neurogenetics
https://doi.org/10.1007/s10048-021-00651-8

ORIGINAL ARTICLE

Pitt–Hopkins syndrome: phenotypic and genotypic description of four unrelated patients and structural analysis of corresponding missense mutations

Tingting Zhao[1] · Georgi Z. Genchev[1,2,3] · Shengnan Wu[4] · Guangjun Yu[1] · Hui Lu[1,3] · Jincai Feng[5]

Received: 7 April 2021 / Accepted: 30 May 2021

Abstract

Pitt–Hopkins syndrome is an underdiagnosed neurodevelopmental disorder which is characterized by specific facial features, early-onset developmental delay, and moderate to severe intellectual disability. The genetic cause, a deficiency of the TCF4 gene, has been established; however, the underlying pathological mechanisms of this disease are still unclear. Herein, we report four unrelated children with different de novo mutations (T606A, K607E, R578C, and V617I) located at highly conserved sites and with clinical phenotypes which present variable degrees of developmental delay and intellectual disability. Three of these four missense mutations have not yet been reported. The patient with V617I mutation exhibits mild intellectual disability and has attained more advanced motor and verbal skills, which is significantly different from other cases reported to date. Molecular dynamics simulations are used to explore the atomic level mechanism of how missense mutations impair the functions of TCF4. Mutations T606A, K607E, and R578C are found to affect DNA binding directly or indirectly, while V617I only induces subtle conformational changes, which is consistent with the milder clinical phenotype of the corresponding patient. The study expands the mutation spectrum and phenotypic characteristics of Pitt–Hopkins syndrome, and reinforces the genotype–phenotype correlation and strengthens the understanding of phenotype variability, which is helpful for further investigation of pathogenetic mechanisms and improved genetic counseling.

Keywords Pitt–Hopkins syndrome · TCF4 · Missense mutations · Phenotypic diversity · Molecular dynamics simulation

Background

Pitt–Hopkins syndrome (PTHS, MIM #610,954) is a neurodevelopmental disorder characterized by early-onset developmental delay, moderate to severe intellectual disability, specific facial features, and breathing anomalies [1, 2]. It was first described in 1978 [1] and in 2007 its cause was proposed to be haploinsufficiency of the TCF4 gene (transcription factor 4, OMIM *602,272) located on 18q21.2 [3, 4]. TCF4 is known to be taking part in a variety of biological process such as B- and T-cell development [5, 6], epithelial mesenchymal transition [7], and the functioning of the central and peripheral nervous system [8, 9]. It encodes a broadly expressed basic helix–loop–helix (bHLH) protein, also called E-protein, which is a DNA binding domain that binds to the enhancer-box (E-box) element 5'-CANNTG-3' (where N = any nucleotide). The bHLH domain is the only functionally characterized domain of the TCF4 protein, which functions as a homo- or heterodimer [4, 10]. Recently, the structure of human TCF4

✉ Hui Lu
huilu@sjtu.edu.cn

✉ Jincai Feng
fengjc@shchildren.com.cn

1 Shanghai Engineering Research Center for Big Data in Pediatric Precision Medicine, Center for Biomedical Informatics, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China

2 Bulgarian Institute for Genomics and Precision Medicine, Sofia, Bulgaria

3 SJTU-Yale Joint Center for Biostatistics, Department of Bioinformatics and Biostatistics, School of Life Sciences and Biotechnology, Shanghai Jiao Tong University, Shanghai, China

4 Molecular Diagnostic Laboratory, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China

5 Department of Rehabilitation, Shanghai Children's Hospital, Shanghai Jiao Tong University, Shanghai, China

Published online: 14 June 2021

C-terminal bHLH domain in complex with 13-bp oligonucleotide containing the E-box sequence was published (PDB 6OD3) [11], which provides a solid structural basis to study the effects of PTHS-associated mutations in the TCF4 bHLH domain.

There are about 500 PTHS patients reported to date [2], but the prevalence of the disease is still unknown and it is widely believed that PTHS is underdiagnosed. Diverse pathogenic variants have been found in patients with PTHS, such as nonsense, frameshift, splice-site, and missense mutations as well as translocations and large deletions encompassing all or part of TCF4 [2–4, 12–14]. The most common ones are gene disruptive or truncating variants including frameshifting, nonsense, and splice-site single nucleotide variants and large deletions at 18q21 encompassing the TCF4 gene [2, 15]. The rarity of the syndrome, complexity of the TCF4 gene, and similarity to other well-recognized syndromes such as Angelman syndrome or Rett syndrome hinder a deeper understanding of the underlying pathological mechanisms [2, 16]. Up to now, only breathing anomalies in patients with PTHS are considered to be caused by impaired noradrenergic neuronal development due to defective TCF4 interaction with the ASCL1-PHOX-RET pathway [3, 13]. Known pathogenic missense variants usually cluster in the highly conserved bHLH domain encoded by exon 18 [2, 17]. Related studies show that the impact of missense mutations ranges from subtle deficiencies to dominant-negative effects and presents to a varying extent, which may be contributing to the phenotypic variability of PTHS patients [18]. However, detailed phenotype data for the vast majority of PTHS cases is unavailable.

