Journal of Theoretical and Computational Chemistry
Vol. 19, No. 6 (2020) 2041003 (17 pages)

DOI: 10.1142/S0219633620410035

World Scientific
www.worldscientific.com

Computational analysis of hereditary spastic paraplegia mutations in the kinesin motor domains of KIF1A and KIF5A

Vidhyanand Mahase*, Adebiyi Sobitan*, Christina Johnson*, Farion Cooper II*, Yixin Xie†, Lin Li†,‡ and Shaolei Teng*,§

*Department of Biology, Howard University
Washington, D.C., 20059 USA

†Computational Science Program
University of Texas at El Paso
El Paso, Texas 79902, USA

‡Department of Physics
University of Texas at El Paso
El Paso, Texas 79902, USA
§shaolei.teng@howard.edu

Received 17 February 2020
Accepted 14 April 2020
Published 5 August 2020

Hereditary spastic paraplegias (HSPs) are a genetically heterogeneous collection of neurodegenerative disorders categorized by progressive lower-limb spasticity and frailty. The complex HSP forms are characterized by various neurological features including progressive spastic weakness, urinary sphincter dysfunction, extra pyramidal signs and intellectual disability (ID). The kinesin superfamily proteins (KIFs) are microtubule-dependent molecular motors involved in intracellular transport. Kinesins directionally transport membrane vesicles, protein complexes, and mRNAs along neurites, thus playing important roles in neuronal development and function. Recent genetic studies have identified kinesin mutations in patients with HSPs. In this study, we used the computational approaches to investigate the 40 missense mutations associated with HSP and ID in *KIF1A* and *KIF5A*. We performed homology modeling to construct the structures of kinesin–microtubule binding domain and kinesin–tubulin complex. We applied structure-based energy calculation methods to determine the effects of missense mutations on protein stability and protein–protein interaction. The results revealed that the most of disease-causing mutations could change the folding free energy of kinesin motor domain and the binding free energy of kinesin–tubulin complex. We found that E253K associated with ID in KIF1A decrease the protein stability of kinesin motor domains. We showed that the HSP mutations located in kinesin–tubulin complex interface, such as K253N and R280C in KIF5A, can destabilize the kinesin–tubulin complex. The computational analysis provides useful information for understanding the roles of kinesin mutations in the development of ID and HSPs.

Keywords: Kinesin mutation; kinesin–tubulin complex; intellectual disability; protein stability; protein–protein interaction.

§Corresponding author.
Vidhyanand Mahase and Adebiyi Sobitan contributed equally to this work..

1. Introduction

Hereditary spastic paraplegia (HSP) describes an assorted group of inherited neurodegenerative or neurodevelopmental disorders in which the prominent neurological symptoms and signs are spasticity and weakness in the lower extremities. HSP is a multiplex with more than 70 genetic subtypes.[1] It is involved in all patterns of hereditary inheritance and non-Mendelian parental conduction. The most shared pathological feature is a retrograde distal axonopathy of the elongated descending motor fibers of the corticospinal tract and posterior columns caused by molecular mechanisms that affect membrane vesicular trafficking, organelle morphogenesis and distribution, axonal transport, lipid metabolism, mitochondrial functions, and different steps of the myelination process.[1,2] Complex forms of HSP comprises of intellectual disability (ID), epilepsy, and extrapyramidal features such as parkinsonism, chorea, dystonia.[2] ID is the most frequent cause of severe disability and a leading socioeconomic healthcare problem in western countries.[3] Natural forms of the disease are usually autosomal dominant (AD-HSP).[4] Autosomal recessive HSPs (AR-HSPs) have a unique clinical paragon, including peripheral neuropathy, movement disorders, cognitive dysfunction and an assortment of other features.[3]

Recent studies revealed that nonsynonymous Single Nucleotide Polymorphisms (nsSNPs) in kinesins are associated with the neurological disorders including HSP and ID.[5] Kinesins are molecular motor proteins that carry cargo along microtubule tracks, thus playing important roles in intracellular transport and cell division.[6] Kinesins use the chemical energy of ATP to generate motile force to directional move various cargos, including mRNAs, protein complexes and membrane organelles, along microtubules.[7] Kinesins are responsible for the transport in the neuronal axons and dendrites. They control different sorts of transports in synaptic proteins, Golgi–plasma membrane, Golgi–ER and lysosome, and play important functions for the elongation of neurites and polarization of neurons. For example, KIF1A is known to transfer cargo toward the positive ends of the microtubules and translate axons in the transport pathway. Disruption of kinesin-related axonal transport may cause vesicle-trafficking problems, alter specific packages interactions and create flaws in retrograde survival signs, essentially leading to neuronal death and loss of human brain function.[8] Kinesins consist of a conserved motor domain (MD), a stalk domain and a cargo-binding region.[9] The motor domain hydrolyzes ATP and reversibly binds to microtubules to generate directed motion.[10] Each kinesin member of the family uses ATP hydrolysis to execute various functions, using both its own electrical motor domains to produce distinct properties. Based on the location of motor domain, kinesin family proteins are comprised of three major groups: N-terminal motor domain KIFs (N-KIFs), middle motor domain KIFs (M-KIFs), and C-terminal motor domain KIFs (C-KIFs).[9] The different kinesins are distinguished by the mechanisms that they use to recognize and bind to their specific cargos. KIF1A and KIF5A are N-KIF motors composed of an N-terminal motor domain followed by a neck coil–coil domain.[11] Different from KIF5A, KIF1A has a pleckstrin homology

(PH) domain in its C-terminal. Kinesins can regulate many broader biological processes such as higher brain function, tumor suppression and developmental patterning.

Numerous experiments have demonstrated that KIF1A and KIF5A perform critical roles in controlling spindle association, finalizing cytokinesis chromosome alignment, and completing chromosome condensation.[12] Mutations in the KIF1A and KIF5A genes were described in two different clinically and genetically heterogeneous groups of neurodegenerative diseases.[13,14] The mutations in kinesin motor domain can cause abnormal accumulations of target proteins in the nervous system and disrupt the transport of cargoes through axons.[5] For example, the genetic screening of KIF5A in patients with neurodegenerative disorders identified that the missense mutations in the motor and stalk domains motors are implicated in axonal Charcot-Marie-Tooth type 2 and spastic paraplegia type 10.[15] Understanding the roles of kinesin mutations can provide important information for identifying the targets causes neurological disorders.

