Disease-associated missense variants in ZBTB18 disrupt DNA binding and impair the development of neurons within the embryonic cerebral cortex

Isabel A. Hemming1,2,3* | Olivier Clément1,2,3,4* | Ivan E. Gladwyn-Ng1,2 | Hayley D. Cullen1,2,3,5 | Han Leng Ng4 | Heng B. See1,2 | Linh Ngo1,2 | Daniela Ulgiati4 | Kevin D.G. Pfleger1,2,6 | Mark Agostino5,7,8 | Julian I-T. Heng5,6,7

1Molecular Medicine Division, QEII Medical Centre, The Harry Perkins Institute of Medical Research, Nedlands, Australia
2Centre for Medical Research, The University of Western Australia, Crawley, Australia
3Faculty of Health and Medical Sciences, Medical School, The University of Western Australia, Crawley, Australia
4School of Biomedical Sciences, Faculty of Health and Medical Sciences, The University of Western Australia, Crawley, Australia
5Curtin Health Innovation Research Institute, Curtin University, Bentley, Australia
6ARC Centre for Personalised Therapeutics Technologies, Melbourne, Australia
7School of Pharmacy and Biomedical Sciences, Curtin University, Bentley, Australia
8Curtin Institute for Computation, Curtin University, Bentley, Australia

Correspondence
Mark Agostino and Julian I-T. Heng, Curtin Health Innovation Research Institute, Curtin University, Bentley, Australia.
Email: Mark.Agostino@curtin.edu.au (M. A.); Julian.Heng@curtin.edu.au (J. I-T. H.)

Funding information
National Health and Medical Research Council, Grant/Award Number: 1011505; Curtin Research Fellowship, Grant/Award Number: CRF130006; Raine Priming Grant from the Raine Medical Research Foundation; National Health and Medical Research Council of Australia; Telethon-Perth Children’s Hospital Research Fund; NHMRC RD Wright Biomedical Research Fellowship, Grant/Award Number: 1085842; National Computational Merit Allocation Scheme, Grant/Award Number: pa6/pawsey0196

Abstract
The activities of DNA-binding transcription factors, such as the multi-zinc-finger protein ZBTB18 (also known as RP58, or ZNF238), are essential to coordinate mammalian neurodevelopment, including the birth and radial migration of newborn neurons within the fetal brain. In humans, the majority of disease-associated missense mutations in ZBTB18 lie within the DNA-binding zinc-finger domain and are associated with brain developmental disorder, yet the molecular mechanisms explaining their role in disease remain unclear. To address this, we developed in silico models of ZBTB18, bound to DNA, and discovered that half of the missense variants map to residues (Asn461, Arg464, Glu486) predicted to be essential to sequence-specific DNA contact, whereas others map to residues (Leu434, Tyr447, Arg495) with limited contributions to DNA binding. We studied pathogenic variants to residues with close (N461S) and limited (R495G) DNA contact and found that each bound DNA promiscuously, displayed altered transcriptional regulatory activity in vitro, and influenced the radial migration of newborn neurons in vivo in different ways. Taken together, our results suggest that altered transcriptional regulation could represent an important pathological mechanism for ZBTB18 missense variants in brain developmental disease.

KEYWORDS
brain disorder, missense variant, neurodevelopment, transcription factor, ZBTB18
1 | INTRODUCTION

The activities of DNA-binding transcription factors are crucial to the formation of the mammalian cerebral cortex, during which, coordinated waves of gene expression orchestrate a step-wise process of neurogenesis, cell migration, and circuit formation within the developing organ (Gupta, Tsai, & Wynshaw-Boris, 2002; Nord, Pattabiraman, Visel, & Rubenstein, 2015). ZBTB18 encodes a transcriptional repressor of the broad complex tramtrack bric-a-brac (BTB) zinc-finger family, comprising an N-terminal BTB domain for protein–protein interaction, as well as four Cys2-His2-like (C2H2) zinc fingers at its C-terminus for DNA binding (Mitchelmore et al., 2002). In mice, Zbtb18 is essential to neurodevelopment (Okado et al., 2009; Xiang et al., 2011), including the radial positioning of newborn postmitotic neurons within the developing cerebral cortex (Heng et al., 2013; Ohtaka-Maruyama et al., 2013). Notably, loss of Zbtb18 results in derepression of downstream target genes, including the cytoskeleton gene Rnd2, which, in turn, leads to defective neuronal migration within the embryonic cerebral cortex (Heng et al., 2013; Ohtaka-Maruyama et al., 2013). We reported that ZBTB18 directly binds enhancer motifs within the Rnd2 gene identified as E1 and E2 (Heng et al., 2013), to modulate its levels for the control of radial migration by immature neurons. The transcriptional repressor functions for ZBTB18 are essential to its role in the regulation of multiple target genes essential to embryo development (Ohtaka-Maruyama et al., 2013; Xiang et al., 2011).

In humans, genetic mutations in ZBTB18 are associated with structural brain abnormalities, neuronal migration disorder and intellectual disability (Cohen et al., 2017; de Munnik et al., 2014; Deciphering Developmental Disorders, 2017; Depienne et al., 2017; Hemming et al., 2016; van der Schoot et al., 2018). More than half of all pathogenic variants identified in ZBTB18 are predicted to be truncating, suggesting that haploinsufficiency/loss-of-function represents a general pathological mechanism for disease (Depienne et al., 2017). However, of the remaining variants that are not predicted to be truncating, it is noteworthy that the majority of these map to the C-terminal zinc-finger DNA-binding domain of the polypeptide (Figure S1 and Table S1). This suggests that missense variants could influence the functions for ZBTB18 in DNA binding and transcriptional regulation which, in turn, leads to brain developmental disease.

Here, we have used homology modeling, molecular dynamics (MD) simulations, and molecular mechanics-generalized Born/GB/SA) calculations to investigate the structural and functional properties of wild-type ZBTB18 and two disease-associated variant—NP_991331.1:p.Arg495Gly (R495G; Rauch et al., 2012) and NP_991331.1:p.Asn461Ser (N461S; Farwell et al., 2015; rs797044885) bound to cognate DNA motifs. Through this approach, we propose key residues within ZBTB18 mediating its contact with DNA, and demonstrate that disease-associated missense variants to contact and nonbase-contact residues within ZBTB18 influence its sequence-specific DNA binding, as well as its transcriptional activity in vitro and neuronal functions in vivo.

2 | MATERIAL AND METHODS

2.1 | Homology modeling of ZBTB18 zinc-finger domain in complex with DNA probe sequences

The structures of the wild-type ZBTB18 zinc-finger region from residues 373-501 in complex with DNA fragments from the E1 and E2 probe sequences were prepared by homology modeling. These motifs were derived from a previously characterized 3’ regulatory enhancer sequence for Rnd2, a downstream target gene which mediates neuronal migration during brain development (Heng et al., 2013). Separate templates for each of the four C2H2 subdomains comprising the ZBTB18 zinc-finger domain were identified using BLAST searches of each subdomain against the Protein Data Bank (PDB) using Prime Structure Prediction (Table S2). The relevant C2H2 subdomains of the templates were overlaid to the relevant C2H2 subdomains of the designed zinc-finger protein Aart (PDB 2I13; Segal, Crotty, Bhakta, Barbas, & Horton, 2006), which afforded the highest sequence identity to the overall ZBTB18 zinc-finger domain. The Composite/Chimera feature of prime was used to build protein–DNA complexes for the wild-type ZBTB18 zinc-finger domain, initially retaining the DNA from PDB 2I13. This DNA was trimmed to a 15 base pair fragment, representing the length of DNA most closely bound by the zinc-finger domain.