Herein, we report four PTHS patients with molecularly confirmed missense mutation (3 novel and 1 reported) in the exon18 of TCF4 and describe their phenotype in detail. These four cases harboring missense mutation show variable degrees of developmental delay and intellectual disability. Moreover, we explore how these missense mutations impair the functions of TCF4 by using molecular dynamics (MD) simulation and protein structure analysis tool. Furthermore, we aim to explain the phenotypic variability of PTHS patients with the effect of these missense mutations.

Methods

Patients

This study was approved by the Ethics Committee of Children's Hospital of Shanghai. Retrospective clinical data was collected on four patients with a molecular diagnosis of Pitt–Hopkins syndrome (PTHS) referred to the Shanghai Children's Hospital. Informed consent for publication of clinical data and genetic test results identified by whole exome sequencing (WES) was obtained. A clinical diagnostic score was assigned to each patient based on the latest proposed international scoring system [2], where score ≥ 9 indicates molecular confirmation and score of 6 to 8 with the presence of the facial characteristics raises suspicion for PTHS and requires further confirmation by molecular testing. The mutation coordinates were given according to the TCF4 transcript variant 1 (NM_001083962.2) and GRCh37 for all patients. Related therapeutic interventions utilized were also discussed.

Bioinformatics analysis

We investigated the effects of missense mutations on exon 18 from the prospective of protein structural analysis. The crystal structure complex of human TCF4 C-terminal bHLH domain and 13-bp oligonucleotide containing an E-box sequence (PDB code: 6OD3) were downloaded from the Protein Data Bank (PDB) [11]. To strengthen the reliability of our MD results, mutation R578H which had been investigated before [18] was included in our analysis. Complexes of DNA-mutant protein (T606A, K607E, V617I, R578C, and R578H) were built with PyMOL 2.4 [19] based on the wild-type (WT) structure. Commonly used computational methods (including PolyPhen-2 [20], PROVEAN [21], FATHMM [22], MutationTaster [23], and SIFT [24]) were used to predict the pathogenicity of the variants. ProtParam [25] (https://web.expasy.org/protparam/) was utilized to study the physical and chemical properties of the protein.

MD simulations were performed using the Amber16 package [26] following a five-step protocol including PDB preparation, minimization, heating, equilibration, and production. The force field parameters retrieved from ff14SB and bsc1 for the protein and DNA were applied. Each ensemble was solvated in a box of the *opc* water model [27] with a padding of 10 Å. Appropriate amounts of Na^+ and Cl^- ions were added to each ensemble to mimic the physiological environment and ensure zero net charges. After minimization, heating, and equilibration, the production runs were simulated at 298 K for 60 ns through 30 million steps of NPT simulations with 2-fs step size. Five independent trajectories of 60 ns for each of the 6 molecular complexes were simulated. For each simulation, sampling was conducted every 5 ps (12,000 snapshots for 60-ns simulations). In total, 1.8 μs of trajectories was collected for the WT and five mutant systems.

The embedded tools in Amber were used to analyze the resultant trajectories. Average structures were defined as the structure ensembles of lowest energy among all five trajectories for the last 10 ns [28, 29]. Binding free energies between TCF4 bHLH domain and DNA sequence for the last 10 ns in all the simulation trajectories were calculated separately with Molecular Mechanics-Generalized Born Surface Area

(MM/GBSA) method [30] using MMPBSA.py [31] and the average value was used as an indicator of the TCF4-DNA binding affinity.

Results

Patient description

All four patients showed typical facial features (Table 1), but with variable degrees of disease disabilities. All patients had no family history of intellectual disability and other neurological or psychiatric disorders. The identified novel variants were absent in large population databases such as dbSNP and the Exome Aggregation Consortium (ExAC) database [32].

Patient 1 is a 4.5-year-old Chinese boy who carries a novel TCF4 missense mutation c.1816A > G (p.T606A). He showed the typical facial features (Table 1) and severe intellectual disability delay with absent or limited speech, but was able to comprehend simple task instructions and play with others. Developmentally, he was looking up at 2 months, rolling at 5 months, sitting at 10 months, standing at 14 months, and walking at 17 months, but in an unstable manner until the present time. Moreover, he exhibited severe language delay; he was able to babble at the age of 3 years and had 10 words at 4 years. He had breathing regulation anomalies (intermittent hyperventilation and/or apnea), myopia, and a history of constipation, but no sleep problems. He recently began to receive speech therapy and stem cell therapy. Retrospective analysis of patient 1 phenotype and comparison with PTHS diagnostic criteria [2] showed 12 points (8 cardinal and 4 supportive points).

Patient 2 is a 12-year-old Chinese girl identified with a novel TCF4 missense mutation c.1819A > G (p.K607E). The girl is the second child of a nonconsanguineous pair and has a healthy sister. The proband was delivered at term with hypoxia. The proband showed the typical facial features (Table 1) and suffered from constipation, myopia, constipation, breathing regulation anomalies, and severe intellectual disability delay. She spoke her first word at 2 years old and developed 2-word phrases until present. She could walk at 4 years, but with an unstable, ataxic gait. She also has learning disabilities and cognitive difficulties. Retrospective analysis of patient 2 phenotype and comparison with PTHS diagnostic criteria [2] showed 12 points (8 cardinal and 4 supportive points).