Crystallographic investigation showed that kinesin motor domain has a single α/β arrowhead-shaped structure, which shares high similarity with the catalytic core of the motor protein myosin.[16] The alanine-scanning mutagenesis revealed that the kinesin residues that alter microtubule binding are located in three loops that cluster together in a patch of the motor domain.[17] The improvement of protein structure modeling makes it possible to predict the effects of kinesin mutations by mapping them on the corresponding kinesin models. The structure-based approaches can be used to quantitatively assess the effects of disease-causing mutations on protein stability and protein–protein interaction.[18,19] In a recent work, we investigated the electrostatic force changes upon kinesin mutations.[20] We showed that the mutations in motor domains can significantly affect the electrostatic component of the binding force and the force changes can be used to discriminate the damaging mutations and neutral variations. In this study, we used the computational approaches to investigate the kinesin mutations associated with neurological disorders including HSP and ID. We applied structure-based energy calculation methods to predict the effects of missense mutations on stability of kinesin motors and kinesin–microtubule binding affinity. The analytical results provide useful information for characterizing the kinesin missense mutations and elucidating the genetic mechanisms by which the coding mutations cause neurological disorders.

2. Methods and Computational Details

2.1. *Structure preparation*

In this study, we investigated the protein structures of human KIF1A and KIF5A. Both proteins have no available structures in Protein Data Bank (PDB).[21] However, the experimentally determined structures of mouse kinesin motor are available. The sequences of KIF1A (Entry: Q12756) and KIF5A (Entry: Q12840) were downloaded

from UniProt.[22] The pairwise sequence alignment of N-terminal motor domains (1-360 aa) of KIF1A and KIF5A was performed using EMBOSS Water.[23] SWISS-MODEL[24] was used to construct the human kinesin structures based on their homologues structures in mouse. Only the top model from multiple models provided by SWISS-MODEL was selected as the corresponding structure. The structure model of human KIF1A was built based on the mouse KIF1A head–microtubule complex structure (PDB ID: 2HXH; sequence similarity = 95.9%). Crystal Structure of the mouse KIF5C motor domain (PDB ID: 3WRD; sequence similarity = 91.2%) was used as the template for KIF5A. The high sequence similarities between the target proteins and the corresponding templates suggest that the models obtained by SWISS-MODEL are very reliable. Profix[25] was used to fix the experimentally determined structures missing heavy atoms. NAMD[26] was applied for a 10,000-step minimization to each kinesin motor structure, and the CHARMM[27] force field and the Generalized Born (GB) implicit solvent model were used in the minimizations. The structure alignment of KIF1A and KIF5A motor domains was carried out using FATCAT.[28]

The structure of the human kinesin–tubulin complex is not available experimentally. However, it was shown that despite their differences, all kinesins bind to microtubule using the same binding interface.[29] Thus, we built the kinesin–tubulin complex structures by homology. The cryo-EM map of microtubule-bound kinesin-5 motor domain (PDB ID: 4AQW)[30] was used as a complex template for all kinesin–tubulin complexes. Chimera[31] was utilized to build complex structures by aligning the N-termini of each kinesin to the structure of recombinant neuronal human tubulin (PDB ID 5JCO).[32] The protein structural pictures were generated using the PyMOL molecular visualization system (http://www.pymol.org/).

2.2. *Mutation collection*

The missense mutations in the kinesin motor domains were extracted from dbNSFP v3.4.[33] To determine the effects of the mutations on protein structure, we collected all mutations that can be mapped onto the corresponding motor domain structures. Human Gene Mutation Database (HGMD)[34] and ClinVar[35] 2017 versions were used for searching for disease-causing mutations. A total of 40 damaging mutations, including 22 KIF1A mutations and 18 KIF5A mutations, were investigated in this study. Of the 22 KIF1A mutations, 14 are associated with ID,[34] 3 contribute to HSP,[34] 5 are marked as pathogenic in ClinVar.[35] All 18 KIF5A mutations are related to HSP.[34]

2.3. *Free energy changes calculations*

Two types of energy calculations, the change of the folding free energy and the change of the binding free energy, were carried out to predict the effects of kinesin mutations on protein stability of kinesin motor domain and binding affinity of kinesin–tubulin complex.

The effects of missense mutations on protein stability were assessed by calculating the change of the folding free energy. The Single Amino Acid Folding free Energy Changes (SAAFEC) method[36] was used for folding energy calculations based on unbound kinesin structures. SAAFEC uses the Molecular Mechanics Poisson–Boltzmann Surface Accessibility (MM/PBSA) approach combined with knowledge-based (KB) terms of protein biophysical characteristics. The folding free energy (ΔG) is computed as

$$\Delta G_{(\text{folding})} = G_{(\text{folded})} - G_{(\text{unfolded})}. \tag{1}$$

Thus, the folding energy difference ($\Delta\Delta G$) between the structure with the mutation (MUT) and wild-type (WT) structure is calculated using

$$\Delta\Delta G_{(\text{stability})} = \Delta G_{(\text{folding})}\text{MUT} - \Delta G_{(\text{folding})}\text{WT}, \tag{2}$$

$$\Delta\Delta G_{(\text{stability})} = [G_{(\text{folded})}\text{MUT} - G_{(\text{unfolded})}\text{MUT}] - [G_{(\text{folded})}\text{WT} - G_{(\text{unfolded})}\text{WT}]. \tag{3}$$

For free folding energy calculation, the unfolded protein is considered as two parts including a three-residues segment with the mutation in the center and all other residues.[37] Thus, the folding energy change can be computed as

$$\Delta\Delta G_{(\text{stability})} = [G_{(\text{folded})}\text{MUT} - G_3\text{MUT}] - [G_{(\text{folded})}\text{WT} - G_3\text{WT}]. \tag{4}$$

