomethylation of the TNDM1 ICR both associated with Transient Neonatal Diabetes Mellitus type 1 result in loss of hZFP57 binding to the TNDM1 locus, likely causing *PLAGL1* activation.

1. Introduction

Transient Neonatal Diabetes Mellitus type 1 (TNDM1; OMIM 601410) is a rare form of diabetes, in which hyperglycaemia remits at a median age of approximately 3 months, but in half of the cases it recurs in adolescence or adulthood. It is caused by overexpression of the imprinted genes *PLAGL1* and *HYMAI*[1,2] on human chromosome 6q24. The expression of these genes is normally restricted to the paternal allele as a result of maternal DNA methylation of the promoter region (TNDM1 imprinted control region or TNDM1 ICR). TNDM1 is not associated with mutations of *PLAGL1* or *HYMAI*, but rather with their overexpression via three known mechanisms: (1) paternal uniparental disomy of chromosome 6 (40%), (2) duplication of the imprinted TNDM1 region on chromosome 6q24 (32%), or (3) relaxation of imprinting with maternal hypomethylation of the TNDM1 ICR in the absence of any detectable DNA or chromosome anomaly *in cis* (28%). Despite their common epimutation, patients with hypomethylation of the TNDM1 ICR are clinically heterogeneous and can be divided in two subgroups: about 50% have no other epimutations while in the others loss of methylation at 6q24 is part of a complex hypomethylation at multiple imprinted loci (hypomethylation at imprinted loci, HIL). The majority of these TNDM1 HIL cases were found to have recessive loss of function mutations in the gene encoding the zinc finger transcription factor ZFP57 on human chromosome 6p22.3 [3].

The human ZFP57 (hZFP57) is a protein of 516 amino acids (the mouse homologue has 421 amino acids), including a predicted Kruppel-associated box (KRAB) domain and six complete classical Cys_2His_2 zinc fingers (ZFs); in contrast only three of the classical Cys_2His_2 ZFs are present in the mouse protein [4], with the first and second ZFs of the mouse protein being homologous to the second and third of the human.

Loss of ZFP57 function in the developing mouse zygote is partially lethal, while eliminating both maternal and zygotic functions of ZFP57 causes complete embryonic lethality [5]. The Zfp57 knockout mouse model confirms that in the mouse, as in human, the protein is required for post-fertilization maintenance of the parent of origin-specific methylation of the ICRs. Human ZFP57 mutations affect several maternally methylated ICRs while inactivation of the mouse *Zfp57* causes imprinting defects at both maternally and paternally methylated loci. In mice, absence of ZFP57 also results in defective establishment of maternal methylation at the imprinted *Snrpn* locus [3,5].

* Corresponding authors. Address: Dipartimento di Scienze e Tecnologie Ambientali, Biologiche e Farmaceutiche, Seconda Università degli Studi di Napoli, Via Vivaldi 43, Caserta, 81100 Italy. Fax: +39 081 6132706 (A. Riccio), +39 0823 274605 (P.V. Pedone).
E-mail addresses: andrea.riccio@unina2.it (A. Riccio), paolov.pedone@unina2.it (P.V. Pedone).
[1] These authors have contributed equally to this work.

0014-5793
http://dx.doi.org/10.1016/j.febslet.2013.02.045

In mouse Embryonic Stem Cells (ESCs), ZFP57 binds all known ICRs where it recruits the corepressor KAP1, the H3K9 methyltransferase SETDB1 and the heterochromatin protein HP1-gamma [6]. This protein complex was shown to be required for maintaining the parent of origin-specific DNA methylation, histone H3K9me3 and gene expression at many of these loci [6–8]. Recruitment of ZFP57/KAP1 at the ICRs is mediated by ZFP57 recognition of a methylated TGC^{met}CGC hexanucleotide motif; the first and second Cys_2His_2 ZFs of the mouse protein are sufficient for the high affinity recognition of this DNA target site [6].

Here, we demonstrate that similarly to the mouse first and second ZFs, the third and fourth ZFs of hZFP57 specifically recognizes the TGC^{met}CGC hexanucleotide. In addition, we define the zinc finger residues that are important for hZFP57–DNA interaction and demonstrate that two hZFP57 mutations (R248H and H277N) associated with TNDM1 affect the binding to methylated DNA. By showing that in the mouse ZFP57 represses *Plagl1* expression and that hZFP57 binds in vitro three DNA sequences of the TNDM1 ICR in a DNA methylation-dependent manner, we indicate a molecular mechanism by which genetic and epigenetic mutations cause TNDM1.