C2H2 zinc-finger domain structures in complex with DNA similar in sequence to the E1 sequence were then identified; these structures guided the placement of the desired DNA sequences with respect to the protein. The PDB was searched for C2H2 zinc-finger domain structures (PFAM PF00096) with DNA bound. Pairwise alignment of the E1 sequence to each DNA sequence from the identified structures was performed using the needle tool of EMBOSS (Rice, Longden, & Blanchon, 2000), this identified the DNA contained in PDB 5KE6 as affording the best alignment (Hashimoto et al., 2017). PDB 5KE6 was then aligned to the initially built DNA-protein complex (i.e., the complex of ZBTB18 with the DNA from PDB 2I13). A sequence alignment between the DNA in PDB 5KE6 and the DNA in PDB 2I13 was generated based on the structural alignment between the two DNA fragments. Combining this structure-based sequence alignment with the sequence alignment of the DNA probe sequences with the template DNA allowed an alignment between the DNA probe sequences and the DNA from PDB 2I13 to be generated (Table S3). The DNA from PDB 2I13 was then mutated in UCSF Chimera to the relevant sequences based on this alignment.

After model building and inclusion of the appropriate DNA sequences, the complex was energy minimized using Prime Minimization in a step-wise manner; non-template residues were first minimized, followed by all protein and ion components of the complex (i.e., the entire structure, excluding DNA), and finally, the entire complex. Erroneous cis-amides (i.e., cis-amides not
derived from the template or immediately before a proline residue) and incorrect stereochemistry generated during the model building process were identified and corrected using Maestro. The structures of the zinc-finger domain with each of the E1 and E2 probe sequences were produced in this manner; complexes with mutant proteins were prepared by mutating the relevant residue(s) in these representative structures using the mutation facilities of Maestro and used without further minimization.

2.1.1 Molecular dynamics simulations

Parameterization of the complexes was performed using AmberTools (Salomon-Ferrer, Case, & Walker, 2013). Simulations were performed using GROMACS 5.0.2 (Abraham et al., 2015). ZBTB18 was parameterized with the AMBER ff14SB force field (Maier et al., 2015), whereas bound DNA was parameterized with the parmbsc0 parameter set (Perez et al., 2007) with ε/OL1 (Zgarbova et al., 2013) and χ/OL4 (Krepl et al., 2012) modifications. The Zinc AMBER force field (Peters et al., 2010) was used to parameterize the zinc centers in ZBTB18. The resulting topology was then ported to GROMACS format using acpype (Sousa da Silva & Vranken, 2012) and the remaining system setup completed in GROMACS.

Complexes were solvated in TIP3P water (Jorgensen, Chandrasekhar, Madura, Impey, & Klein, 1983) in a dodecahedral box with a minimum of 10 Å distance from the protein to the box edge. The system was charge neutralized with the addition of sodium ions, with further sodium and chloride ions added to a concentration of 0.1 M. Equilibrations of the system in the NVT and NPT ensembles were adapted from previously described procedures (Shields, Laughton, & Orozco, 1997; Shields, Laughton, & Orozco, 1998; Perez et al., 2007); briefly, short simulations (0.1 ns) with gradually reducing position restraints on the complex heavy atoms were performed (initially an NVT simulation with 10,000 kJ/(mol nm²) applied, followed by NPT simulations with 2,000 kJ/(mol nm²) applied, gradually decreasing by 400 kJ/(mol nm²)). After this, the production MD simulation was performed in the NPT ensemble for 50 ns, with a target temperature of 300 K and a target pressure of 1 atm. The modified Berendsen thermostat (V-rescale) was employed for temperature coupling (Bussi, Donadio, & Parrinello, 2007). The Parrinello–Rahman barostat was employed for pressure coupling (Parrinello & Rahman, 1981). The smooth particle mesh Ewald method was employed for long-range electrostatics (Essmann et al., 1995). All bonds were constrained using the LINCS algorithm (Hess, Bekker, Berendsen, & Fraaije, 1997). Coordinates were saved every 10 ps.

Five simulations of each complex were performed, starting from new random velocities in each replicate. For binding energy calculations, the final 10 ns of three of these simulations were used, determined by examining the root-mean-squared deviation (RMSD) for heavy atoms over this time range; simulations, where the RMSD appeared stable and maintaining under 6.0 Å, were selected for binding energy calculations.

2.1.2 Binding free energy calculations and decomposition

Binding energies were calculated using the (MM-GB/SA) approach, facilitated by the MMPBSA.py tool of AmberTools (Miller et al., 2012). The single-trajectory protocol was utilized (Gohlke, Kiel, & Case, 2003), with the following equation:

$$\Delta G_{\text{bind}} = G_{\text{complex}} - G_{\text{protein}} - G_{\text{ligand}} = \Delta H + \Delta G_{\text{solvation}} - T\Delta S$$

where $$\Delta G_{\text{vdw}}$$ is the molecular mechanics the van der Waals interaction energy, $$\Delta G_{\text{ele}}$$ is the molecular mechanics electrostatic interaction energy, $$\Delta G_{\text{GB}}$$ is the change in polar desolvation energy upon complex formation and $$\Delta G_{\text{SA}}$$ is the change in nonpolar desolvation energy upon complex formation. The entropic term ($-T\Delta S$) was not calculated, due to high computational cost and poor accuracy. The modified GB model developed by Onufriev, Bashford, & Case (2004) (gb = 5) was used to calculate the polar desolvation energy; this GB model is suggested as a suitable “general purpose” GB model and has been previously applied to protein–DNA complexes (Rusco & Onufriev, 2006). The nonpolar component of the desolvation energy was calculated via solvent accessible surface areas calculated with the LCPO method (Weiser, Shenkin, & Still, 1999). MM-GB/SA energies were also decomposed per-residue (idecomp = 1). The GROMACS tool g_cluster was used to identify the most frequently observed structures from the regions of the simulations submitted for MM-GB/SA analysis. The cutoff for the g_cluster calculation was set to 0.25 Å.

2.2 Mouse handling and animal ethics license

All animal procedures were conducted according to standard operating procedures approved by the Animal Ethics Committees (Perkins Institute: AE021, University of Western Australia: RA/3/100/1361), as well as guidelines provided by the National Health and Medical Research Council of Australia. Unsexed C57BL/6 J wild-type mouse embryos were used for in utero electroporation studies.