A positive value of $\Delta\Delta G$ suggests that the mutation can destabilize the monomer protein and a positive value of $\Delta\Delta G$ indicates that it makes the protein more stable. The change of the folding free energy caused by amino acids mutations is an indicator of how mutations affect the stability of the corresponding kinesin. Any deviation of the protein stability from wild type may be affecting the function of kinesins.[20]

The binding energy difference between the mutant and wild-type structures was used to estimate the effect of the kinesin mutations on protein–microtubule interactions. The Single Amino Acid Mutation change of Binding Energy (SAAMBE) method[38] was used to compute the change of the binding free energy of kinesins and tubulins. The binding free energy was the difference between the energy of the complex (AB) and the individual monomers (A and B)

$$\Delta\Delta G_{(\text{binding})} = \Delta G_{(\text{folding: AB})} - \Delta G_{(\text{folding: A})} - \Delta G_{(\text{folding: B})}. \tag{5}$$

Changes in the binding energy ($\Delta\Delta\Delta$G) between the mutant and wild-type kinesins are calculated by

$$\Delta\Delta\Delta G_{(\text{binding})} = \Delta\Delta G_{(\text{binding})}\text{MUT} - \Delta\Delta G_{(\text{binding})}\text{WT} \tag{6}$$

A positive value of $\Delta\Delta\Delta G$ indicates that the mutation weakens the binding affinity, whereas a negative value suggests that the mutation strengthens the binding affinity. Any alteration of the binding affinity is expected to affect kinesin–tubulin interaction.

We utilized mCSM-PPI2[39] Predictions to identify the mutations located in interface of kinesin–tubulin complex. The Distance to Interface was used to classify the mutations as two groups: interface (Distance < 12 Å) and noninterface (Distance > 12 Å).

2.4. *Mutation pathogenicity and conservation*

We applied Rare Exome Variant Ensemble Learner (REVEL)[40] to predict the damaging effects of missense mutation on kinesin function. REVEL applies random forest on the data sets of rare variants and incorporates 18 damaging prediction scores to construct the prediction models. REVEL shows better performance for analysis of mutation pathogenicity compared with other ensemble tools such as CADD[41] and DANN.[42] We utilized the PhyloP[43] to analyze the evolutionary conservation of kinesins residues. PhyloP[43] computes evolutionary conservation at distinct alignment sites. The scores are compared to the evolution expected under neutral drift. Positive scores reflect slower evolution than expected, and such residues that are predicted to be conserved. Negative scores suggest faster evolution than expected, and such sites that are predicted to be nonconserved. PhyloP scores are useful to estimate signs of selection at particular nucleotides or classes of nucleotides.

3. Results and Discussion

3.1. *KIF1A and KIF5A motor domains*

In this work, we applied homology modeling methods to create motor domain structures of KIF1A and KIF5A. Both kinesin proteins are N-terminal motor domain KIFs (Fig. 1(a)). The sequence alignment in Fig. 1(b) suggests that their motor domains have a 56.9% similarity. All of 40 damaging mutations are in the conserved region and 22 mutations (55%) have identical residues in the sequence alignment (Fig. 1(b)). Since the kinesins have similar microtubule binding sites,[29] the human kinesin–tubulin complex structures were built by aligning each kinesin to the structures available in other species. The mutant structures of kinesin motor domains and kinesin–tubulin were built by introducing the corresponding mutations to wild-type structures. The kinesin–tubulin complex structures of KIF1A and KIF5A are shown as examples in Fig. 1(c). The kinesin motor domain has a positively charged binding interface, so it can form a binding funnel with the tubulin dimer chains which have a negatively charged binding interface. The structure alignment shows that the structures of KIF1A and KIF5A motor domains are significantly similar ($P = 0.00e + 00$; raw score $= 691.73$) and have 315 equivalent positions with an RMSD of 2.74 (Fig. 1(d)).

3.2. *Effects of mutations on protein stability*

We computed the changes of the folding free energy ($\Delta\Delta G$) and changes of binding free energy ($\Delta\Delta\Delta G$) introduced by 40 pathogenic mutations in the data set.