Patient 3 is a 2.5-year-old Chinese girl, with a novel variant c.1849G > A (p.V617I) located in exon 18 of TCF4. She is one of the mildest affected patients reported to date. She was rolling at 4 months, sitting at 7 months, and crawling at 13 months. She showed development delay on her initial intelligence test at 20 months; she received rehabilitation training at 21 months, including physical therapy, speech therapy, and occupation therapy. She was able to walk at 21 months and at present she can ambulate independently and run normally, although she cannot jump. She had 20 words at 2 years and had 2-word phrases on an occasional basis. She exhibits slightly worse command and cognition than her peers, and is able to count to 10 now. She began to eat and dress by herself, although is not good at it. In addition, she was described to usually have a smiling appearance, but an anxious or agitated disposition. Notably, she has never demonstrated any abnormal breathing or seizure-like activity. Retrospective analysis of patient 3 phenotype and comparison with PTHS diagnostic criteria [2] showed 5 points (4 cardinal and 1 supportive points).

Patient 4 is a 1.5-year-old Chinese boy with a reported TCF4 missense mutation c.1732C > T (p.R578C). He showed typical facial features (Table 1). He was looking up at 2 months, rolling at 4 months, sitting at 11 months, crawling at 13 months, and pulling to stand at 16 months and beginning to babble. He can understand basic instructions. Repetitive movements were also described as he was easily getting excited, beating the table or toys and shaking his head. Reported was neither constipation nor seizure-like activity, and his vision had not yet been evaluated. No abnormalities were found on brain magnetic resonance imaging (MRI) and electroencephalography examination. He recently received rehabilitation training including physical therapy, brain acupuncture, and electrotherapy. At the same time, he also took an injection of mouse nerve growth factor and sevoflurane oral solution. Retrospective analysis of patient 3 phenotype and comparison with PTHS diagnostic criteria [2] showed 6 points (4 cardinal and 2 supportive points).

Table 1 Physical characteristics of the four reported patients

ID	Gender	Age	Square forehead	Thin lateral eyebrows	Wide nasal bridge/ridge/tip	Flared nasal alae	Full cheeks/prominent midface	Wide mouth/full lips/cupid bow upper lip	Thickened/overfolded helices	Slender fingers
1	Male	4.5	1	1	1	1	1	1	1	1
2	Female	12	1	1	1	1	1	1	1	1
3	Female	2.5	1	1	1	0	1	1	1	1
4	Male	1.5	1	1	1	1	1	1	1	1

Class in Vertebrata	Species	TCF4 bHLH Region helix 1 helix 2	GenBank#
Mammalia	*Homo sapiens*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	CBY80191
Mammalia	*Pan troglodytes*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	XP_016789244
Aves	*Gallus gallus*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	Q90683
Reptilia	*Anolis carolinensis*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	XP_016850729
Amphibia	*Xenopus tropicalis*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	NP_001096226
Osteichthyes	*Lepisosteus oculatus*	KERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	XP_015217707
Chondrichthyes	*Rhincodon typus*	RERRMANNARERLRVRDINEAFKELGRMVQLHLKSDKPQTKLLILHQAVAVILSLEQQVRE	XP_020365982

Class in Vertebrata	Species	TCF12 bHLH Region helix 1 helix 2	GenBank#
Mammalia	*Homo sapiens*	KERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRE	XP_011520261
Mammalia	*Pan troglodytes*	KERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRE	XP_016783136
Aves	*Gallus gallus*	KERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRE	XP_015134250
Reptilia	*Anolis carolinensis*	RERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRE	XP_016853254
Amphibia	*Xenopus tropicalis*	KERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILNLEQQVRE	XP_002940299
Osteichthyes	*Lepisosteus oculatus*	RERRMANNARERLRVRDINEAFKELGRMCQLHLKSEKPQTKLLILHQAVAVILSLEQQVRE	XP_015198688
Chondrichthyes	*Rhincodon typus*	RERRVANNARERLRVRDINEAFKELGRMCQLHLNSDKPQTKLLILHQAVSVILNLEQQVRE	XP_020369765

Fig. 1 Sequences of the basic helix-loop-helix (bHLH) regions in proteins related to TCF4 are shown. TCF4 mutations investigated in this work are highlighted in red

Bioinformatics analysis of corresponding missense mutations

The four identified missense mutations are highly conserved in all classes of Vertebrata, and in the related transcription factor TCF12 (Fig. 1). The pathogenicity of the variant prediction results was consistent in the majority of algorithms used and showed to be damaging/deleterious (Supplementary Table 1). Table 2 shows the physical and chemical properties of the wild-type and mutant bHLH domain predicted by ProtParam. Mutant R578C and K607E showed a decrease in the theoretical pI. Moreover, R578C also showed a modest increase in the instability index and grand average of hydropathicity (GRAVY).

To explore the molecular mechanism underlying variable disease severity, we first mapped these four mutations in exon18 to the homodimer structure models in PDB 6OD3 (Fig. 2a) and investigated the molecular consequences of the pathogenetic mutations with MD simulations. The root mean squared deviation (RMSD) was calculated for every trajectory (Supplemental Fig. 1) showing that 60 ns of simulation time was sufficient for the equilibration of the WT and mutants at the NPT ensemble.