2.3 DNA constructs and cloning

A Zbtb18 shRNA (short hairpin RNA) vector used in this study has been previously reported (Heng et al., 2013). The cDNA sequence encoding isoform 1 of human ZBTB18 (NM_205768.2) and ZBTB18 (R495G) were commercially synthesized (Integrated DNA Technologies, Coralville, IA). Both cDNAs were individually subcloned into pCIG-FLAG vector in which the GFP cassette was excised by restriction digestion and subsequent cloning (pCIG-F[NG]; Heng et al., 2013). The N461S variant was generated via site-directed mutagenesis using a GeneArt® Site-Directed
Mutagenesis PLUS kit (A14604; Life Technologies, Carlsbad, CA) on the pCIG-F[g] ZBTB18(WT) construct with the following primers: sense, 5′-TCCAGTACTCGCAAGCTAGCCCAGCTGC-3′; antisense, 5′-GC ATGGCCGCTCAGGCTGTGGGTAATGGA-3′. All constructs were verified by Sanger sequencing (Australian Genome Research Facility, Perth, Australia).

2.4 | Cell culture and transfection

Human (HeLa, HEK293T) and mouse (P19 embryocarcinoma) cell lines were cultured in complete growth media (DMEM [10313-021; Gibco by Life Technologies, Carlsbad, CA], 2 mM l-glutamine [25030-081; Gibco by Life Technologies, Carlsbad, CA], 10% fetal bovine serum [SFBSF; Bovogen Biological, Keilor East, Victoria, Australia]), in an incubator at 37°C supplied with humidified 5% CO2. HeLa cells were seeded in 500 μl of growth media in a 24-well plate, at a density of 50,000 cells per well for localization experiments. P19 cells were seeded in 500 μl of growth media in a 24-well plate, at a density of 25,000 cells per well for luciferase assays. HEK293T cells were seeded in 2 ml of growth media in a 6-well plate, at a density of 500,000 cells per well for electrophoretic mobility shift assay (EMSA) experiments and initial western blotting, and in a 48-well plate at a density of 25,000 cells per well for experiments with MG132 exposure. After seeding of 24 hr, cells were transiently transfected with Lipofectamine 2000 (11668-019; Invitrogen Life Technologies, Carlsbad, CA) as per manufacturer’s instructions. The media was changed 5 hr posttransfection (media without Lipofectamine). For immunostaining experiments, HeLa cells were cultured on uncoated 13 mm round coverslips (CS13100; Grale by Trajan, Ringwood, Victoria, Australia) and fixed onto the coverslips at either 24 or 48 hr posttransfection. P19 cells were collected 48 hr posttransfection for luciferase assays. Nuclear extract from HEK293T cells was collected 48 hr posttransfection for EMSAs. For MG132 experiments HEK293T cells were treated with 10 μM MG132 (C2211; Sigma-Aldrich, St. Louis, MO) at 40 hr posttransfection, and then washed with 1x PBS buffer and lysed using KALB buffer at either 0, 4, or 8 hr after MG132 treatment.

2.5 | Luciferase assays

Firefly luciferase reporter constructs harboring a synthetic decamerised binding site (BS10), as well as a construct comprising mouse Rnd2 3′-enhancer sequence (Heng et al., 2008), were used. A Renilla luciferase construct was cotransfected to normalize for transfection efficacy (Promega, Madison, WI). For competition assays between WT, N461S, and R495G, 250 ng of each ZBTB18 construct was added, together with 50 ng of firefly reporter construct, and 25 ng of Renilla construct, per well. DNA was added to 100 μl Opti-MEM I (31985-070; Gibco by Life Technologies, CA) per sample. The four constructs were pCIG-F[g]-empty, pCIG-F[g]-ZBTB18(WT), pCIG-F[g]-ZBTB18(N461S), and pCIG-F[g]-ZBTB18(R495G). A total of 48 hr posttransfection, P19 cells were washed with 500 μl 1x PBS and lysed with 250 μl 1x Passive lysis buffer as per the manufacturer’s instructions (E1960; Dual Luciferase Assay Kit, Promega, Madison, WI). The lysate was then centrifuged at 15000g for 15 min at 4°C. 20 μl of supernatant for each 24-well was added to two 96-wells, and reagent added to each well manually. Firefly and Renilla signals were read in each well of the 96-well plate in order of rows (5 s recordings per well), using a VictorLight plate reader (Perkin Elmer, Waltham, MA) according to the manufacturer’s instructions.

2.6 | Western blot analysis

HEK293T cells were collected 48 hr after transfection and western blotting. Briefly, cells were washed with 1x PBS buffer and were lysed using KALB buffer (1% Triton X detergent, 150 mM NaCl, 0.02% sodium azide, 1 mM ethylenediaminetetraacetic acid (EDTA), 50 mM Tris-HCl pH 7.4, and protease inhibitor (4693159001; Roche, Basel, Switzerland)), and protein concentration measured using Bradford reagent (500205; Bio-Rad, Hercules, CA). Protein samples were analyzed on a 10% sodium dodecyl sulfate (SDS) gel, transferred to nitrocellulose membrane, and incubated overnight with primary antibodies: mouse anti-β-actin (A5441; Sigma-Aldrich; 1:5,000), mouse anti-FLAG (Sigma-Aldrich; F1804, 1:5,000), mouse anti-ZBTB18 (H00010472-M04; Abnova, Jhongli, Taiwan; 1:5,000), rabbit anti-FLAG (2368; Cell Signaling, Danvers, MA; 1:5,000) and rabbit anti-ZBTB18 (12714-1-AP; Proteintech, Chicago, IL; 1:5,000). Membranes were incubated with antirabbit IRDye 800LT or antimouse IRDye 800 secondary antibodies (LI-COR Biosciences, Lincoln, NE; 1:10,000 for each) for 2 hr before analysis using the Odyssey imaging system (LI-COR Biosciences).

2.7 | Electrophoretic mobility shift assays

EMSA were performed as previously described (Cruickshank et al., 2015), using nuclear extracts from HEK293T transfected cells collected using the NE-PER™ Nuclear and Cytoplasmic Extraction Kit according to manufacturer’s instructions (78833; Life Technologies, CA). In short, 2 μg of crude nuclear protein extract was incubated with 75 fmol of double-stranded oligonucleotides in EMSA reaction buffer (4% Ficoll, 20 mM HEPES [pH 7.9], 1 mM EDTA [pH8.0], 15 mM DTT, 0.5 μg Poly dIdC) for 30 min, before running on a 6% nondenaturing polyacrylamide gel (6% Bis-acrylamide, 2.5% glycerol, 0.75× TBE buffer) for 90 min at 150 V. Commercially sourced, HPLC-purified 5′-biotinylated oligonucleotides (Sigma-Aldrich) comprising sense and antisense strands of DNA stretches encompassing the binding sites of E1 (sense 5′-CCGTGGGTCTCTCCTCTTAAATAC AGATGTCGTCTGCTGCA-TCA-3′; antisense 5′-TATGGCAGACAGCTACCTTTTTCGAGATACGAA-C-5′); E1mut (sense 5′-CCGTGGGTCTCTCCTCTTAAATAC AGATGTCGTCTGCTGCA-TCA-3′; antisense 5′-TAT GGCAGACAGCTACCTTTTTCGAGATACGAA-C-3′); E2 (sense 5′-TTCATCACAAGGAGGAGCAGATGGGAGTAGG
GAAG-3’; antisense 5’-CTTCCCTACTCCCATCTGCCCTCCCTTT GTGTTGATCAA-3’; E2mut (sense 5’-TTTGGATCCCAAAGGGAG GGCCAGTGAGGATGGGAAG-3’; antisense 5’-CTTCCCTACT CCTACCTGCCCCCTCTTGGTGGATCAA-3’). Strand pairs were annealed by resuspension in 1× TE buffer and heated to 95°C for 10 min, then slowly cooled to room temperature. Oligonucleotides were run on a 6% polyacrylamide gel and double-stranded oligonucleotides were extracted from the gel in elution buffer (0.1% SDS, 0.5 M ammonium acetate, 10 mM magnesium acetate) at room temperature for 16 hr. Oligonucleotides were precipitated with 0.5 M ammonium acetate, 10 mM magnesium acetate and air dried thoroughly. Oligonucleotides were resuspended in 100% ethanol at 4°C for 16 hr, then pelleted by centrifugation at 16,000 g for 30 min at 4°C and washed with 800 μl of 70% ethanol and air dried thoroughly. Oligonucleotides were resuspended in nuclease-free water and diluted to 25 pmol/μl stock just before use. Biotin detection was undertaken with a Chemiluminescent Nucleic Acid Detection Module Kit according to the manufacturer’s instructions (89880; Thermofisher, Waltham, MA).