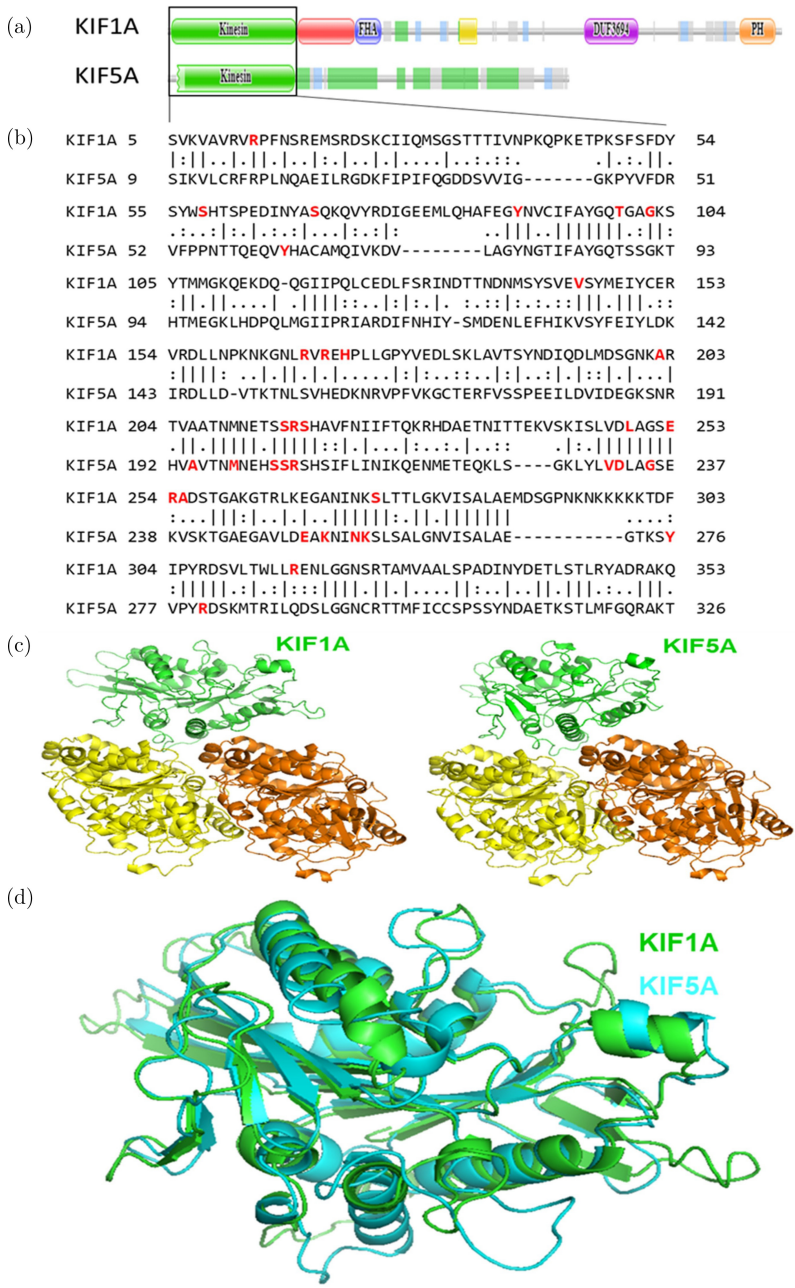

Fig. 1. (Color online) KIF1A and KIF5A motor domains. (a) Schematic functional domains of KI1A and KIF5A from Pfam.[44] (b) The pairwise sequence alignment of N-terminal motor domains (1–360 aa) of KIF1A and KIF5A from EMBOSS Water.[23] The positions of disease-causing mutations are colored with red. (c) The overall structure of KIF1A and KIF5A motor domains (green) with a tubulin dimer (A chain: orange; I chain: yellow). (d) The structure alignment of KIF1A and KIF5A motor domains from FATCAT[28] (KIF1A: green; KIF5A: cyan).

The energy changes of 22 KIF1A and 18 KIF5A disease-causing mutations were shown in Table 1. The $\Delta\Delta G$ was calculated to access the effects of kinesin mutations on motor stability. Of 40 KIF1A and KIF5A mutations, 30 (75%) have positive $\Delta\Delta G$ values and the remaining 10 (25%) have negative values, suggesting the most disease-causing mutations can decrease protein stability. 34 of 40 (85%) pathogenic mutations have large effects ($|\Delta\Delta G| > 0.5$ kcal/mol) on protein stability. The largest folding energy change ($\Delta\Delta G = 12.948$ kcal/mol) is introduced by E253K associated with ID in *KIF1A*. The HSP mutation E251K related to HSP in *KIF5A* has the second largest $\Delta\Delta G$ value at 7.924 kcal/mol. The results implicate that both mutations can significantly destabilize the kinesin motor domains. In contrast, G102D implicated in ID in *KIF1A* mutation has the largest negative folding free energy change at -4.498 kcal/mol. The HSP mutations located in R280 of *KIF5A* cause the significant folding energy changes (R280H: $\Delta\Delta G = -3.136$ kcal/mol; R280C: $\Delta\Delta G = -2.732$ kcal/mol; R280L: $\Delta\Delta G = -2.133$ kcal/mol). The negative $\Delta\Delta G$ values suggest that these mutations may stabilize the kinesin monomer structure.

3.3. *Effects of mutations on protein–protein interaction*

The $\Delta\Delta\Delta G$ was computed to estimate the effects of kinesin mutations on kinesin–tubulin Interaction. Of 40 pathogenic mutations, 36 (90%) have $\Delta\Delta\Delta G > 0$ and remaining 4 (10%) have negative $\Delta\Delta\Delta G$ values, indicating that most of disease-causing mutations can decrease kinesin–tubulin binding affinity. 26 of 40 (65%) pathogenic mutations have large effects ($|\Delta\Delta\Delta G| > 0.5$ kcal/mol) on protein–protein interaction stability. The largest binding energy change ($\Delta\Delta\Delta G = 2.194$ kcal/mol) is introduced by A255V in KIF1A. The R280C in KIF5A has the second largest $\Delta\Delta\Delta G$ value at 2.14 kcal/mol, and $\Delta\Delta\Delta G$ value caused by R280H located in the same position is 1.416 kcal/mol. These three mutations are related to HSP. The large positive $\Delta\Delta\Delta G$ values indicate that these mutations can significantly reduce the binding affinity of kinesin–tubulin complex. In contrast, E253K associated with ID in KIF1A has the largest negative $\Delta\Delta\Delta G$ at -0.789 kcal/mol, suggesting that the mutation makes kinesin–tubulin complex more stable.

We identified 19 interface mutations (Distance to Interface < 12 Å) and 21 noninterface mutations (Distance to Interface > 12 Å). We observed the $|\Delta\Delta G|$ and $|\Delta\Delta\Delta G|$ medians of interface mutations are higher compared to those values of noninterface mutations (Fig. 2). Not surprisingly, the difference of $|\Delta\Delta\Delta G|$ between the interface mutations and noninterface mutations is statistically significant (P-value $= 0.0003$), which indicates that the interface mutations have large effects on binding affinity of kinesin–tubulin complex. However, the difference of effects on protein stability ($|\Delta\Delta G|$) between interface mutations and noninterface mutations is statistically indistinguishable (P-value $= 0.3363$). In addition, 23 mutations cause large changes on both folding energy ($|\Delta\Delta G| > 0.5$) and binding energy ($|\Delta\Delta\Delta G| > 0.5$). We computed Pearson correlation coefficient r between $\Delta\Delta G$ and

Table 1. Effects of kinesin pathogenic mutations in KIF1A and KIF5A.