K607 and R578 are highly conserved amino acids and located at the beginning site of helix 1 and the middle site

Table 2 Analysis of wild-type and mutant bHLH domain by ProtParam

	Molecular weight	Theoretical pI	Instability index[a]	Aliphatic index	Grand average of hydropathicity (GRAVY)[b]
Wild type	7218.46	10.48	31.08	108.69	−0.621
R578C	7165.42	9.85	41.44	108.69	−0.507
T606A	7188.44	10.48	31.08	110.33	−0.580
K607E	7219.41	9.97	35.63	108.69	−0.615
V617I	7232.49	10.48	31.08	110.33	−0.616

[a]Instability index: instability index > 40, unstable; instability index < 40, stable

[b]The negative value of GRAVY stands for hydrophilic protein and lower negative value represents higher hydrophilicity

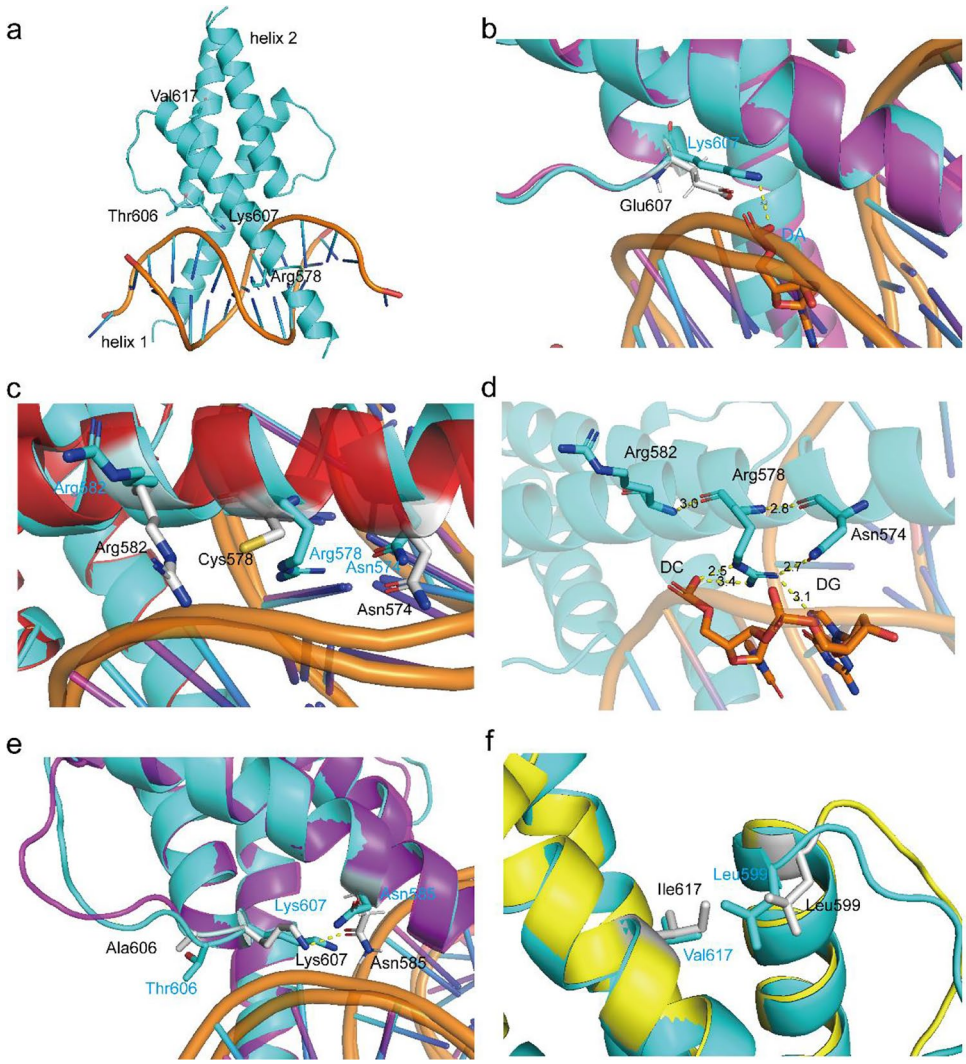

Fig. 2 The overall WT crystal structure and mutant models after MD simulation. **a** Mapping of PTHS-causing mutations in TCF4 crystal structure. V617 is located at the middle of helix 2, which is far away from the DNA. T606 sits at a conjunction site between a loop and helix 1 of the TCF4 bHLH domain near the basic DNA-binding region. K607 located at the beginning site of helix and R578 at the middle site of helix 1, which are both in the basic DNA-binding region. **b** The model of K607E mutant after MD simulation. **c** The model of R578C mutant after MD simulation. **d** R578 interacts with two neighboring phosphate groups (cytosine and guanine of the 5'-CACGTG -3'). **e** The model of T606A mutant after MD simulation. **f** The model of V617I mutant after MD simulation. Yellow dash lines represented hydrogen bonds. The cyan model representing the crystal structure and mutant average structure is colored by another color. Side chains are shown as sticks with carbon atoms colored in blue for the crystal structure and off-white for the mutant structure

of helix 1 respectively, which are both in the basic DNA-binding region (Fig. 2a). It is known that both of them are binding in the major groove of the DNA and are in contact with DNA phosphate groups [11]. Mutation K607E introduces a large negative charge in place of the positive charge at position 607. As observed in the MD simulation for this mutant (Fig. 2b), the favorable interaction between K607 and the DNA backbones disappears and the negatively charged DNA is pushed away from the protein. Similarly, mutation R578C introduces the polar glutamic acid in place of the positively charged arginine and interactions with the negatively charged DNA phosphate group are removed (Fig. 2c). Cascaded conformational changes are observed during the simulation. R582 which was not involved in the binding of either base or phosphate backbone of DNA replaces partly R578 to contact a phosphate group of the backbone of cytosine. Moreover, the N574●●●R578 interaction (Fig. 2d) which bridges the two phosphate groups neighboring the central cytosine [11] and works as stabilizing interaction at the protein-DNA interface is also lost, which leads to a slight movement of the helix α2 away from the DNA major groove. The R578C mutant is unstable as indicated by ProtParam.