2.8 | In utero electroporation and tissue collection for immunostaining

In utero electroporation was performed on E14.5 embryos as described previously (Haas et al., 2016). Plasmids were delivered at a final concentration of each species at 1 μg/μl, premixed with Fast Green to visualize the site of injection. After 3 days of electroporation, the pregnant dam was euthanatized by cervical dislocation, embryos (E17.5) were decapitated, and their brains were dissected then fixed overnight at 4°C in 4% paraformaldehyde in PBS. Then, the brains were equilibrated in 30% sucrose in PBS for 3 days at 4°C, before embedding in OCT mounting medium, frozen on dry ice and finally stored at −80°C. Coronal sections 16-μm thick were prepared from each brain using a cryostat (Leica, Biosystems, Wetzlar, Germany), and recovered on Superfrost-PLUS glass slides (Menzel-Glas by Trajan, Ringwood, Victoria, Australia). Slides were dried for at least 60 min at ambient temperature conditions before storage at −80°C until immunostaining was performed, as follows. Briefly, after a blocking step with 10% of normal goat serum in PBS, brain sections were incubated overnight at 4°C with the primary antibodies: chicken anti-ZBTB18 (Sigma-Aldrich, F-1804, 1:500). Then sections were incubated with fluorescein secondary antibodies (488-Goat anti-Chicken (Invitrogen, Carlsbad, CA; A11039, 1:1000); 568-Goat antimouse (Invitrogen, A11031, 1:1000)) at room temperature for 2 hr before sections were counterstained with 4’-Diamidino-2-Phenylindole (DAPI; Invitrogen, D1306, 1:1000). A coverslip was mounted on each slide using Fluorescent Mounting Medium (DAKO, Berlin, Germany).

2.9 | Microscopy, imaging, quantification studies

Images of brain sections were captured on an epifluorescence microscope (Olympus, Shinjuku, Tokyo, Japan) equipped with a CCD camera (SPOT, Sterling Heights, MI) for neuronal migration, and on a confocal microscope (Nikon, Minato, Tokyo, Japan) for the analysis of neuronal shape. Subdivisions of the embryonic cortex (VZ/SVZ, IZ, and CP) were identified based on cell density as visualized with 4’-6-diamidino-2-phenylindole staining, as described previously (Haas et al., 2016). Cell counting was performed blind to the condition of each split sections of electroporated brains using ImageJ software.

2.10 | Statistics

Data from all experiments are expressed as the mean ± standard error of the mean. A Kruskal-Wallis analysis was employed to test multiple conditions, whereas a Mann-Whitney U test was applied for statistical analyses between two treatments. All statistics were performed through Prism software (GraphPad, San Diego, CA).

3 | RESULTS

3.1 | Structural prediction and binding energy calculations of ZBTB18 zinc-finger domain for wild-type and missense variants, in complex with DNA probe sequences

We used homology modeling, MD simulations, and MM-GB/SA calculations to investigate the structural and functional properties of the C-terminal zinc-finger domain of ZBTB18 (residues 373-501) bound to two DNA regulatory motifs, E1 and E2, within the Rnd2 gene (Heng et al., 2008; Heng et al., 2013). To explore sequence-specific binding properties, we also modeled ZBTB18 binding to mutated versions of each motif, named E1mut and E2mut, which we also previously reported as refractory to ZBTB18-mediated transcriptional repression (Heng et al., 2013). The predicted structure of the ZBTB18 zinc-finger domain comprises four C2H2-type subdomains, featuring a moderately sized insertion between the first two C2H2 subdomains relative to the template structures (Figure 1, Video S1 and Figures S2A-C). MD simulations indicated the stability of the majority of protein–DNA complexes, which were typically reproducible when simulations were commenced from new random velocities (Figure S3A,D,G,J).

Based on our MM-GB/SA calculations, we found that wild-type ZBTB18 formed energetically more favorable interactions with both E1 and E2, compared to the mutated sequences E1mut and E2mut (Table 1). Per-residue binding energy decomposition identified residues—including Tyr458, Asn461, and His493—as critical to wild-type ZBTB18 to bind E1 and E2 (Figure 2, Table 2 and Table S4). For example, Asn461 contributes as much as 7.2% and 7.6% of the total binding energy of ZBTB18 to E1 and E2, respectively. Importantly, binding energy contributions made by these residues were markedly lower for ZBTB18 modeled with E1mut and E2mut mutants, with the exception of His493, which displayed an increased contribution when complexed with E2mut, altogether suggesting roles for each residue in sequence-specific binding (Figure 2b). Based on our models of ZBTB18 with E1 and E2, we also observed that half of the disease-associated missense variants map to residues (Asn461, Arg464, Glu486) essential to sequence-specific DNA.
contact, whereas others map to residues exhibiting comparably reduced (Tyr447, Arg495) or negligible (Leu434) contributions to DNA-binding (Table 2).

Next, we focused our attention on disease-associated ZBTB18 variants which featured residues with close (N461S) or limited (R495G) DNA contact. Separate models were generated for N461S (Figure 3a,b, and Figure S2D–F) and R495G (Figure 3c,d, and Figure S2G–I) binding E1, E1mut, E2, and E2mut. Binding energy calculations suggested that N461S exhibited more selective binding for E1 over E1mut, and preferential binding for E2mut over E2 (Table 1). On the other hand, R495G exhibited similar binding energies for each of the E1, E1mut, and E2 motifs, suggesting nonselective binding for all three sequences. Calculations with E2mut indicated that R495G bound with substantially lower energy compared to E1, E1mut, and E2 motifs.

We performed binding energy decomposition to understand the interactions taking place in the complexes in finer detail. In the case of a N461S substitution, we observed a reduced contribution to the binding energy of this position between E1 and E2, concomitant with a structural rearrangement of Tyr458 and His493, which augmented their binding with E1mut and E2mut, as well as enhanced the contribution to the binding energy of His493 with E2 (Figure 3b; Table 2). In contrast, Arg495 weakly contributes to binding of E1 and E2, and substitution to glycine eliminates any contribution to binding energy made by this position (Figure 3d, Table 2). However, with this substitution, we also observed reduced contributions made by Tyr458 and Asn461 when binding to the E1 and E2 motifs, as well as an increase in contribution to E2 binding by His493, altogether suggesting indirect effects of this missense variant on DNA binding.