Mutation†	Energy Change (kcal/mol)		Distance to interface (Å)			Disease*
	$\Delta\Delta G$	$\Delta\Delta\Delta G$		REVEL	phyloP	
KIF1A						
R13H	0.603	1.268	19.140	0.953	7.648	HSP
S58L	0.794	0.403	25.030	0.938	9.750	ID
S69L	1.365	0.743	30.250	0.576	6.219	HSP
Y89D	2.519	0.778	20.880	0.954	9.083	ID
T99M	0.370	0.113	13.710	0.933	9.741	ID
G102D	−4.498	1.574	19.530	0.980	7.706	ID
V144F	−0.225	0.455	20.080	0.946	7.706	ID
R167C	1.837	0.721	7.850	0.917	5.514	ID
R169K	0.948	0.257	3.270	0.690	5.989	Pathogenic
H171P	−0.022	1.089	5.220	0.950	7.815	Pathogenic
A202P	1.730	0.777	22.000	0.860	7.706	ID
S215R	0.579	0.296	15.050	0.980	7.728	ID
R216H	1.961	0.857	10.460	0.950	7.620	ID
R216P	0.656	1.932	10.460	0.960	7.620	ID
R216C	0.517	0.884	10.460	0.960	9.628	ID
S217C	1.370	0.789	14.440	0.930	9.628	Pathogenic
L249Q	0.924	0.008	12.790	0.990	9.039	ID
E253K	12.948	−0.789	4.660	0.950	7.670	ID
R254W	2.959	0.730	2.570	0.810	3.040	Pathogenic
A255V	1.543	2.194	3.200	0.810	9.693	HSP
S274L	1.397	0.610	7.850	0.840	9.746	Pathogenic
R316W	3.954	0.577	20.370	0.910	2.704	ID
KIF5A						
Y63C	0.773	0.098	29.440	0.900	9.325	HSP
A194P	4.261	0.681	16.650	0.903	9.823	HSP
M198T	1.028	0.080	15.580	0.925	7.938	HSP
S202N	−1.104	0.366	17.390	0.846	9.900	HSP
S203C	1.096	−0.012	15.950	0.882	7.827	HSP
R204W	2.654	−0.149	12.710	0.852	0.568	HSP
R204Q	−0.075	0.996	12.710	0.924	9.900	HSP
V231L	0.015	0.446	18.650	0.900	9.814	HSP
D232N	5.392	0.053	17.830	0.902	9.814	HSP
G235E	−1.798	1.748	10.610	0.906	9.768	HSP
E251K	7.924	−0.681	10.500	0.890	10.003	HSP
K253N	−1.826	1.401	4.010	0.701	1.092	HSP
N256S	0.067	2.110	4.150	0.916	9.325	HSP
K257N	0.886	1.711	5.850	0.825	3.160	HSP
Y276C	1.257	0.287	1.390	0.882	9.197	HSP
R280C	−2.732	2.140	1.880	0.902	3.122	HSP
R280H	−3.136	1.416	1.880	0.883	9.864	HSP
R280L	−2.133	1.987	1.880	0.883	9.864	HSP

Notes: †Bold: Mutations with $|\Delta\Delta G| > 0.5$ kcal/mol and $|\Delta\Delta\Delta G| > 0.5$ kcal/mol; *Intellectual disability (ID), Hereditary spastic paraplegia (HSP), Pathogenic: ClinVar clinical significance value.

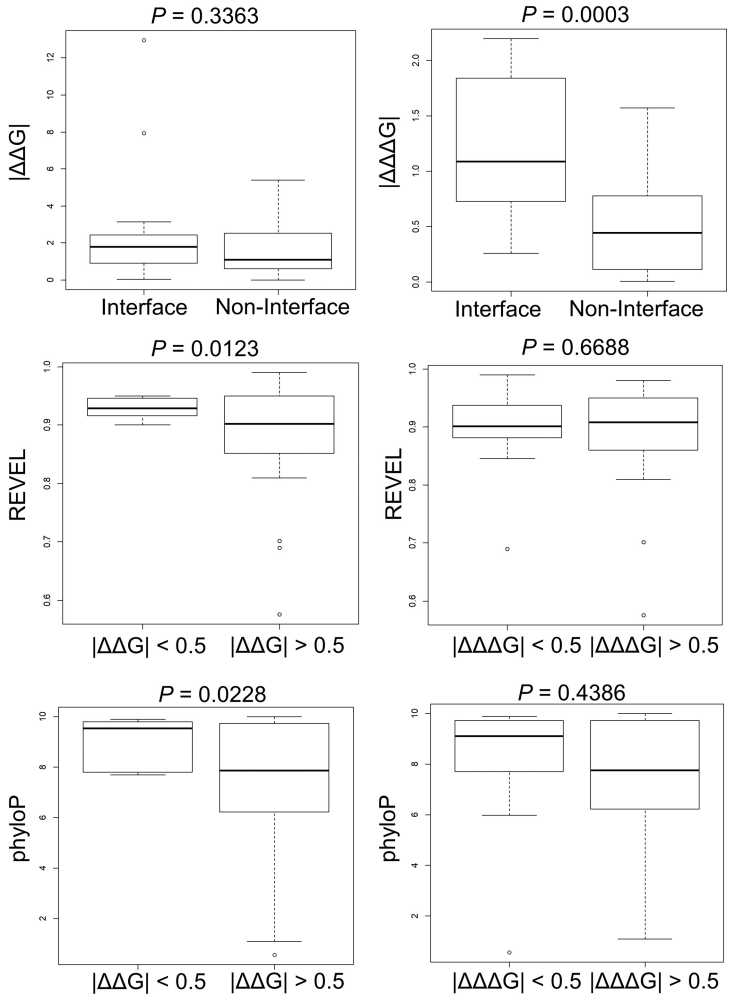

Fig. 2. Boxplots for the energy changes (kcal/mol) of kinesin pathogenic mutations.

$\Delta\Delta\Delta G$. We observed a moderate negative association ($r = -0.6307$) between $\Delta\Delta G$ and $\Delta\Delta\Delta G$ for all 40 mutations. The association is stronger in 19 interface mutations ($r = -0.7849$), which implies that the effects of kinesin mutations on motor domain stability and kinesin–tubulin binding affinity are negatively corrected in interface region.