The protein-DNA binding free energy for mutation K607E and R578C is significantly higher than that of the wild type (Table 3), signifying that the DNA binding affinity greatly drops. This is consistent with the evidence related to R578H, whose molecular consequences on DNA binding were examined experimentally by using in vitro translated proteins [18]. Hence, this further suggests that mutations K607E and R578C are set to abrogate DNA binding completely.

T606 sits at a conjunction site between a loop and helix 1 of the TCF4 bHLH domain near the basic DNA-binding region (Fig. 2a). When the polar T606 is replaced by a non-reactive alanine, the interactions between the polar side chain of the threonine and the DNA are removed, and the distance between the affected amino acid and DNA increases (Fig. 2e). Our MD results reveal that K607 tends to interact

Table 3 TCF4-DNA binding free energy with MM/GBSA

Contribution	WT	K607E	T606A	V617I	R578C	R578H
VDW	− 127.393	− 119.572	− 123.444	− 122.432	− 116.061	− 110.392
EEL	− 6510.051	− 5483.254	− 6430.466	− 6497.942	− 6007.643	− 5910.047
EGB	6477.221	5482.199	6405.493	6463.631	5998.923	5911.729
ESURF	− 18.806	− 17.638	− 18.219	− 18.288	− 17.327	− 16.336
ΔG_{Gas}	− 6637.444	− 5602.826	− 6553.910	− 6620.374	− 6123.704	− 6020.440
ΔG_{Solv}	6458.415	5464.561	6387.274	6445.343	5981.596	5895.392
ΔG_{Total}	− 179.029	− 138.265	− 166.636	− 175.031	− 142.108	− 125.047

[a] *VDW*, van der Waals energy; *EEL*, electrostatic energy; *EGB*, polar solvation energy; *ESURF*, non-polar solvation energy; ΔG_{Gas}, (VDW + EEL) gas-phase free energy; ΔG_{Solv}, (EGB + ESRUF) solvation free energy; ΔG_{Total}, (GGas + GSolv) total free energy

with E585 rather than with the DNA. Accordingly, DNA-binding free energy by T606A is higher compared to the binding energy of the WT (Table 3). Hence, the DNA binding affinity is affected. Moreover, the loop beside the mutation T606A undergoes a conformational change as well (Fig. 2e). We hypothesize that the T606A mutation together with the conformational change would very likely abolish the activity to recognize the DNA binding partner.

V617 is located at the middle of helix 2, and is further removed from the DNA (Fig. 2a). The replacement of V617 with an isoleucine residue does not change the hydrophobic property at this position but the length of side chain is enlarged, which leads to a slight movement of the beginning of helix α1 and induces subtle changes in the loop (Fig. 2f). Accordingly, mutation V617I does not destroy the hydrophobic microenvironment but the increased size of this substitution very likely causes local steric hindrance which may disturb local structure and folding energies.

Discussion

In this study, we reported four patients identified with a different de novo missense mutation in the bHLH domain of TCF4, and with variable degrees of developmental delay and intellectual disability. MD simulations were used to explore how these missense mutations specifically impair the function of TCF4. Mutations T606A, K607E, and R578C are found to affect DNA binding directly or indirectly, while V617I only induced subtle conformational changes. This is consistent with the fact that patient 3, who is afflicted with mutation V617I, has a milder clinical phenotype compared with the other cases reported and exhibited less prominent motor delays and to some degree language delay.

All of the patients in this work presented with typical facial gestalt, which is the main key feature for the diagnosis of PTHS and distinguishes it from other syndromes [2, 17]. Compared with the majority of cases reported in the literature where speech is often absent and there are severe motor delays (independent ambulation after 5 years of age or not at all) [14, 17, 33–37], the patients showed variable disease phenotype. In particular, patient 3 who has a missense mutation in exon18 exhibited mild special facial features and attained more advanced motor and verbal skills with mild intellectual disability. Similarly, in a Caucasian male with PTHS caused by a single-pair deletion in exon 19 of the TCF4 gene, with mild to moderate intellectual disability, use of full sentences and independent ambulation were reported [38]. More mildly affected individuals with TCF4 point mutations will be identified with the development of genetic tests [12]. Previous studies report that individuals with variants starting from exon 9 to exon 20 tend to present a typical PTHS phenotype [2, 39, 40], while those with variants affecting exons 1 to 5 in the TCF4 gene are associated with mild intellectual disability. These abovementioned unique cases showed that this conclusion has not held true in all PTHS cases. In addition, it is notable that testing for PTHS of patient 3 and patient 4 would not have been recommended based on the latest proposed international clinical diagnostic scores [2], partly because the patients were too young to show the symptoms, which indicates that the available systems and approaches are still limited especially for young patients [14, 33]. Hence, although there are published diagnostic guidelines for PTHS [2], there is still a need for additional descriptions of novel, especially unique, clinical cases [15] to expand the spectrum of PTHS.