When modeled with E1mut, Tyr458 also shows a five-fold elevation in binding energy contribution in the R495G variant (−3.5 kcal/mol for this residue within R495G complexed with E1mut, vs. −0.7 kcal/mol within WT ZBTB18 complexed with E1mut; see Table 2). Therefore, our models suggest that missense variants to residues relevant to close (N461S) or limited (R495G) DNA contact differentially affect the capacity for ZBTB18 to bind DNA, and could lead to spurious binding.

3.1.1 Missense variants in ZBTB18 disrupt sequence-specific DNA-binding and transcriptional repression

To substantiate our molecular predictions, we performed EMSAs with DNA probes comprising each motif. We found that WT ZBTB18 formed distinct molecular complexes with a DNA probe comprising E1 (Figure 4a, arrows), with an additional complex

![Image 1](image1.png)

**FIGURE 1** Homology model for ZBTB18 in complex with a 15-residue duplex derived from E1 sequence (TAACAGATGTCTGTC), shown in “front” (C2H2 domains 1 and 4 facing viewer) and “rear” (C2H2 domains 2 and 3 facing viewer) views. Zinc Fingers 1-4 (colored as N- to C-terminal rainbow, blue to red), as well as zinc atoms (gray spheres) and DNA residues (dA: white; dC: pink; dG: pale green; dT: tan) are shown. Residues of ZBTB18 impacted by disease-associated missense variants (Leu434, Tyr447, Asn461, Asn464, Glu486, and Arg495) are shown in the lateral panels, with hydrogen bonds in dashed green lines. Arg464, Gln486, and Asn461 are rotated slightly around the x-axis relative to the full complex views. See Figure S1B and Table S1 for the degree of amino acid sequence conservation and binding energy decomposition for these six residues, respectively. C2H2, Cys2-His2.

| ZBTB18 variant | ΔΔGbind (kcal/mol) |
|----------------|-------------------|
| DNA motifs     | E1    | E1mut | E2     | E2mut | E1 | E1mut | E2      | E2mut |
| WT             | −119.8 ± 0.3 | −117.1 ± 0.3 | −125.2 ± 0.3 | −94.0 ± 0.3 | −119.8 ± 0.3 | −117.1 ± 0.3 | −125.2 ± 0.3 | −94.0 ± 0.3 |
| N461S          | −129.6 ± 0.3 | −119.2 ± 0.4 | −108.1 ± 0.3 | −124.5 ± 0.4 | −129.6 ± 0.3 | −119.2 ± 0.4 | −108.1 ± 0.3 | −124.5 ± 0.4 |
| R495G          | −126.4 ± 0.3 | −126.0 ± 0.3 | −126.9 ± 0.4 | −102.0 ± 0.3 | −126.4 ± 0.3 | −126.0 ± 0.3 | −126.9 ± 0.4 | −102.0 ± 0.3 |
formed by R495G (Figure 4a, arrowhead). In agreement with our modeling results, these complexes were either detected at very low intensity or not detected at all in assays with WT ZBTB18 and E1mut (Figure 4b). In contrast to WT ZBTB18, however, each missense variant formed distinct complexes with the E1mut motif (Figure 4b, arrow and closed arrowhead), with R495G forming a prominent high-molecular-weight complex (Figure 4b, open arrowhead). In the case of the E2 motif, ZBTB18 and its missense variants formed molecular complexes (Figure 4c, arrows), with additional complexes formed by R495G (Figure 4c, open and closed arrowheads). However, such protein–DNA complexes were not identified in parallel experiments using a probe comprising the E2mut motif (Figure 4d). Together, these modeling and in vitro data indicate that ZBTB18 binds its DNA motifs E1 and E2 in a sequence-specific manner, whereas its missense variants bind E1, E2, as well as E1mut, but not E2mut.

To understand the impact of each missense variant ZBTB18 on transcriptional regulation in vitro, we performed luciferase reporter assays using a previously validated synthetic promoter construct (termed BS10; Aoki et al., 1998), as well as a construct harboring the Rnd2 3′ enhancer (comprising E1 and E2 motifs; Heng et al., 2013). We found that wild-type ZBTB18 suppressed luciferase reporter

### TABLE 2 Binding energy decompositions for selected residues in ZBTB18-DNA complexes

| Protein DNA | ΔGbind (kcal/mol) | N461S | R495G |
|-------------|-------------------|-------|-------|
| Leu434      | ΔE1, ΔE1mut, ΔE2, ΔE2mut |       |       |
| Tyr447      | -2.1              | -0.5  | +0.0  |
| Tyr458      | -3.2              | -0.7  | -3.4  |
| Asn/Ser461  | -8.6              | -2.3  | -9.6  |
| Arg464      | -4.0              | -3.1  | -4.9  |
| Glu486      | -3.1              | -2.7  | -4.3  |
| His493      | -5.9              | -0.3  | -0.7  |
| Arg/Gly495  | -0.7              | -2.5  | -0.3  |

Note: Three residues (Tyr458, Asn461, and His493) with significant contributions to DNA-binding are indicated in gray. Binding energy calculations for additional residues (Leu434, Tyr447, Arg464, Glu486, Arg495) impacted by pathogenic missense variants are also listed. A comprehensive list describing binding energy decomposition values for all residues making a major contribution to the binding energy in at least one complex is provided as Table S4.
activity, whereas N461S failed to suppress reporter activity mediated by the BS10 synthetic enhancer, but stimulated luciferase signal mediated by the Rnd2 3′ enhancer (Figure 4e,f). Unexpectedly, we found that R495G significantly enhanced luciferase signal in both reporter constructs. The capacity for N461S and R495G to stimulate reporter activity was extinguished when equal quantities of the WT construct were cotransfected (Figure S4). Next, we explored transcriptional reporter activity of ZBTB18 and each missense variant using reporter constructs bearing either $E_1^{\text{mut}}$, $E_2^{\text{mut}}$, or both $E_1^{\text{mut}}$ and $E_2^{\text{mut}}$ motifs within the Rnd2 3′ enhancer sequence. For wild-type ZBTB18, transcriptional repression was only abolished in experiments with a reporter bearing both $E_1^{\text{mut}}$ and $E_2^{\text{mut}}$ motifs, confirming its sequence-specific binding to native $E_1$ and $E_2$ motifs (Figure 4g). In contrast, statistically significant transcriptional activation by N461S was observed in constructs bearing $E_1+E_2$ or $E_1^{\text{mut}}+E_2^{\text{mut}}$ motifs, but not an $E_1^{\text{mut}}+E_2$ motif or