3.4. *Mutation pathogenicity and conservation*

REVEL was applied to predict the harmful effects of kinesin missense mutations. The REVEL score can range from 0 to 1, with 1 reflecting the most likelihood of pathogenicity. 38 (95%) of 40 disease-causing kinesin mutations with REVEL score > 0.75 are predicted to have damaging effects on protein function. The REVEL

scores of mutations with large folding energy changes ($|\Delta\Delta G| > 0.5$|,kcal/mol) and binding energy changes ($|\Delta\Delta\Delta G| > 0.5$ kcal/mol) were compared with those mutations with small effects, respectively. Figure 2 shows the difference between the mutations with large effects on protein stability ($—\Delta\Delta G| > 0.5$ kcal/mol) and the mutations with $|\Delta\Delta G| < 0.5$ kcal/mol is statistically significant (P-value = 0.0123). However, we did not find a statistically significant difference for the effects of two groups ($|\Delta\Delta\Delta G| > 0.5$ and $|\Delta\Delta\Delta G| < 0.5$) on the binding affinity of kinesin–tubulin complex (P-value = 0.6688).

The PhyloP scores were used to access the conservation of kinesin mutations (Table 1). As shown in Fig. 2, the PhyloP medians of the mutations with small folding energy changes ($|\Delta\Delta G| < 0.5$ kcal/mol) and binding energy changes ($|\Delta\Delta\Delta G| < 0.5$ kcal/mol) are lower than those mutations with large energy changes. There is a statistically significant difference for the effects of variants on protein stability (P-value = 0.0228), which suggests that the KIF1A and KIF5A mutations with small effects on kinesin monomer domains are more conserved in mammalians. The findings indicate that the residues with low mutability have limited effect on protein stability, which is consistent with the previous study showing that the large shifts of mutational effects on stability are rare during protein evolution.[45] We did not observe a statistically significant result on the protein–protein interaction of kinesin–tubulin complex (P-value = 0.4386).

3.5. *Intellectual disability (ID) and KIF1A mutations*

Intellectual disability (ID) is the most common neurodevelopmental disorder characterized by cognitive delays in intellectual functioning and adaptive behavior. The mutant genes are important risk factors for the ID. In this study, we investigated 14 kinesin mutations associated with ID in *KIF1A*. KIF1A is a microtubule plus end-directed motor protein that belongs to the kinesin-3 family and is involved in the anterograde transport of synaptic vesicle precursors supported by the GTP-Rab3 adaptor.[46] Various heterozygous de novo missense mutations in *KIF1A* have been identified in patients who present intellectual disability along with atrophy and spastic paraplegia. Lee *et al.* applied exome sequencing to sequence 14 patients with ID and other complex neurologic phenotypes such as epilepsy, peripheral neuropathy and spastic paraparesis.[47] They identified 11 de novo missense mutations in the motor domain of KIF1A. They found two recurrent mutations, T99M and E253K, in unrelated cases. E253K was carried by the most severely affected patients. In this study, E253K caused the largest folding energy change ($\Delta\Delta G$) at 12.948 kcal/mol, suggesting that the mutation can significantly destabilize the KIF1A motor domain. As shown in Fig. 3(a), E253 is located in a loop and can form a back-door salt bridge. The mutation of negatively charged glutamic acid to positively charged lysine may disturb the salt bridge. The folding energy changes caused by the nearby mutations R254W and A255V are $\Delta\Delta G = 2.959$ kcal/mol and $\Delta\Delta G = 1.543$ kcal/mol, respectively. These mutations may break the back-door salt bridge and alter the local

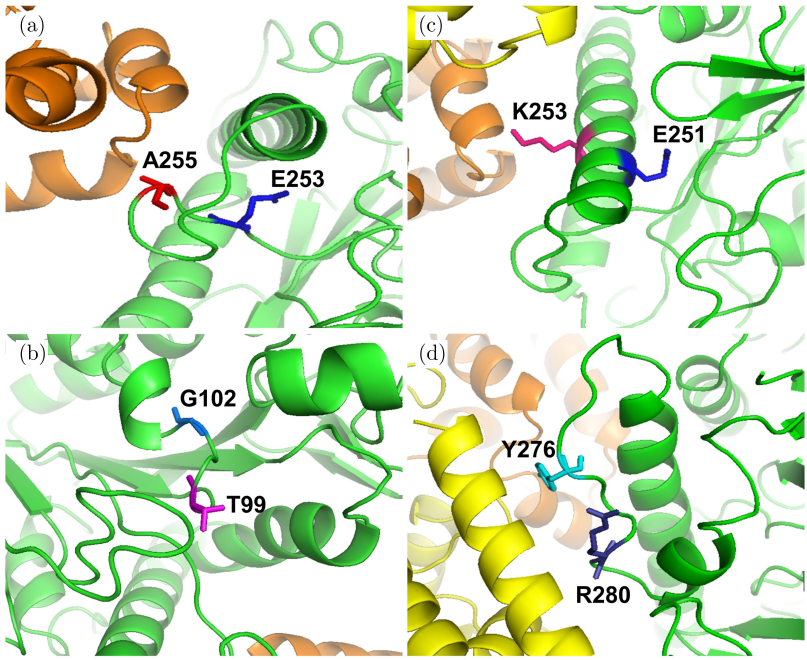

Fig. 3. (Color online) Structural representation of pathogenic mutations in KIF1A (a & b) and KIF5A (c & d) motor domains (green) with a tubulin dimer (A chain: orange; I chain: yellow). (a) Locations of E253 (blue) and A255 (red) in KIF1A. (b) Locations of T99 (pink) and G102 (marine) in KIF1A. (c) Locations of E251 (blue) and K253 (hot pink) in KIF5A. (d) Locations of Y276 (cyan) and R280 (deep blue) in KIF5A.

structure of KIF1A thereby destabilizing the motor domain. The A255V mutation is located at the interaction interface of KIF1A and tubulin (Fig. 3(a)). It introduced the largest binding energy change at $\Delta\Delta\Delta G = 2.194\,\text{kcal/mol}$, indicating that the mutation could significantly destabilize kinesin–tubulin complex. Interestingly, the binding energy change caused by E253K is $\Delta\Delta\Delta G = -0.789\,\text{kcal/mol}$. The results suggest that even nearby mutations can have different effects on kinesin–tubulin interaction. We also noted that the ID mutations in the same region have different effects on the stability of the KIF1A motor domain (Fig. 3(b)). T99M has a small effect on protein stability ($\Delta\Delta G = 0.37\,\text{kcal/mol}$). In contrast, G102D has the largest negative folding free energy change at $\Delta\Delta G = -4.498\,\text{kcal/mol}$. G102 is in a nucleotide-binding pocket which is important for ATP binding (Fig. 3(b)). The mutation of nonpolar glycine to polar aspartic acid may induce a significant conformational change for this binding pocket and disrupt the ATP binding of the motor domain.