These four cases show that different missense mutations within the same exon cause different degrees of intellectual disability including motor delay and speech delay. To explain the phenotypic variability of these PTHS patients with corresponding missense mutations, we used MD simulation to further explore each missense mutation consequence and attempt to match it to the patient's phenotype. To strengthen the reliability of our MD results, mutation R578H which has been investigated before [18] is included in our analysis and our results are consistent with the molecular consequences of R578H examined experimentally. Moreover, to further substantiate the pathogenicity of these missense

mutations, additional commonly used computational methods were used here. The prediction results are consistent in the majority of algorithms used. However, those computational algorithms tend to be overly dependent on the degree of conservation [41] and lack functional mechanism-based interpretations. In contrast, MD can study protein-DNA interactions and interpret dysfunction at the molecular level with atomic resolution; thus, we believe it can become an integral part of meta-prediction approaches [42].

As it is known that E-proteins usually form homodimers and heterodimers with tissue-specific bHLH factors [4, 10] and it has been verified that the impact of mutations in one subunit is similar to that in both subunits [18], with the goal to focus the MD functional study herein, we explored the consequences of mutation of the intra-TCF4 heterodimers that contain one WT and one mutant TCF4 subunit. Combining the MD results with the mutation functional analysis, we demonstrated that not all pathogenetic missense mutations in exon18 of TCF4 gene will result in the complete loss of function and invariable disease severity as in [18]. For example, patient 3 — who is harboring missense mutation in exon18 which does not affect DNA binding — presents mild disease phenotype, which helps to explain the phenotypic differences and the milder signs and symptoms.

No specific medication is so far known to be generally effective in individuals with PTHS [2, 43]. Hence, it is necessary to determine appropriate intervention services and educational strategies based on developmental assessments upon initial diagnosis [2]. Owing to the fast-developing sequencing technology and widespread use of WES in clinical application, the detection of mutations can improve our ability to accurately and early diagnose the patients, especially those mildly affected individuals [12] which will inform the timely intervention for individuals with PTHS and lead to improved treatment results. Patient 3 is a good example who received rehabilitation at her very young age (21 months) and gained great improvement such as independent ambulation and higher cognition in the following year. Similarly, patient 4 also received rehabilitation training. Figuring out how such missense mutations impair the protein function of TCF4 is important for further understanding of the genetic etiology and is also helpful for a more accurate diagnosis, prognosis, and management of this rare and severe disease.

Conclusions

The results of this study expand the mutation spectrum and phenotypic characteristics of PTHS. In addition, the structural analysis performed confirms that not all pathogenetic missense mutations in exon18 of TCF4 gene result in the complete loss-of-function and thereby affect the clinical phenotype or disease severity of the patients. A considerable effort is still needed to understand the phenotype variability and pathogenetic mechanisms, thus fostering the development of specific medications or therapeutic strategies to treat and manage this rare neurodevelopmental disorder.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s10048-021-00651-8.

Acknowledgements We wish to thank the patients and their families for their participation in this study.

Author contribution Conceptualization, J.F.; methodology, J.F. and H.L.; software, T.Z. and G.Z.G.; validation, T.Z. and G.Z.G.; formal analysis, J.F., T.Z., and G.Z.G.; investigation, J.F., T.Z., and G.Z.G.; writing—original draft preparation, T.Z. and G.Z.G.; writing—review and editing, J.F., H.L., and S.W.; visualization, T.Z. and G.Z.G.; supervision, J.F. and G.Y.; project administration, J.F. and T.Z.; funding acquisition, G.Y. All the authors have read and agreed to the published version of the manuscript.

Funding This research was funded by Shanghai Municipal Commission of Health and Family Planning (Grant No. 2018ZHYL0223).

Availability of data and materials All data generated or analyzed during this study are included in this published article and its supplementary information files.

Code availability Not applicable.

Declarations

Ethics approval The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Ethical Committee of Children's Hospital of Shanghai and informed consent was obtained from all the patients.

Competing interests The authors declare no competing interests.

References

1. Pitt D, Hopkins I (1978) A syndrome of mental retardation, wide mouth and intermittent overbreathing. Aust Paediatr J 14:182–184. https://doi.org/10.1111/jpc.1978.14.3.182
2. Zollino M, Zweier C, Van Balkom ID, Sweetser DA, Alaimo J, Bijlsma EK, Cody J, Elsea SH, Giurgea I, Macchiaiolo M et al (2019) Diagnosis and management in Pitt-Hopkins syndrome: First international consensus statement. Clin Genet 95:462–478. https://doi.org/10.1111/cge.13506
3. Zweier C, Peippo MM, Hoyer J, Sousa S, Bottani A, Clayton-Smith J, Reardon W, Saraiva J, Cabral A, Gohring I et al (2007) Haploinsufficiency of TCF4 causes syndromal mental retardation with intermittent hyperventilation (Pitt-Hopkins syndrome). Am J Hum Genet 80:994–1001. https://doi.org/10.1086/515583
4. Zweier C, Sticht H, Bijlsma EK, Clayton-Smith J, Boonen SE, Fryer A, Greally MT, Hoffmann L, den Hollander NS, Jongmans M et al (2008) Further delineation of Pitt-Hopkins syndrome: phenotypic and genotypic description of 16 novel patients. J Med Genet 45:738–744. https://doi.org/10.1136/jmg.2008.060129