**FIGURE 3** Structural analysis of DNA complexes with ZBTB18 missense variants. (a) Model of ZBTB18 N461S missense variant zinc fingers 1-4 bound to the E1 sequence. (b) Major interactions of Tyr458, Ser461, and His493 of the N461S variant with $E_1$, $E_1^{\text{mut}}$, $E_2$, and $E_2^{\text{mut}}$ sequences. Hydrogen bonds (dashed green lines), coordination bonds (dashed purple lines), and residues involved in hydrophobic contacts in at least one panel (transparent spheres) highlighted. Views rotated slightly relative to panel (a) to improve the visibility of contacts. In all panels, the protein is colored rainbow (blue to red) from N- to C-terminal, with zinc atoms as gray spheres, and DNA residues colored as follows: dA, white; dC, pink; dG, pale green; dT, tan. (c) Model of ZBTB18 R495G missense variant zinc fingers 1-4 bound to the E1 sequence. (d) Major interactions of Tyr458, Asn461, and His493 of N461S variant with $E_1$, $E_1^{\text{mut}}$, $E_2$, and $E_2^{\text{mut}}$ sequences. Hydrogen bonds (dashed green lines), coordination bonds (dashed purple lines), and residues involved in hydrophobic contacts in at least one panel (transparent spheres) highlighted. Views rotated slightly relative to panel (c) to improve visibility of contacts.
an E1+E2mut motif. When compared to WT ZBTB18, the N461S variant has reduced the capacity for repression in the context of E1+E2, E1mut+E2 and E1+E2mut motif pairs, and activation in the context of the E1mut+E2mut motif pair (Figure 4h, summarized in Figure S5B). These data suggest that N461S variant has reduced repressor activity compared with WT, and may influence transcriptional activation in some contexts. Conversely, R495G robustly stimulated reporter activity in all four reporter constructs, suggesting this variant has a strong capacity for activation in all contexts (Figure 4i, Figure S5C).

Next, we performed transfection experiments utilizing FLAG-tagged ZBTB18 constructs to explore the impact of each missense variant on subcellular localization and expression. After 24-hr transfection, we observed reduced nuclear localization of N461S compared to wild-type ZBTB18, concomitant with a corresponding increase in pancellular signal distribution (Figure S6A–C). The
proportion of cells with a punctate distribution of R495G within the nucleus was also significantly increased compared with wild-type ZBTB18 (Figure S6D). These effects were transient and they were not observed 48-hr posttransfection (Figure S6E–G). We also routinely observed significantly fewer FLAG-N461S immunopositive cells compared with wild-type and R495G-transfected cells suggesting that N461S might be unstable (Figure S7A,B). Furthermore, low steady-state levels of FLAG-N461S protein from lysates of transiently transfected cells were observed (Figure S7C,D). To further investigate the hypothesis that the N461S variant might be unstable, we examined FLAG-tagged ZBTB18 expression in transiently transfected cells treated with the proteasome inhibitor MG132 for up to 8 hr followed by western blotting analysis (Figure S7F). In this experiment, we reproducibly observed a very low immunoblotted FLAG signal in the lysate of N461S transfected cells compared to WT and R495G conditions at 0 hr of MG132 treatment (Figure S7E). However, after 4 hr of MG132 treatment, the FLAG signal was significantly increased for N461S, compared with no MG132 treatment. After 8-hr MG132 exposure, N461S, WT, and R495G variant levels were significantly increased (Figure S7F). These results indicate that the N461S variant exhibits an exaggerated protein instability compared with WT, and is prone to degradation in cells.

### 3.1.2 Disease-associated missense variants to ZBTB18 disrupt radial migration within the embryonic cerebral cortex

The transcriptional regulatory activities of ZBTB18 are essential to mammalian neurodevelopment, and suppression of Zbtb18 impairs radial migration of cortical projection neurons from their birthplace within the germinal ventricular zone (VZ), including their multipolar-to-bipolar transition within the intermediate zone (IZ) as they reach the cortical plate (CP) of the embryonic cerebral cortex (Heng et al., 2013; Ohtaka-Maruyama et al., 2013). Hence, we assessed the capacity for ZBTB18 and its missense variants to influence the radial migration of E14.5-born cortical neurons within the E17.5 cortex through a series of in vitro electroporation experiments. Treatment with Zbtb18 shRNA disrupted the capacity for cells to migrate into the CP; however, this impairment was corrected by codelivery of wild-type ZBTB18 (Figure 5a–g), consistent with previous findings (Heng et al., 2013; Ohtaka-Maruyama et al., 2013). Strikingly, we found that codelivery of N461S significantly augmented the radial migration of Zbtb18 shRNA-treated cells into the CP and markedly potentiated their intracortical positioning within the upper CP (Figure 5f,g). We could not detect N461S immunofluorescence in rescued cells (Figure S8), consistent with the notion that this protein variant is unstable (see Figure S7), and beyond the limit of immunodetection in electroporated cortical cells. In contrast, codelivery of R495G exacerbated their migration defect (Figure 5f,g).

Next, we analyzed the cell shapes of IZ and CP neurons across each treatment group and found that knockdown of Zbtb18 led to a significant reduction in the proportion of uni/bipolar-shaped cells in the CP and a concomitant increase in the proportion of multipolar-shaped cells, as well as significantly shortened, leading neurites (Figure 5h–o). Codelivery of WT ZBTB18 resulted in significant restorative effects on cell shape and leading neurite length (Figure 5m–o). In contrast, codelivery of N461S did not affect IZ cell shape but instead led to an increase in multipolar-shaped cells within the CP with shorter leading neurites (Figure 5o), suggestive of dysregulated locomotion which results in their enhanced localization within the upper CP. On the other hand, codelivery of R495G led to an increase in multipolar-shaped cells in the IZ, which we interpret as a disruption of multipolar-to-bipolar transition within the cortical subcompartment. Cell shape profiles and leading neurite lengths of the few
cells arriving within the CP were significantly different to wild-type ZBTB18 rescue, consistent with a detrimental impact of the R495G expression in CP cells. Taken together, missense variants of ZBTB18 impair radial migration in a cell autonomous fashion, and disrupt neurite outgrowth in immature cortical neurons.

4 | DISCUSSION

In this study, we have applied molecular modeling approaches to investigate the binding of ZBTB18 to two native DNA motifs (E1 and E2) within the 3′-enhancer region for the migration-promoting gene, Rnd2; as well as their mutagenised counterparts (E1mut and E2mut). We demonstrate that MM-GB/SA analysis is effective to model ZBTB18 wild-type and variant complexes with different DNA sequences. This is remarkable considering the difficulties associated with the development of accurate DNA force fields for simulations, and the assumptions required for use of the MM-GB/SA approach (Genheden & Ryde, 2015). A strength to our approach has been to incorporate information from multiple templates, as well as using a sequence homology-based approach to identifying relevant DNA structural templates, during the initial homology...
modeling. As a result, we have developed stable DNA-protein complexes and calculated binding energies for ZBTB18 and its variants, with observations substantiated by biochemical observations in this study.