3.6. *Hereditary spastic paraplegias and KIF5A mutations*

Hereditary spastic paraplegias (HSPs) are a genetically heterogeneous collection of neurodegenerative disorders categorized by progressive lower-limb spasticity and

frailty.[48–50] HSP is characterized by the axonal degeneration of motor and sensory neurons that is at the peak at distal ends of the longest axons of the central nervous system. In this research, we studied the effects of 21 mutations associated with HSP on kinesin structure and function. Of these 21 mutations, 3 are located in KIF1A and 18 are located in KIF5A. Recent reports suggest that the most common form of HSP, pure *SPG4,* is caused by the damaged regulation of the microtubule cytoskeleton, of which the kinesin family is involved.[50,51] Particularly, *SPG10* is an autosomal dominant form of HSP that is caused by mutations in *KIF5A*.[52] KIF5A is a neuronally-enriched isoform of the Kinesin-1 family of cellular transport motors. KIF5A is believed to be in control of intracellular vesicular movements for the processive nature of their movement.[52] The kinesin heavy chain (KHC) proteins, of which *KIF5A* is a part of, constitute a large portion of the multi-subunit complex that perform as a plus-end–directed microtubule motor implicated in anterograde transport of membranous organelles in nerve axons.[33,50,53] Currently, there are many loci at which the *KIF5A* gene can be examined for phenotypic effects in multiple forms of HSP. 23 separate mutations in the motor domain of KIF5A have been identified in patients with a complex form of HSP.[41] E251K in KIF5A has been identified in HSP patients with axonal neuropathy but no pyramidal signs, which is related to the phenotype of Charcot–Marie–Tooth disease type 2.[15] Here, E251K introduced the second largest $\Delta\Delta G$ value at 7.924 kcal/mol. As shown in Fig. 3(c), E251 resides in a helix and participates in the salt bridge confirmation. When the negatively charged glutamic acid is replaced by the positively charged lysine, the mutation may disturb the salt bridge and destabilize the motor domain. Interestingly, K253N, which lies two residues away from E251K, can stabilize the KIF5A monomer structure ($\Delta\Delta G = -1.826$ kcal/mol). The mutation is located at the interaction interface of KIF5A (Fig. 3(c)). The binding energy change caused by K253N is 1.401 kcal/mol suggesting that the mutation of positively charged lysine to uncharged asparagine acid could weaken the binding infinity of KIF5A kinesin–tubulin complex. K253N is close to the nucleotide binding domain, suggesting it may impact nucleotide binding and hydrolysis release. Three mutations, R280C, R280H, and R280L, located in R280 of KIF5A (Fig. 3(d)) have similar effects with K253N. They can stabilize the motor domain but destabilize the protein complex. In addition to the reduced kinesin–tubulin binding affinity, K253 and R280 sites have been associated with the microtubule gliding velocity and ATPase activity.[14] However, Y276C can reduce the stability of the KIF5A motor domain ($\Delta\Delta G = 1.257$ kcal/mol) and has an opposite effect compared with these three mutations in R280. The results suggest the different effects of HSP mutations in KIF5A on kinesin mechanochemistry.

4. Conclusion

In this work, we employed the homology modeling approach to construct the kinesin–microtubule binding domain and kinesin–tubulin complex structures of KIF1A

and KIF5A. The structure-based energy calculation methods were applied to determine the effects of kinesin missense mutations on protein stability and protein–protein interaction. We observed most of disease-causing mutations (75%) can decrease the stability of kinesin motor domain and 85% pathogenic mutations have large effects ($|\Delta\Delta G| > 0.5$ kcal/mol) on protein stability. The binding energy calculation shows that 90% mutations have $\Delta\Delta\Delta G > 0$ and 65% mutations have large effects ($|\Delta\Delta\Delta G| > 0.5$ kcal/mol) which indicate that the majority of deleterious mutations can decrease kinesin–tubulin binding affinity. We compared the damaging and conservation scores of kinesin mutations with large energy changes with those mutations with small effects. The results suggest that the mutations with large effects on the binding affinity of kinesin–tubulin complex could cause significant damaging effects on kinesin function (P-value $= 0.0486$). We found the kinesin mutations with small effects on the stability of kinesin motor domain are more conserved in mammalians (P-value $= 0.0228$). The results revealed the mutations associated with neurological disorders could alter folding free energy. For example, ID-related mutation E253K in KIF1A caused the $\Delta\Delta G$ at 12.948 kcal/mol, suggesting that the mutation can significantly destabilize the protein stability of kinesin–microtubule binding domains. The harmful mutations associated with HSP, such as K253N and R280C in KIF5A, located in complex interface can reduce the binding affinity of kinesin–tubulin complex. The results provide useful evidence for characterizing the kinesin mutations implicated in ID and HSP and elucidating the genetic mechanisms by which the point mutations cause these neurological disorders.

Acknowledgments

We thank Professor Emil Alexov at Clemson University for supports and critical comments. We acknowledge the Howard University Junior Faculty Writing & Creative Works Summer Academy. This work was supported by the Howard University startup funds (U100193) and National Science Foundation DBI 2000296 and IIS 1924092. This project was supported (in part) by the National Institute on Minority Health and Health Disparities of the National Institutes of Health under Award Number 2U54MD007597. This research was also supported by Grant SC1GM132043-01 from National Institutes of Health; Grant 5U54MD007592 from the National Institutes on Minority Health and Health Disparities (NIMHD), a component of the NIH. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. First two authors contributed equally to this work.

References

1. Guidubaldi A *et al.*, Novel mutations in SPG11 cause hereditary spastic paraplegia associated with early-onset levodopa-responsive Parkinsonism, *Mov Disord* **26**:553–556, 2011.