2. Materials and methods

2.1. Cloning and Purification of the Proteins

DNA fragments encoding the different segments of the hZFP57 protein were generated by PCR from cDNA hIPS. Oligonucleotides were synthesized on the basis of the published sequence (NP_001103279.2) which replaced the original XP_294093 which lacks 20 residues at the N-terminus. The following oligonucleotides were used as primers: primer 1, 5′-CGGGATCCGGGCCACCCTTTTTTGCTACACCTGTGGCAAATG-3′, and primer 2, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATC-3′, for hZFP57zf1-4 (residues 172–287). In order to generate the hZFP57zf3-4 (residues 227–287), we substituted the Cys232 of the protein with a phenylalanine, present in the mouse protein, by PCR-mediated mutagenesis [9]. This substitution increased the stability and easy of handling of the protein without changing its DNA binding properties (Supplementary Fig. 1) and so we used it in the hZFP57zf3-4 protein and in all the mutants. For cloning of the coding sequence for hZFP57zf3-4, we used primer 3, 5′-CGGGATCCGGGGAGAGGCCCTTCTTTTGCACGCTCTG-3′, and primer 2. The plasmid including the coding sequence of the hZFP57zf3-4 was used as a template to generate by PCR-mediated mutagenesis the ZF3 point mutants of position −1 (D242A), +2 (S244A), +3 (G245A) and +6 (R248A) and the ZF 4 point mutants in position −1 (D270A), +2 (S272A), +3 (E273A) and +6 (R276A). The primers 4, 5′-CGGGATCCGGGGAGAGGCCCTTCTTTTGCACGCTCTGTGACAAGACCTACTGTGCTGCTTCTGGAC-3′, and primer 2 were used to generate $_{hZFP57zf3-4}$D242A; primer 5, 5′-CGGGATCCGGGGAGAGGCCCTTCTTTTGCACGCTCTGTGACAAGACCTACTGTGATGCTGCTGGACTAAGTC-3′, and primer 2 to generate $_{hZFP57zf3-4}$S244A; primer 6, 5′-CGGGATCCGGGGAGAGGCCCTTCTTTTGCACGCTCTGTGACAAGACCTACTGTGATGCTTCTGCACTAAGTCGTC-3′, and primer 2 to generate $_{hZFP57zf3-4}$G245A; primer 7, 5′-CGGGATCCGGGGAGAGGCCCTTCTTTTGCACGCTCTGTGACAAGACCTACTGTGATGCTTCTGGACTAAGTGCTCACCGCCGCG-3′, and primer 2 to generate $_{hZFP57zf3-4}$R248A; primer 3 and primer 8, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATCTTCTGGTGGCGTTTGAGCTCAGACTGGGCCCGGAAGCTC-3′, for $_{hZFP57zf3-4}$D270A; primer 3 and primer 9, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATCTTCTGGTGGCGTTTGAGCTCAGCCTGGTCCCGG-3′, for $_{hZFP57zf3-4}$S272A; primer 3 and primer 10, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATCTTCTGGTGGCGTTTGAGCGCAGACTGGTCC-3′, for $_{hZFP57zf3-4}$E273A; primer 3 and primer 11, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATCTTCTGGTGGGCTTTGAGCTCAG-3′, for $_{hZFP57zf3-4}$R276A. Primer 1, primer R248H for, 5′-GGACTAAGTCATCACCGCCGCGTCCATCTGGGTTACCGG-3′, primer R248H rev 5′-GATGGACGCGGCGGTGATGACTTAGTCC AGAAGC-3′ and primer 2 were used to generate by a double-step PCR-mediated mutagenesis reaction the TND mutant $_{hZFP57zf1-4}$R248H; primer 1 and primer 12, 5′-AACATGGTCGACTCACACTGGCTCCTGGTTTTGGTGTATCTTCTGGTTGCGTTTGAGC-3′, for the TND mutant $_{hZFP57zf1-4}$H277N.

All of the PCR products were digested with the restriction enzymes *Bam*HI and *Sal*I and cloned into a *Bam*HI/*Sal*I-digested pGEX-4T-3 (GE Healthcare) expression vector and the proteins were expressed as a glutathione S-transferase (GST) fusion. The fusion proteins were expressed in the *Escherichia coli* BL21 host strain, induced for 2 h in the presence of 1 mM IPTG at 37 °C. The GST fusion proteins were purified as previously published [10].

2.2. Electrophoretic Mobility Shift Analysis (EMSA)

10 pmol of the purified proteins were incubated for 10 min on ice with 2.5 pmol of the specified duplex oligonucleotides in the presence of 25 mM HEPES (pH 7.9), 50 mM KCl, 6.25 mM $MgCl_2$, 1% NP-40 and 5% glycerol. After incubation, the mixture was loaded on a 5% polyacrylamide gel (29:1 acrylamide:bisacrylamide ratio) and run in 0.5% TBE (200 V for 75 min). The gels were than stained with SYBR Green (Invitrogen) and acquired with Typhoon Trio^{++} scanner (GE Healthcare). The 5′-TGCCGC-3′ DNA probe used in the EMSA experiments was obtained by annealing the following single strands oligonuceotides: 5′-GTAATATTGCCGCAGTAATA-3′ and 5′-TATTACTGCGGCAATATTAC-3′; the 5′-TGC^{met}CGC-3′ DNA probe was obtained by annealing the following methylated single strands oligonucleotides: 5′-GTAATATTGC^{met}CGCAGTAATA-3′ and 5′-TATTACTG^{met}CGGCAATATTAC-3′.

2.3. DNA methylation and gene expression analyses

1 μg genomic DNA extracted from mouse wild-type ESCs (line A3) and *Zfp57* −/− ESCs (clones 2E9 & 1A12, [11,12] was treated using the EpiTect Bisulfite Kit (Qiagen). Sodium bisulfite converts unmethylated cytosine residues to uracil, while methylated cytosine is unchanged. The *Plagl1* ICR (mm9 Chrm10:12810201-12811599) was amplified by PCR using bisulfite converted DNA as template. The primers used were as follows: Plagl1_BIS_F1 5′-AYGTAGATTTTTAGATTGAGTTGAGGTTAT-3′ and Plagl1_BIS_R1 5′-CCTACCAAACRAATAAAATTAAAAAAC-3′. The PCR products were cloned in Topo pCR2.1 vector (Topo-TA cloning kit, Invitrogen) and the clones sequenced. For expression analysis, we extracted RNA from mouse ESCs lines A3, 2E9 and 1A12, using the Trizol reagent (Invitrogen). cDNA was synthesized using 1 μg of total RNA and Quantitech Reverse Transcription Kit (Qiagen), according to the protocol of the manufacturer. *Plagl1* expression was examined by iQ SYBR Green Supermix (BioRad) quantitative real-time RT-PCR. Reactions were run on a CFX96 Real-Time System + C1000 Thermal Cycler (Biorad) using the default cycling conditions. Relative expression was determined using the ΔΔCT method, and gene expression values were normalized against the expression of GAPDH. The PCR primers were as follows: Plagl1RTcDNA1for: 5′-GCCTCACAGCTGGCTTTCCTGCTCTCAC-3′; Plagl1RTcDNA1rev: 5′-GTGGGGAGTCACTACAGCCTGACTCCAG-3′; GapdhRTFor: 5′-GCAAGGCTTCCGTGCTCTCGAGAGAG-3′; GapdhRTRev: 5′-GCCCGAGGACAATAAGGCTCAAGGGC-3′.

2.4. Chromatin immuno-precipitation (ChIP)

Mouse ES cells were cultured at 80% confluence and cross-linked with 1% formaldehyde for 10 min at room temperature.