A recent study by van der Schoot et al. (2018) identified via exome sequencing ZBTB18 variants at Arg464 in patients with neurological abnormalities and used homology modeling against a single template structure and a model DNA fragment to illustrate its potential for direct involvement in DNA binding. While our study also identified Arg464 as a major contributor to DNA binding by ZBTB18, our rigorous modeling approach utilizing multiple templates to develop a model bound to cognate DNA motifs can provide fine detail into the energetic contributions to binding made by each residue, including Asn461 and His493. Crucially, we applied this modeling approach also to enable us to investigate the impact of ZBTB18 missense variants on DNA binding. Indeed, we found that substitution of Asn461 to serine alters the DNA-binding specificity of ZBTB18, and disrupts the capacity for this transcription factor to mediate transcriptional repression and radial migration. Furthermore, our models of ZBTB18 with mutated DNA motifs indicated that residues with very limited DNA contact, such as Arg495, might be relevant to the binding of spurious DNA motifs. Indeed, we find that the missense variant R495G dramatically altered the DNA-binding specificity, transcriptional repressor activity and neurobiological functions of ZBTB18. It has recently been established that nonbase-contacting residues are crucial to the evolution of C2H2 zinc-finger DNA binding (Najafabadi et al., 2017) and our results suggest that molecular modeling approaches are capable of informing the structural and functional consequences of substitution variants and nonbase-contacting residues.

Our biochemical characterization of N461S and R495G variants demonstrates that each variant displays distinct effects on its subcellular localization, as well as its capacity to mediate transcriptional repression in a luciferase reporter assay. The altered capacity for N461S to traffic into the nucleus could arise as a consequence of disruptions to its association with nuclear import proteins, thereby influencing its availability within the nucleus to mediate transcription. Also, we reported low steady-state levels of N461S observed in transfected cells and electroporated neurons, demonstrating that the N461S variant in unstable both in vitro and in vivo. This suggests that the N461S variant has a pathophysiological role with very low expression levels in cells. In the case of the R495G variant, we find that the presence of this variant influences the sub-nuclear localization of ZBTB18, leading to its accumulation in nuclear puncta, suggestive of protein aggregation. Over time, we find that the distribution patterns of N461S and R495G are not significantly different to wild type, hence we conclude that each missense variant likely causes a transient delay in the appropriate localization of ZBTB18 within cells, leading to asynchronization of its nuclear functions, ultimately disrupting its capacity to orchestrate neurodevelopment.

In addition to transient disruptions in subcellular localization, we find that the transcriptional repressor functions for each missense variant are disrupted in different ways. For example, we find that the N461S variant has lost the capacity for suppressing transcription via the BS10 and Rnd2 3’ enhancer constructs in our luciferase assays, and can potentiate transcriptional activation in certain contexts. We speculate that the impaired capacity for ZBTB18-mediated repression, together with low steady-state levels, explains the pathological impact of the N461S variant, leading to an imbalance in the regulation of downstream target genes. In contrast, we find that the R495G variant appears to have acquired strong transcriptional activation potential in vitro, with attendant consequences on gene regulation in cells. Given that ZBTB18 mediates its transcriptional regulatory functions in part through its recruitment of corepressors such as Dnmt3a (Fuks, Burgers, Godin, Kasai, & Kouzarides, 2001), we further speculate that N461S and R495G are altered in their capacity to bind protein partners for transcriptional regulation, and may recruit protein complexes that potentiate gene expression. Our capacity to visualize such changes in protein–DNA complexes in our EMSA experiments was confounded by the presence of endogenous ZBTB18 binding to the E1 motif in the lysates of mock-transfected cells. Also, our experiments with E2mut motifs did not yield specific complexes for ZBTB18 and its missense variants. Nevertheless, we did observe unique protein–DNA complexes formed by N461S on the E1mut motif, as well as by R495G on the E1, E1mut, and E2 motifs. Thus, we surmise that luciferase reporter activity mediated by N461S can be explained by its preferential binding to E1, E1mut, and E2 motifs, whereas R495G recruits additional factors to potentiate reporter activity upon binding to these motifs (Figure S5). As such, the “loss-of-repression” variant N461S and the “gain-of-transactivation” variant R495G are postulated to influence gene expression during neurodevelopment, likely through dysregulation of downstream genes that influence the radial positioning and differentiation of neurons.

Through a series of functional assays, we found that the presence of N461S and R495G affects the development and positioning of newborn neurons within the embryonic cerebral cortex in different ways. Notably, the impact on transcriptional regulation during radial migration is likely to be two-fold. First, the presence of each missense variant could lead to spurious binding of DNA regulatory sequences, resulting in changes in gene expression levels beyond ZBTB18 targets. Secondly, the presence of each missense variant has a unique effect on radially migrating neurons within the embryonic cerebral cortex. Thus, our results support a scenario in which the “loss-of-repression” variant, N461S, has a different impact on radial migration, in contrast to the “gain-of-transactivation” variant, R495G (see Figure S10 for an illustrative summary). In such a model, a loss-of-repression owing to the presence of the N461S variant influences gene expression regulation within radially migrating neurons, leading to their exuberant migration and preferential insertion within the Cortical Plate as multipolar-shaped cells, suggestive of disrupted radialglial-guided locomotion. In contrast, the R495G variant potentiates gene expression so as to severely impair their radial migration, culminating in a failure in the multipolar-to-bipolar transition of cortical cells, and their accumulation within the IZ. Indeed, we previously reported that too much or too little Rnd2 disrupts radial migration, such that its overexpression in cortical neurons impairs the capacity for cells to adopt appropriate morphologies to undergo movement (Heng et al., 2008; Pacary et al., 2011). Consistent with this interpretation, our analysis of IZ and CP
neurons in the rescue assays revealed that the presence of the R495G variant augmented the multipolar phenotype of Zbtb18 shRNA-treated neurons, thereby suggesting an additive effect on cell shape which could underlie its negative impact on radial migration. In contrast, treatment with N461S did not have such an effect in Zbtb18 shRNA-treated cells in the IZ or the CP. In addition, Zbtb18 shRNA-treated CP neurons coelectroporated with R495G extended longer leading neurites while N461S treatment did not have such an effect, further suggesting different cellular mechanisms that underlie the aberrant migration of treated neurons in each respective condition. Based on our results, we surmise that the presence of each substitution variant has a significant effect on the development of neurons within the mammalian brain. Future studies will address how each substitution variant disrupts the long-term positioning, dendritic branching and synaptic connectivity of cortical neurons within the postnatal cortex.

It has been reported that more than half of all pathogenic variants identified in ZBTB18 are predicted to be truncating, suggesting that haploinsufficiency/loss-of-function represents a general pathomechanism (Depienne et al., 2017). Yet, the majority of disease-associated missense variants to ZBTB18 lie within the C-terminal zinc-finger DNA-binding domain. Therefore, our investigation of the N461S and R495G variants has been crucial to demonstrate that altered transcriptional regulation also represents an important pathological mechanism for ZBTB18 variants in brain developmental disease. Altogether, our in silico models and functional studies demonstrate that missense variants to base-contact and non-base-contact residues influence DNA binding, transcriptional repression, and neurodevelopmental signaling.

The differences in functions for each pathogenic missense variant draws parallels with the differences in phenotypic presentation of the two human subjects diagnosed with ZBTB18 variants (Cohen et al., 2017; Rauch et al., 2012). Particularly, while intellectual disability was diagnosed in both subjects, the subject harboring an N461S variant was also diagnosed with microcephaly and hypoplasia of the corpus callosum (Cohen et al., 2017). In contrast, the subject diagnosed with the R495G variant was diagnosed with macrocephaly. We speculate that the presence of missense variants to the zinc-finger domain of ZBTB18 influences its capacity to bind DNA in a sequence-specific fashion, affect its transcriptional regulatory functions, and negatively impact on neural development to cause a spectrum of brain disorders in children. Functional studies of common missense variants that lie within the DNA-binding domain of ZBTB18 will also be important to clarify their potential impact on transcriptional regulation in human health and disease.