5. Zhuang, Cheng, Weintraub (1996) B-Lymphocyte development is regulated by the combined dosage of three basic helix-loop-helix genes, E2A, E2–2, and HEB. Mol Cell Biol. https://doi.org/10.1128/MCB.16.6.2898
6. Bergqvist E, Saarikettu C, Grundstron H (2000) The basic helix-loop-helix transcription factor E2–2 is involved in T lymphocyte development. Eur J Immunol 30:2857–2863. https://doi.org/10.1002/1521-4141(200010)30:10%3C2857::AID-IMMU2857%3E3.0.CO;2-G
7. Cano A, Portillo F (2010) An emerging role for class I bHLH E2–2 proteins in EMT regulation and tumor progression. Cell Adh Migr 4:56–60. https://doi.org/10.4161/cam.4.1.9995
8. Bertrand N, Castro DS, Guillemot F (2002) Proneural genes and the specification of neural cell types. Nat Rev Neurosci 3:517–530. https://doi.org/10.1038/nrn874
9. Heng JIT, Tan SS (2010) The role of class I HLH genes in neural development–have they been overlooked? BioEssays 25:709–716. https://doi.org/10.1002/bies.10299
10. Corneliussen B, Thornell A, Hallberg B, Grundstrm T (1991) Helix-loop-helix transcriptional activators bind to a sequence in glucocorticoid response elements of retrovirus enhancers. J Virol 65:6084–6093. https://doi.org/10.1128/jvi.65.11.6084-6093.1991
11. Yang J, Horton JR, Li J, Huang Y, Zhang X, Blumenthal RM, Cheng X (2019) Structural basis for preferential binding of human TCF4 to DNA containing 5-carboxylcytosine. Nucleic Acids Res 47:8375–8387. https://doi.org/10.1093/nar/gkz381
12. Rosenfeld JA, Leppig K, Ballif BC, Thiese H, Erdie-Lalena C, Bawle E, Sastry S, Spence JE, Bandholz A, Surti U et al (2009) Genotype-phenotype analysis of TCF4 mutations causing Pitt-Hopkins syndrome shows increased seizure activity with missense mutations. Genet Med 11:797–805. https://doi.org/10.1097/GIM.0b013e3181bd38a9
13. de Pontual L, Mathieu Y, Golzio C, Rio M, Malan V, Boddaert N, Soufflet C, Picard C, Durandy A, Dobbie A et al (2009) Mutational, functional, and expression studies of the TCF4 gene in Pitt-Hopkins syndrome. Hum Mutat 30:669–676. https://doi.org/10.1002/humu.20935
14. Goodspeed K, Newsom C, Morris MA, Powell C, Evans P, Golla S (2018) Pitt-Hopkins syndrome: a review of current literature, clinical approach, and 23-patient case series. J Child Neurol 33:233–244. https://doi.org/10.1177/0883073817750490
15. Sparber P, Filatova A, Anisimova I, Markova T, Voinova V, Chuhrova A, Tabakov V, Skoblov M (2020) Various haploinsufficiency mechanisms in Pitt-Hopkins syndrome. Eur J Med Genet 63. https://doi.org/10.1016/j.ejmg.2020.104088
16. Kousoulidou L, Alexandrou A, Papaevripidou I, Evangelidou P, Tanteles G, Anastasiadou VC, Sismani C (2019) Two unrelated individuals carrying rare mosaic deletions in TCF4 gene. Am J Med Genet A 179:134–138. https://doi.org/10.1002/ajmg.a.60692
17. Whalen S, Héron D, Gaillon T, Moldovan O, Rossi M, Devillard F, Giuliano F, Soares G, Mathieu-Dramard M, Afenjar A et al (2012) Novel comprehensive diagnostic strategy in Pitt-Hopkins syndrome: clinical score and further delineation of the TCF4 mutational spectrum. Hum Mutat 33:64–72. https://doi.org/10.1002/humu.21639
18. Sepp M, Pruunsild P, Timmusk T (2012) Pitt-Hopkins syndrome-associated mutations in TCF4 lead to variable impairment of the transcription factor function ranging from hypomorphic to dominant-negative effects. Hum Mol Genet 21:2873–2888. https://doi.org/10.1093/hmg/dds112
19. Delano WL (2002) The PyMOL Molecular Graphics System, DeLano Scientific, San Carlos
20. Adzhubei IA, Schmidt S, Peshkin L, Ramensky VE, Gerasimova A, Bork P, Kondrashov AS, Sunyaev SR (2010) A method and server for predicting damaging missense mutations. Nat Methods 7:248–249. https://doi.org/10.1038/nmeth0410-248
21. Choi Y, Chan AP (2015) PROVEAN web server: a tool to predict the functional effect of amino acid substitutions and indels. Bioinformatics 31:2745–2747. https://doi.org/10.1093/bioinformatics/btv195
22. Shihab HA, Gough J, Cooper DN, Stenson PD, Barker GL, Edwards KJ, Day IN, Gaunt TR (2013) Predicting the functional, molecular, and phenotypic consequences of amino acid substitutions using hidden Markov models. Hum Mutat 34:57–65. https://doi.org/10.1002/humu.22225
23. Schwarz JM, Cooper DN, Schuelke M, Seelow D (2014) MutationTaster2: mutation prediction for the deep-sequencing age. Nat Methods 11:361–362. https://doi.org/10.1038/nmeth.2890
24. Kumar P, Henikoff S, Ng PC (2009) Predicting the effects of coding non-synonymous variants on protein function using the SIFT algorithm. Nat Protoc 4:1073–1081. https://doi.org/10.1038/nprot.2009.86
25. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A (2003) ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res 31:3784–3788. https://doi.org/10.1093/nar/gkg563
26. Case DA, Cheatham TE 3rd, Darden T, Gohlke H, Luo R, Merz KM Jr, Onufriev A, Simmerling C, Wang B, Woods RJ (2005) The Amber biomolecular simulation programs. J Comput Chem 26:1668–1688. https://doi.org/10.1002/jcc.20290
27. Izadi S, Anandakrishnan R, Onufriev AV (2014) Building water models: a different approach. J Phys Chem Lett 5:3863–3871. https://doi.org/10.1021/jz501780a
28. Zhang J, Luo H, Liu H, Ye W, Luo R, Chen HF (2016) Synergistic modification induced specific recognition between histone and TRIM24 via fluctuation correlation network analysis. Rep 6:24587. https://doi.org/10.1038/srep24587
29. Roe DR, Cheatham TE 3rd (2013) PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. J Chem Theory Comput 9:3084–3095. https://doi.org/10.1021/ct400341p
30. Fogolari F, Brigo A, Molinari H (2003) Protocol for MM/PBSA molecular dynamics simulations of proteins. Biophys J 85:159–166. https://doi.org/10.1016/S0006-3495(03)74462-2
31. Miller BR III, Mcgee TD Jr, Swails JM, Homeyer N, Gohlke H, Roitberg AE (2012) MMPBSA.py: an efficient program for end-state free energy calculations. J Chem Theory Comput. https://doi.org/10.1021/ct300418h
32. Lek M, Karczewski KJ, Minikel EV, Samocha KE, Banks E, Fennell T, O'Donnell-Luria AH, Ware JS, Hill AJ, Cummings BB et al (2016) Analysis of protein-coding genetic variation in 60,706 humans. Nature 536:285–291. https://doi.org/10.1038/nature19057
33. De Winter CF, Baas M, Bijlsma EK, Van Heukelingen J, Routledge S, Hennekam RCM (2016) Phenotype and natural history in 101 individuals with Pitt-Hopkins syndrome through an internet questionnaire system. Orphanet J Rare Dis 11:1–12. https://doi.org/10.1186/s13023-016-0422-2
34. Sweatt JD (2013) Pitt-Hopkins syndrome: intellectual disability due to loss of TCF4-regulated gene transcription. Exp Mol Med 45:e21. https://doi.org/10.1038/emm.2013.32
35. Balkom IDCV, Vuijk PJ, Franssens M, Hoek HW, Hennekam RCM (2012) Development, cognition, and behaviour in Pitt-Hopkins syndrome. Dev Med Child Neurol 54:925–931. https://doi.org/10.1111/j.1469-8749.2012.04339.x
36. Laura T, Mari S, Timmusk T, Palgi M (2015) Introducing Pitt-Hopkins syndrome-associated mutations of TCF4 to Drosophila daughterless. Biology open 4:1762–1771. https://doi.org/10.1242/bio.014696
37. Rannals MD, Page SC, Campbell MN, Gallo RA, Mayfield B, Maher BJ (2016) Neurodevelopmental models of transcription factor 4 deficiency converge on a common ion channel as