Cross-linking was quenched with 125 mM glycine, and whole-cell extracts were prepared for the chromatin immunoprecipitations. The fragmented chromatin from 7×10^6 cells was used for each reaction. DNA was immunoprecipitated with 8 μg of anti-ZFP57 (by Abcam Ab 45341) and with 7 μg of anti-KAP1 (by Abcam Ab 22553). DNA from the immunoprecipitates or from 1% of input was analyzed by real-time PCR using the following primers specific for CpGi 36 and CpGi 72: Plagl1CpG36F: 5′-TCTTTGGGGTCCAAGCAA-3′; Plagl1CpG36R 5′-ACGAGGTCGTTCCCTTGTC-3′; Plagl1CpG 72F: 5′-TGAGTCCGGCTAGGGTAGG-3′; Plagl1CpG72R: 5′-GAGCCTTTTGCTGCATCTCT-3′. The human embryonal carcinoma 2102Ep cells were processed for ChIP using the MAGnify Chromatin Immunoprecipitation System (Life Technologies). Briefly, the cells were fixed with 1% formaldehyde for 10 min at room temperature and the chromatin was sheared by sonication to an average fragment size of 100–500 bp with the Covaris instrument. DNA was immunoprecipitated with 5 μg of anti-KAP1 (by Abcam Ab10483) according to the MAGnify kit protocol. ChIP and input DNA were amplified by qPCR using primers specific for the TNDM1 ICR (CpGi 118): F-GTCACGTGGCAGGAGGAG and R-CTGTACCTGG GCGACCTTG, and for the upstream CpGi 113: F-GGCTAAGCTGGCACTTGAGA and R-GTCTGCAGCTCTTCGCTTCT. A fraction of the input chromatin (1%) was processed together with a control rabbit anti-IgG and qChIP values were given as percent of input.

3. Results

Despite numerous data suggesting similar functions in vivo of human (hZFP57) and mouse ZFP57 (mZFP57) [3,5,13], these proteins share poor sequence identity outside the regions encompassing the KRAB domain (that binds the cofactor KAP1) and the third and fourth human and first and second mouse ZFs (Fig. 1). In proteins containing multiple fingers, it is clear that adequate specificity and affinity can be provided by a few critical fingers, usually separated by a canonical 7 amino acids spacer [14,15]. In hZFP57 the 6 ZFs are divided in two distant clusters of canonically spaced domains, comprising the NH_2-terminal four and the COOH-terminal two ZFs. To define the minimal region able to bind DNA, fragments encoding the four NH_2-terminal ZFs of hZFP57 (hZFP57zf1-4) and ZFs 3 and 4 (hZFP57zf3-4) were cloned into the pGEX-4T-3 expression vector. The proteins were expressed and tested by EMSA on oligonucleotides containing the TGCCGC motif. We observed that hZFP57zf1-4 as well as hZFP57zf3-4 are able to bind this DNA sequence (Fig. 2A and B) with a clear preference for the methylated site, indicating that the two conserved ZFs play a critical role for DNA interaction. Notably, EMSA performed with hemimethylated oligonucleotides showed that the methylated C recognized by hZFP57 is on the reverse strand of the DNA target (Fig. 2B lanes 3 and 4).

All the structural and functional studies conducted on animal zinc finger proteins interacting with DNA suggest that these motifs employ the residue that immediately precedes the α-helix (position −1), as well the second (+2), third (+3) and sixth (+6) residues of the helix to contact the base pairs [14,16]. In order to demonstrate if these residues were also important for the interaction of hZFP57 with methylated DNA, EMSAs were performed with hZFP57zf3-4 containing alanine substitutions of the amino acids in positions −1, +2, +3, +6 of ZF3 (D242, S244, G245 and R248) and ZF4 (D270, S272, E273 ad R276). The results shown in Fig. 3A and B clearly indicate that all the single mutations except E273A reduce protein–DNA binding affinity; a dramatic effect is observed with the ZF3 mutants in position +2 ($_{hZFP57zf3-4}$S244A),