ACKNOWLEDGMENTS

I.A.H. and O.C. performed biochemistry and molecular biology studies with J.I.-T.H., H.L.N., H.B.S., D.U., K.D.G.P., H.D.C., and I.E.G-N. performed in utero electroporation experiments with L.N., O.C., and M.A. performed structural modeling calculations in consultation with J.I.-T.H. I.A.H., O.C., J.I.-T.H., I.A.H., O.C., and M.A. wrote the manuscript and prepared figures, with all authors providing comment.

DATA AVAILABILITY

The datasets analyzed during the current study are available from the corresponding author on reasonable request.

FUNDING INFORMATION

Curtin Research Fellowship [CRF120006 to M.A.]; Raine Priming Grant from the Raine Medical Research Foundation [M.A.]; National Health and Medical Research Council of Australia [1011505 to J.I.-T.H.]; Telethon-PeRth Children’s Hospital Research Fund [J.I.-T.H.]; NHMRC RD Wright Biomedical Research Fellowship [1085842 to K.D.G.P.]; Work was supported by computational resources provided by the Australian Government through the Pawsey Supercomputing Center under the National Computational Merit Allocation Scheme [pa6/pawsey0196].

CONFLICT OF INTERESTS

The authors declare that they have no conflict of interests.

ORCID

Mark Agostino http://orcid.org/0000-0002-1799-0392
Julian I-T. Heng http://orcid.org/0000-0002-0378-7078

REFERENCES

Abraham, M. J., Murtola, T., Schulz, R., Páll, S., Smith, J. C., Hess, B., & Lindahl, E. (2015). GROMACS: High performance molecular simulations through multi-level parallelism from laptops to supercomputers. SoftwareX, 1-2, 19-25.
Aoki, K., Meng, G., Suzuki, K., Takashi, T., Kameoka, Y., Nakahara, K., ... Kasai, M. (1998). RP58 associates with condensed chromatin and mediates a sequence-specific transcriptional repression. Journal of Biological Chemistry, 273(41), 26698–26704.
Bussi, G., Donadio, D., & Parrinello, M. (2007). Canonical sampling through velocity rescaling. Journal of Chemical Physics,126(1), 014101.(618,567),(947,592) https://doi.org/10.1063/1.2408420doi:Art1014101
Cohen, J. S., Srivastava, S., Farwell Hagman, K. D., Shinde, D. N., Huether, R., Darcy, D., ... Fatemi, A. (2017). Further evidence that de novo missense and truncating variants in ZBTB18 cause intellectual disability with variable features. Clinical Genetics, 91(5), 697–707. https://doi.org/10.1111/cge.12861
Cruickshank, M. N., Dods, J., Taylor, R. L., Karimi, M., Fenwick, E. J., Quail, E. A., ... Ugliati, D. (2015). Analysis of tandem E-box motifs within human complement receptor 2 (CR2/CD21) promoter reveals cell specific roles for RP58, E2A, USF and localized chromatin accessibility. International Journal of Biochemistry and Cell Biology, 64, 107–119. https://doi.org/10.1016/j.biocel.2015.03.016
Deciphering Developmental Disorders, S. (2017). Prevalence and architecture of de novo mutations in developmental disorders. Nature, 542(7642), 433–438. https://doi.org/10.1038/nature21062
Depienne, C., Nava, C., Keren, B., Heide, S., Rastetter, A., Passemard, S., ... Mignot, C. (2017). Genetic and phenotypic dissection of 1q43q44 microdeletion syndrome and neurodevelopmental phenotypes associated with mutations in ZBTB18 and HNRNPU. Human Genetics, 136(4), 463–479. https://doi.org/10.1007/s00439-017-1772-0
Essmann, U., Perera, L., Berkowitz, M. L., Darden, T., Lee, H., & Pedersen, L. G. (1995). A smooth particle mesh ewald potential. Journal of Chemical Physics, 103, 8577–8592.
van der Schoot, V., de Munnik, S., Venselaar, H., Elting, M., Mancini, G. M. S., Ravenswaaij-Arts, C. M. A., ... Stevens, S. J. C. (2018). Toward clinical and molecular understanding of pathogenic variants in the ZBTB18 gene. *Molecular Genetics & Genomic Medicine, 6*(3), 393–400. https://doi.org/10.1002/mgg3.387

Segal, D. J., Crotty, J. W., Bhakta, M. S., Barbas, C. F., 3rd, & Horton, N. C. (2006). Structure of Aart, a designed six-finger zinc finger peptide, bound to DNA. *Journal of Molecular Biology, 363*(2), 405–421. https://doi.org/10.1016/j.jmb.2006.08.016

Shields, G. C., Laughton, C. A., & Orozco, M. (1997). Molecular dynamics simulations of the d(T·A·T) triple helix. *Journal of the American Chemical Society, 119*, 7463–7469.

Shields, G. C., Laughton, C. A., & Orozco, M. (1998). Molecular dynamics simulation of a PNA·DNA·PNA triple helix in aqueous solution. *Journal of the American Chemical Society, 120*, 5895–5904.

Sousa da Silva, A. W., & Vranken, W. F. (2012). ACPYPE − AnteChamber PYthon Parser interface. *BMC Research Notes, 5*, 367. https://doi.org/10.1186/1756-0500-5-367

Weiser, J., Shenkin, P. S., & Still, W. C. (1999). Approximate atomic surfaces from linear combinations of pairwise overlaps (LCPO). *Journal of Computational Chemistry, 20*(2), 217–230. https://doi.org/10.1002/(SICI)1096-987x(19990130)20:2<217::AID-Jcc4>3.0.CO;2-A

Xiang, C., Baubet, V., Pal, S., Holderbaum, L., Tatard, V., Jiang, P., ... Dahmane, N. (2011). RP58/ZNF238 directly modulates proneurogenic gene levels and is required for neuronal differentiation and brain expansion. *Cell Death and Differentiation, 19*, 692–702. https://doi.org/10.1038/cdd.2011.144doi:cdd2011144

Zgarbova, M., Luque, F. J., Sponer, J., Cheatham, T. E., 3rd, Otýepka, M., & Jurecka, P. (2013). Toward improved description of DNA backbone: Revisiting Epsilon and Zeta Torsion Force Field Parameters. *Journal of Chemical Theory and Computation, 9*(5), 2339–2354. https://doi.org/10.1021/ct400154j

**SUPPORTING INFORMATION**

Additional supporting information may be found online in the Supporting Information section.

---

**How to cite this article:** Hemming IA, Clément O, Gladwyn-Ng IE, et al. Disease-associated missense variants in ZBTB18 disrupt DNA binding and impair the development of neurons within the embryonic cerebral cortex. *Human Mutation*. 2019;40:1841–1855. https://doi.org/10.1002/humu.23803