a potential therapeutic target for Pitt Hopkins syndrome. Rare Diseases 4:43–55. https://doi.org/10.1080/21675511.2016.1220468

38. Tan A, Goodspeed K, Edgar VB (2018) Pitt-Hopkins syndrome: a unique case study. J Int Neuropsychol Soc 24:995–1002. https://doi.org/10.1017/s1355617718000668
39. Bedeschi MF, Marangi G, Calvello MR, Ricciardi S, Leone FPC, Baccarin M, Guerneri S, Orteschi D, Murdolo M, Lattante S et al (2017) Impairment of different protein domains causes variable clinical presentation within Pitt-Hopkins syndrome and suggests intragenic molecular syndromology of TCF4. Eur J Med Genet 60:565–571. https://doi.org/10.1016/j.ejmg.2017.08.004
40. Tripon F, Bogliș A, Micheu C, Streață I, Bănescu C (2020) Pitt-Hopkins syndrome: clinical and molecular findings of a 5-year-old patient. Genes 11. https://doi.org/10.3390/genes11060596
41. Sun H, Yu G (2019) New insights into the pathogenicity of non-synonymous variants through multi-level analysis. Sci Rep 9:1667. https://doi.org/10.1038/s41598-018-38189-9
42. Chi YI, Stodola TJ, De Assuncao TM, Levrence EN, Tripathi S, Dsouza NR, Mathison AJ, Basel DG, Volkman BF, Smith BC et al (2021) Molecular mechanics and dynamic simulations of well-known Kabuki syndrome-associated KDM6A variants reveal putative mechanisms of dysfunction. Orphanet J Rare Dis 16:66. https://doi.org/10.1186/s13023-021-01692-w
43. Liu Y, Guo Y, Liu P, Li F, Yang C, Song J, Hu J, Xin D, Chen Z (2018) A case of Pitt-Hopkins syndrome with de novo mutation in TCF4: clinical features and treatment for epilepsy. Int J Dev Neurosci 67:51–54. https://doi.org/10.1016/j.ijdevneu.2018.03.010