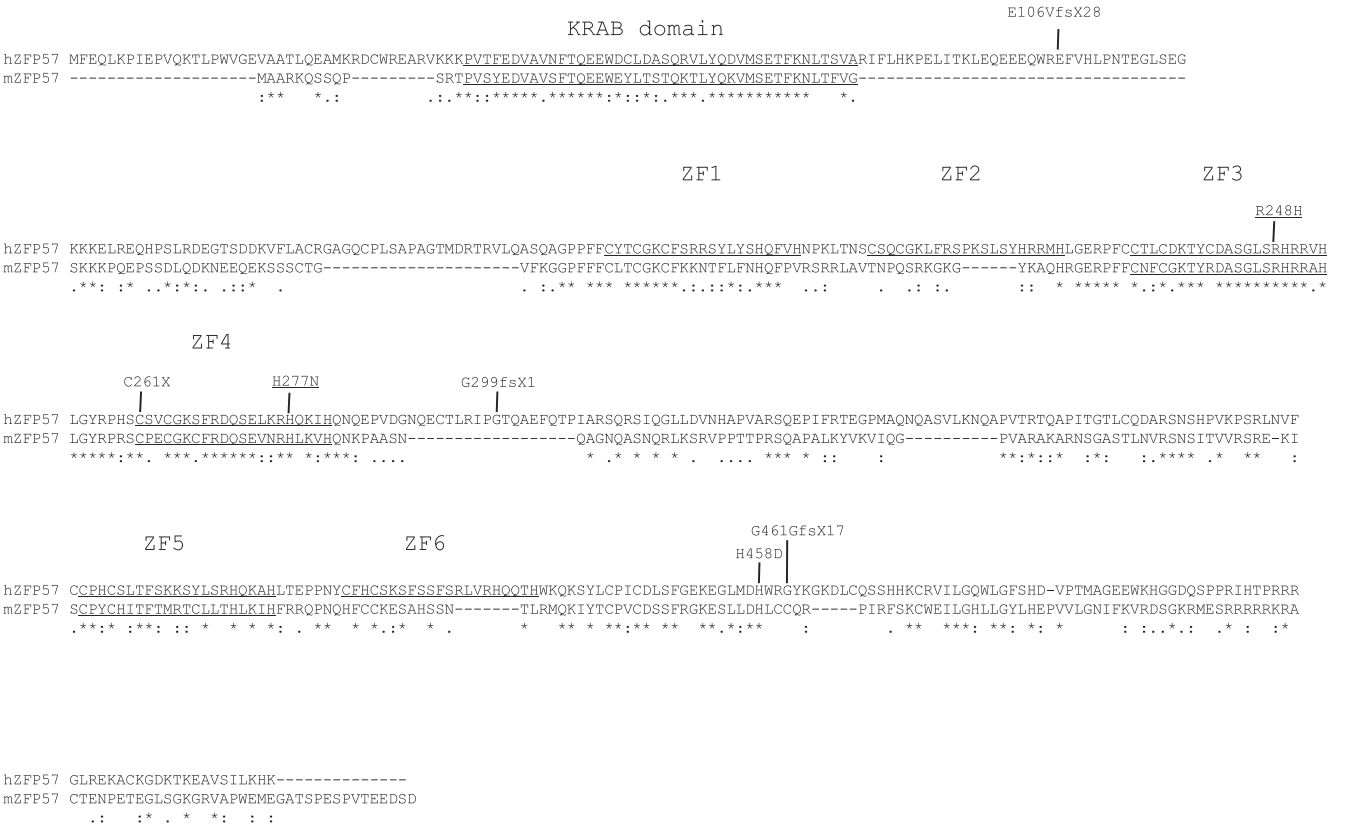


Fig. 1. Alignment of hZfp57 and mZfp57 sequences. The KRAB domain and the Cys_2His_2 ZFs domains are underlined in the two sequences. The six Cys_2His_2 human ZFs domains are indicated. The TNDM1 associated mutations are also indicated. Amino acids identities are indicated by asterisks. Conservative and non-conservative homologies are indicated by double and single dots, respectively.

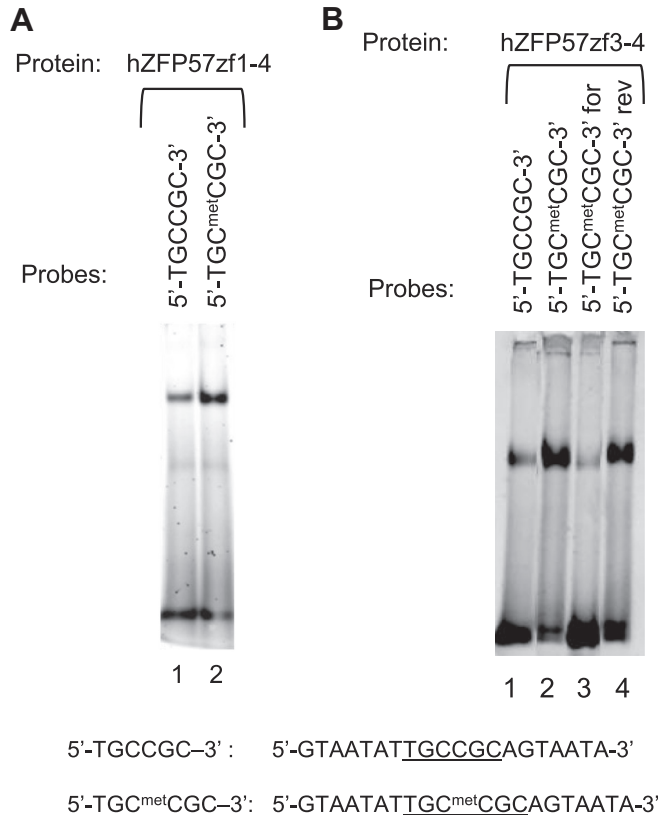


Fig. 2. (A) EMSA analysis of hZFP57zf1-4 DNA binding capability (B) EMSA analysis of hZFP57zf3-4 DNA binding capability. The sequences of the forward strands of the oligonucleotides used as probes are indicated. The 5′-TGC^{met}CGC-3′ for and 5′-TGC^{met}CGC-3′ rev oligonucleotides are methylated only on the forward and reverse strands, respectively.

+3 ($_{hZFP57zf3-4}$G245A) and +6 ($_{hZFP57zf3-4}$R248A) and the ZF 4 mutant in position +6 ($_{hZFP57zf3-4}$R276A).

All the mutants showing a reduced but detectable DNA binding capability have been tested in EMSA for their sensitivity to DNA methylation and demonstrate to prefer the methylated target sites (all the results obtained with the different mutants are summarized in Suppl. Table 1). Interestingly, the ZF4 mutant $_{hZFP57zf3-4}$E273A shows similar affinities toward the methylated TGC^{met}CGC and non-methylated TGCCGC-containing oligonucleotide (Fig. 3C). This result suggests that ZF 4 is involved in the recognition of the methyl-cytosine and that the presence of glutamate in position 273 plays a critical role in reducing the affinity to the unmethylated target.

Mackay and colleagues [3] have shown that patients affected by TNDM1 have mutations in hZFP57 (Fig. 1). Two of these mutations, R248H and H277N, affect residues in ZF 3 and ZF 4 respectively. In order to test the effect of these mutations on hZFP57 DNA binding capability, we have produced mutant versions of hZFP57zf1-4 bearing the mutations R248H and H277N. The EMSA presented in Fig. 4 shows that both of these mutations cause loss of hZFP57 DNA binding capability.

In both mouse and human, *PLAGL1* is transcribed from two CpG island-associated promoters (Fig. 5A and E), of which the downstream one (P1) is methylated and imprinted and corresponds to the ICR, while the upstream one (P2) is unmethylated [17,18]. Consistent with previous results [6], we show that the ZFP57–KAP1 complex binds the imprinted but not the non-imprinted *Plagl1* promoter in mouse ESC (Fig. 5B). In addition, we observed that the *Plagl1* ICR is hypomethylated and *P1agl1* expression derived from the imprinted promoter is up-regulated in mouse *Zfp57* −/− ESCs, indicating that ZFP57 is a repressor of the *Plagl1* gene (Fig. 5C and D). The binding of KAP1 to the human TNDM1 ICR was demonstrated in embryonic kidney HEK293 cells in a ChIPseq study conducted by Farnham and co-workers [19]. By using locus-specific ChIP, we observed that in human embryonal carcinoma 2102Ep

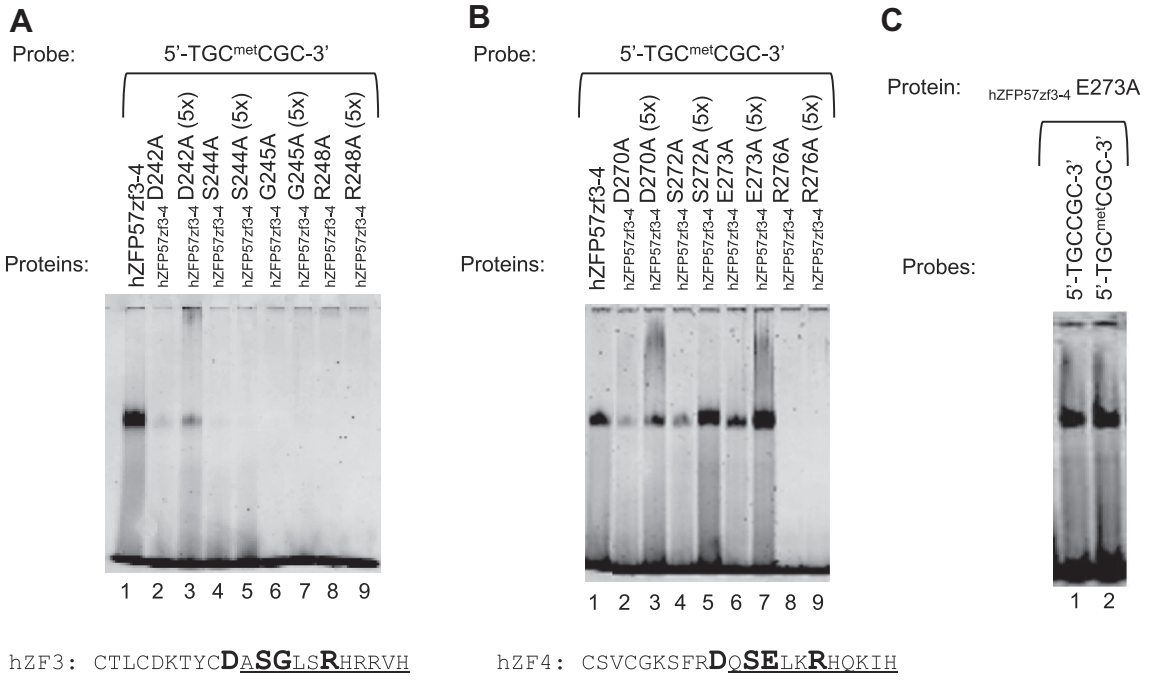
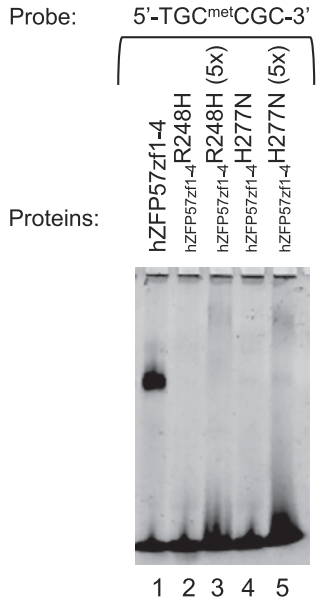


Fig. 3. (A) EMSA DNA binding analysis of hZFP57zf3-4 point mutants at positions −1 ($_{hZFP57zf3-4}$D242A); +2 ($_{hZFP57zf3-4}$S244A), +3 ($_{hZFP57zf3-4}$G245A) and +6 ($_{hZFP57zf3-4}$R248A) of the ZF3 α-helical region. Where indicated (5×) a fivefold excess of the purified proteins was added. (B) EMSA DNA binding analysis of hZFP57zf3-4 point mutants at positions −1 ($_{hZFP57zf3-4}$D270A); +2 ($_{hZFP57zf3-4}$S272A), +3 ($_{hZFP57zf3-4}$E273A) and +6 ($_{hZFP57zf3-4}$R276A) of the ZF4 α-helical region. Where indicated (5×) a fivefold excess of the purified proteins was added. (C) EMSA analysis of the $_{hZFP57zf3-4}$E273A mutant DNA binding sensitivity to cytosine methylation of the TGCCGC target site. The sequences of hZFP57 ZF3 and ZF4 are indicated; the regions which correspond to the α-helices are underlined and the residues which have been mutated to alanine are in bold and in a larger font size.



hZF3: CTLCDKTYCDASGLSRHRRVH

hZF4: CSVCGKSFRDQSELKRHQKIH

Fig. 4. EMSA analysis of the DNA binding capability of the TNDM1 point mutants $_{hZFP571-4}$R248H and $_{hZFP57zf1-4}$H277N. Where indicated (5×) a fivefold excess of the purified proteins was added. The sequences of hZFP57 ZF3 and ZF4 are indicated; the regions which correspond to the α-helices are underlined and the residues which are mutated in TNDM1 are in bold and in a larger font size.

cells KAP1 binds the imprinted *PLAGL1* promoter much more efficiently than the non-imprinted promoter, suggesting an imprinting control mechanism based on ZFP57–KAP1 and similar to the mouse (Fig. 5F).

In order to determine if the TNDM1 ICR is really able to bind hZFP57, we tested by EMSA as DNA targets for hZFP57zf3-4 three DNA sequences containing the TGCCGC motif identified in this region (PLAGL1-1, PLAGL1-2 and PLAGL1-3 see Fig. 5E and Supplementary Fig. 2). The results demonstrate the DNA methylation-dependent binding of hZFP57 to these sequences (Fig. 6). Considering that the PLAGL1-1 and the PLAGL1-2 sites are very close and form a tight palindrome in the *PLAGL1* promoter (see Supplementary Fig. 2) we also tested by EMSA an oligonucleotide containing the palindrome. We found that, hZFP57zf3-4 binds the palindromic repeat and is sensitive to the methylation of the cytosines present in the two hexanucleotides. The methylation of the cytosine in the spacer region does not affect the protein–DNA interaction (Supplementary Fig. 3).

4. Discussion

Mouse and human ZFP57 contain three and six Cys_2His_2 ZFs respectively, two of which, the first and the second of the mouse protein being homologous to the third and fourth of the human protein. Our results on the human ZFP57 protein demonstrate that the sequence extending from glycine 226 to valine 288, including the third and the fourth ZFs, is sufficient for high affinity binding to the methylated DNA target sites. With amino acid substitutions of the canonical α-helical positions −1, +2, +3 and +6 involved in ZF-DNA recognition [14,16], we show which residues of the ZFs 3 and 4 are important for hZFP57–DNA interaction. In addition, we demonstrate that two hZFP57 mutations (R248H and H277N) associated with TNDM1 affect the binding to methylated DNA. Furthermore, we show in mouse ESCs that *Zfp57* inactivation results in loss of DNA methylation and gene activation at the *Plagl1* promoter. Finally, we identify three DNA sequences in the TNDM1 ICR showing DNA methylation-dependent binding to hZFP57, indicating that either mutations in the critical ZF residues or hypomethylation of the target DNA sequences may result in loss of hZFP57 binding and *PLAGL1* activation in TNDM1.

By analyzing single amino acid substitution mutants of the hZFP57 DNA binding domain we demonstrated that substitution of S244, G245 and R248, occupying position +2, +3 and +6 of ZF 3 α-*helix* and R276, located at position +6 of ZF4 helical region, have a dramatic effect on the interaction with DNA. Weaker reduction of the DNA binding capability is observed when D242 (position −1 ZF3), D270 (position −1 ZF4), S272 (position +2 ZF4) are mutated to alanine. While aspartate, serine, glutamate and arginine can act as base determinants, contact the phosphates or alternatively have a structural role to correctly position other residue side chains, a glycine residue, which lacks a side chain, could hardly be involved in a critical contact. For G245 we can then imagine a sort of passive role in DNA recognition in which its presence allows the α-helix to closely approach the nucleic acid. Our mutants analysis also demonstrate that ZF4 residue E273, located in position +3 of the helix, is important to reduce the binding to the unmethylated target sites favoring the preferential recognition of the methylated sequences. While this article was being written, Cheng and co-workers published the structural characterization of mouse ZFP57 ZF1 and ZF2 complexed to the methylated hexanucleotide 5′-TGC^{met}CGC-3′ target site [20]. This study reveals that the methylated DNA target is specifically recognized by residues invariant between mouse and human; those residues are located in the N-terminal portion of the helix and in the preceding loop (positions −2, −1, +2, +3, and +6) of the first and the second mouse ZFP57 zinc fingers. Interestingly, the two methylated cytosines are asymmetrically recognized. One involves hydrophobic interactions with ZF2 Arg176 (that occupies the non-canonical position −2 and corresponds to ZF4 Arg269 in the human ZFP57) and the other requires a layer of water molecules. Overall, the structural and functional findings obtained for the mouse ZFP57 are in very good agreement with the functional results of the human protein here presented, and suggest that ZF4 Arg269 of human ZFP57 play a key role in the specific recognition of the methylated site. Interestingly, ZF1 and ZF2 of human ZFP57 are not conserved in the mouse and are dispensable for high affinity DNA binding to the methylated hexanucleotide 5′-TGC^{met}CGC-3′ target site; further investigations would be needed to clarify the exact role of these extra finger domains, possibly involved in protein–protein interactions or in the recognition of different target sites.

Recently, seven different recessive mutations in hZFP57 have been identified as a cause for variable but stable mosaic hypomethylation at multiple imprinted loci in patients with TNDM1 [3]. Four of these ZFP57 mutations (E106VfsX28, C261X, G299fsX1 and G461GfsX17) generate truncated forms of the protein and three (R248H, H277N and H458D) are missense mutations affecting residues highly conserved among several mammalian species (see Fig. 1). We demonstrated that two of the missense mutations (R248H and H277N), affecting position +6 of ZF 3 and one of the zinc-binding histidines within ZF 4 respectively, alter hZFP57 DNA binding capability. The 2-bp deletion in the KRAB domain (E106VfsX28) causing a frameshift and generating a premature translation termination codon would produce a short protein lacking the entire ZF containing region which certainly is not able to bind DNA. Interestingly, on the basis of our results we can argue that also the truncating substitution C261X, changing the first cysteine of ZF4 into a stop codon, generates a protein unable to recognize the DNA target with adequate affinity because it is missing the

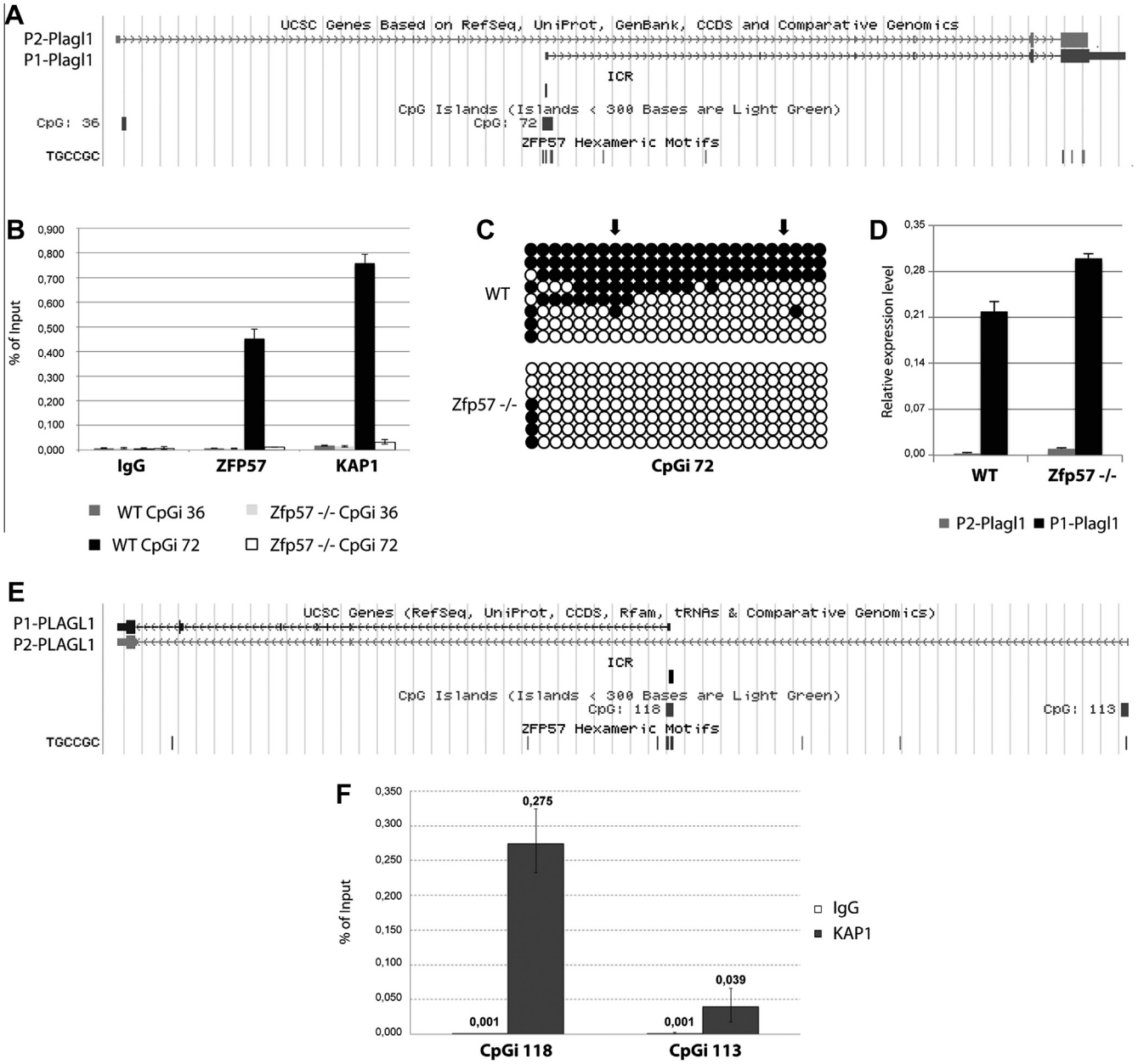


Fig. 5. ChIP, DNA methylation and expression analyses of the *PLAGL1* locus in mouse and human cells. (A) Screen shot of the UCSC genome browser showing the mouse *Plagl1* transcripts derived from the imprinted (P1) and the non-imprinted (P2) promoters. The location of CpGi 36 and CpGi 72 corresponding to a non-methylated CpG island and the *Plagl1* ICR, respectively, and the exameric ZFP57 motifs are indicated. (B) Analysis of ZFP57 and KAP1 binding by ChIP at CpGi 36 and CpGi 72 in mouse wildtype and *Zfp57* −/− ESCs. The amount of target DNA enrichment was estimated by q-PCR and expressed as percent of input. Data represent the average of three independent experiments. (C) DNA methylation analysis of the *Plagl1* ICR (CpGi 72) as assessed by bisulfite sequencing. Each row of circles represents the results of an independent sequencing reaction. Open and filled circles indicate unmethylated and methylated CpGs, respectively. The location of the ZFP57 hexameric target sequences is indicated by arrows. (D) Expression of the *Plagl1* transcripts in mouse wildtype and *Zfp57* −/− ESCs as assessed by quantitative RT-PCR. The RNA was analyzed with primers specific for the P1-*Plagl1* and P2-*Plagl1* transcripts. The results obtained with one clone of *Zfp57* −/− ESCs are shown. The analysis of a second clone gave comparable results (not shown). (E) Screen shot of the UCSC genome browser showing the human *PLAGL1* transcripts derived from the imprinted (P1) and the non-imprinted (P2) promoters. The location of CpGi 113 and CpGi 118 corresponding to a non-methylated CpG island and the TNDM1 ICR, respectively, and the exameric ZFP57 motifs are indicated. (F) Analysis of KAP1 binding by ChIP at CpGi 113 and CpGi 118 in 2102Ep cells. The amount of target DNA enrichment was estimated by q-PCR and expressed as percent of input. Data represent the average of three independent experiments.

entire C-terminus, including the DNA-contacting ZF4 domain. The truncating mutations G299fsX1 and G461GfsX17 and the missense mutation H458D, mapping in the putative seventh ZF domain lacking the fourth coordinating residue, generate proteins including the KRAB domain and the entire DNA-binding ZFs region. Since the clinical features of the patients bearing the ZFP57 mutations in the C-terminal half of the protein are indistinguishable from the others, it is likely that the protein products of these mutant genes are equally not functional. Further studies are required to find out whether these mutations reduce the protein stability or if the C-terminal part of the protein, including ZF5 and ZF6, is directly involved in protein function.

ZFP57 is necessary for the maintenance of DNA methylation at the ICRs and its loss results in activation of some imprinted genes in mouse embryos and ESCs [5,6]. We extended this analysis further and demonstrated that ZFP57 represses *Plagl1* expression

Fig. 6. EMSA DNA binding analysis of hZFP57zf3-4 on the three putative ZFP57 DNA target sequences (PLAGL1-1, PLAGL1-2 and PLAGL1-3) identified in the TNDM1 ICR. The sequences of the oligonucleotides used as probes are indicated.

derived from the imprinted promoter and maintains it methylated in mouse ESCs. A ChIPseq analysis performed by Farnham and co-workers [19] and a locus-specific ChIP performed in this study show the binding of KAP1 to the TNDM1 ICR and in vitro binding assays demonstrate that this sequence is indeed able to bind hZFP57 if methylated.

The causes of isolated loss of *PLAGL1* methylation and multiple hypomethylation of imprinted loci (HIL) without ZFP57 mutations demonstrated in TNDM1 are unknown. They may include environmental factors, reproductive errors or genetic mutations at loci different from *ZFP57*. The elevated incidence of assisted reproduction in patients with TNDM1 and HIL but without ZFP57 mutations, although not statistically significant, may indicate a possible factor in the etiology of at least some of these cases [1]. Whatever the cause of loss of methylation would be, our data suggest that in these patients the lack of methylation on the maternal allele of the 6q24 TNDM1 ICR would reduce the recruitment of the repressor ZFP57/KAP1 complex on this site and alter the expression of the imprinted genes *PLAGL1* and *HYMAI*.

Acknowledgements

This study has been funded by Programma MERIT RBNE08HWLZ_014 (to P.V.P.), MIUR PRIN 2009 and Telethon-Italia grant no. GGP11122 (to A.R.), and Progetto Bandiera MIUR-CNR Epigenomica and People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° 290123 (to A.R. and G.G.).

We thank T. Yokota and T. Akagi for the gift of *Zfp57−/−* cells. We are greatful to Marco Mammucari and Maurizio Muselli for excellent technical assistance.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2013.02.045.

References

[1] Mackay, D.J.G. and Temple, I.K. (2010) Transient neonatal diabetes mellitus type 1. Am. J. Med. Genet. Part C Semin. Med. Genet. 154C, 335–342.
[2] Temple, I.K. and Shield, J.P.H. (2010) 6q24 Transient neonatal diabetes. Rev. Endocr. Metab. Disord. 11, 199–204.
[3] Mackay, D.J.G., Callaway, J.L., Marks, S.M., White, H.E., Acernini, C.L., Boonen, S.E., Dayanikli, P., Firth, H.V., Goodship, J.A., Hahnemann, J.M., Kordonouri, O., Masoud, A.F., Oestergaard, E., Storr, J., Ellard, S., Httersley, A.T., Robinson, D.O. and Temple, I.K. (2008) Hypomethylation of multiple imprinted loci in patients with transient neonatal diabetes is associated with mutations in ZFP57. Nat. Genet. 40, 949–951.
[4] Alonso, M.B., Zoidl, G., Taveggia, C., Bosse, F., Zoidl, C., Rahman, M., Parmantier, E., Dean, C.H., Harris, B.S., Wrabetz, L., Muller, H.W., Jessen, K.R. and Mirsky, R. (2004) Identification and characterization of ZFP-57, a novel zinc finger transcription factor in the mammalian peripheral nervous system. J. Biol. Chem. 279, 25653–25664.
[5] Li, X., Ito, M., Zhou, F., Youngson, N., Zuo, X., Ledeer, P. and Ferguson-Smith, A.C. (2008) A maternal-zygotic effect gene, Zfp57, maintains both maternal and paternal imprints. Dev. Cell 15, 547–557.
[6] Quenneville, S., Verde, G., Corsinotti, A., Kapopoulou, A., Jakobsson, J., Offner, J., Baglivo, I., Pedone, P.V., Grimaldi, G., Riccio, A. and Trono, D. (2011) In embryonic stem cells, ZFP57/KAP1 recognize a methylated hexanucleotide to affect chromatin and DNA methylation of imprinting control regions. Mol. Cell 44, 361–372.
[7] Zuo, X., Sheng, J., Lau, H., McDonald, C.M., Andrade, M., Cullen, D.E., Bell, F.T., Iacovino, M., Kyba, M., Xu, G. and Li, X. (2011) The zinc finger protein ZFP57 requires its cofactor to recruit DNA methyltransferases and maintains the DNA methylation imprint in embryonic stem cells via its transcriptional repression domain. J. Biol. Chem. 287, 2107–2118.
[8] Messerschmidt, D.M., de Vries, W., Ito, M., Solter, D., Fergouson-Smith, A. and Knowles, B.B. (2012) Trim28 is required for epigenetic stability during mouse occyte to embryo transition. Science 335, 1499–1502.
[9] Delidow, B.C., Lynch, J.P., Peluso, J.J. and White, B.A. (1993) Polymerase chain reaction: basic protocols. Methods Mol. Biol. 15, 1–29.
[10] Esposito, S., Baglivo, I., Malgieri, G., Russo, L., Zaccaro, L., D'Andrea, L.D., Mammucari, M., Di Blasio, B., Isernia, C., Fattorusso, R. and Pedone, P.V. (2006) A novel type of zinc finger DNA binding domain in the *Agrobacterium tumefaciens* transcriptional regulator Ros. Biochemistry 45, 10394–10405.
[11] Akagi, T., Usuda, M., Matsuda, T., Ko, M.S., Niwa, H., Asano, M., Koide, H. and Yokota, T. (2005) Identification of Zfp-57 as a downstream molecule of STAT3 and Oct-3/4 in embryonic stem cells. Biochem. Biophys. Res. Commun. 331, 23–30.
[12] Azuma, S. and Toyoda, Y. (1991) Production of a germ-line chimeric mouse derived from newly established embryonic stem cells. Jpn. J. Anim. Reprod. 37, 37–43.
[13] Hirasawa, R. and Feil, R. (2008) A KRAB domain zinc finger protein in imprinting and disease. Dev. Cell 15, 487–488.
[14] Wolfe, S.A., Nekludova, L. and Pabo, C.O. (2000) DNA recognition by Cys2His2 zinc finger proteins. Annu. Rev. Biophys. Biomol. Struct. 29, 183–212.
[15] Renda, M., Baglivo, I., Burgess-Beusse, B., Esposito, S., Fattorusso, R., Felsenfeld, G. and Pedone, P.V. (2007) Critical DNA binding interactions of the insulator protein CTCF: a small number of zinc fingers mediate strong binding, and a single finger–DNA interaction controls binding at imprinted loci. J. Biol. Chem. 282, 33336–33345.
[16] Klug, A. and Schwabe, J.W.R. (1995) Protein motifs 5. Zinc finger. FASEB J. 9, 597–604.

[17] Iglesias-Platas, I., Martin-Trujillo, A., Cirillo, D., Court, F., Guillaumet-Adkins, A., Camprubi, C., Bourc'his, D., Hata, K., Feil, R., Tartaglia, G., Arnaud, P. and Monk, D. (2012) Characterization of novel paternal ncRNAs at the Plagl1 locus, including Hymai, predicted to interact with regulators of active chromatin. PLoS One 7, e38907.
[18] Valleley, E.M., Cordery, S.F. and Bonthron, D.T. (2007) Tissue-specific imprinting of the ZAC/PLAGL1 tumour suppressor gene results from variable utilization of monoallelic and biallelic promoters. Hum. Mol. Genet. 16, 972–981.
[19] Iyengar, S., Ivanov, A.V., Jin, V.X., Rauscher 3rd, F.J. and Farnham, P.J. (2011) Functional analysis of KAP1 genomic recruitment. Mol. Cell. Biol. 31, 1833–1847.
[20] Liu, Y., Toh, H., Sasaki, H., Zhang, X. and Cheng, X. (2012) An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev. 26, 2374–2379.