2. de Souza PVS, de Rezende Pinto WBV, de Rezende Batistella GN, Bortholin T, Oliveira ASB, Hereditary spastic paraplegia: Clinical and genetic hallmarks, *Cerebellum* **16**:525–551, 2017.
3. Abou Jamra R *et al.*, Adaptor protein complex 4 deficiency causes severe autosomal-recessive intellectual disability, progressive spastic paraplegia, shy character, and short stature, *Am J Hum Genet* **88**:788–795, 2011.
4. Faber I, Branco LMT, França Júnior MC, Cognitive dysfunction in hereditary spastic paraplegias and other motor neuron disorders, *Dement Neuropsychol* **10**:276–279, 2016.
5. De Vos KJ, Grierson AJ, Ackerley S, Miller CCJ, Role of axonal transport in neurodegenerative diseases, *Annu Rev Neurosci* **31**:151–173, 2008.
6. Hirokawa N, Takemura R, Molecular motors and mechanisms of directional transport in neurons, *Nat Rev Neurosci* **6**:201–214, 2005.
7. Hirokawa N, Noda Y, Intracellular transport and kinesin superfamily proteins, KIFs: Structure, function, and dynamics, *Physiol Rev* **88**:1089–1118, 2008.
8. Goldstein AY, Wang X, Schwarz TL, Axonal transport and the delivery of pre-synaptic components, *Curr Opin Neurobiol* **18**:495–503, 2008.
9. Miki H, Okada Y, Hirokawa N, Analysis of the kinesin superfamily: Insights into structure and function, *Trends Cell Biol* **15**:467–476, 2005.
10. Goshima G, Cell cycle-dependent dynamics and regulation of mitotic kinesins in drosophila S2 cells, *Mol Biol Cell* **16**:3896–3907, 2005.
11. Hirokawa N, Noda Y, Tanaka Y, Niwa S, Kinesin superfamily motor proteins and intracellular transport, *Nat Rev Mol Cell Biol* **10**:682–696, 2009.
12. Seeger MA, Zhang Y, Rice S E, Kinesin tail domains are intrinsically disordered, *Proteins Struct Funct Bioinforma* **80**: 2437–2446, 2012.
13. Klebe S *et al.*, KIF1A missense mutations in SPG30, an autosomal recessive spastic paraplegia: Distinct phenotypes according to the nature of the mutations, *Eur J Hum Genet* **20**:645–649, 2012.
14. Dutta M, Diehl MR, Onuchic JN, Jana B, Structural consequences of hereditary spastic paraplegia disease-related mutations in kinesin, *Proc Natl Acad Sci USA* **115**: E10822–E10829, 2018.
15. Crimella C *et al.*, Mutations in the motor and stalk domains of KIF5A in spastic paraplegia type 10 and in axonal Charcot-Marie-Tooth type 2, *Clin Genet* **82**:157–164, 2012.
16. Kull FJ, Sablin EP, Lau R, Fletterick RJ, Vale RD, Crystal structure of the kinesin motor domain reveals a structural similarity to myosin, *Nature* **380**:550–555, 1996.
17. Woehlke G *et al.*, Microtubule interaction site of the kinesin motor, *Cell* **90**:P207–P216, 1997.
18. Teng S, Srivastava AK, Schwartz CE, Alexov E, Wang L, Structural assessment of the effects of amino acid substitutions on protein stability and protein protein interaction, *Int J Comput Biol Drug Des* **3**: 334–349, 2010.
19. Teng S, Madej T, Panchenko A, Alexov E, Modeling effects of human single nucleotide polymorphisms on protein-protein interactions, *Biophys J* **96**: 2178–2188, 2009.
20. Li L *et al.*, Forces and disease: Electrostatic force differences caused by mutations in kinesin motor domains can distinguish between disease-causing and non-disease-causing mutations, *Sci Rep* **7**:8237, 2017.
21. Berman H, Henrick K, Nakamura H, Markley JL, The worldwide Protein Data Bank (wwPDB): Ensuring a single, uniform archive of PDB data, *Nucleic Acids Res* **35**:D301–D303, 2007.
22. Wu CH *et al.*, The Universal Protein Resource (UniProt): An expanding universe of protein information, *Nucleic Acids Res* **34**: D187–191, 2006.

23. Madeira F *et al.*, The EMBL-EBI search and sequence analysis tools APIs in 2019, *Nucleic Acids Res* **47**:W636–W641, 2019.
24. Biasini M *et al.*, SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information, *Nucleic Acids Res* **42**:W252–W258, 2014.
25. Xiang Z, Advances in homology protein structure modeling, *Curr Protein Pept Sci* **7**:217–227, 2006.
26. Nelson MT *et al.*, NAMD: A parallel, object-oriented molecular dynamics program, *Int J High Perform Comput Appl* **10**:251–268, 1996.
27. Vanommeslaeghe K *et al.*, CHARMM general force field: A force field for drug-like molecules compatible with the CHARMM all-atom additive biological force fields, *J Comput Chem* **31**:671–690, 2010.
28. Ye Y, Godzik A, FATCAT: A web server for flexible structure comparison and structure similarity searching, *Nucleic Acids Res* **32**:W582–W585, 2004.
29. Mizuno N *et al.*, Dynein and kinesin share an overlapping microtubule-binding site, *EMBO J* **23**:2459–2467, 2004.
30. Goulet A *et al.*, The structural basis of force generation by the mitotic motor kinesin-5, *J Biol Chem* **287**:44654–44666, 2012.
31. Pettersen EF *et al.*, UCSF Chimera — A visualization system for exploratory research and analysis, *J Comput Chem* **25**:1605–1612, 2004.
32. Vemu A *et al.*, Structure and dynamics of single-isoform recombinant neuronal human tubulin, *J Biol Chem* **291**:12907–12915, 2016.
33. Liu X, Wu C, Li C, Boerwinkle E.dbNSFP v3.0: A one-stop database of functional predictions and annotations for human nonsynonymous and splice-site SNVs, *Hum Mutat* **37**: 235–241, 2016.
34. Stenson PD *et al.*, The human gene mutation database: 2008 update, *Genome Med* **1**: 13, 2009.
35. Landrum MJ *et al.*, ClinVar: Public archive of interpretations of clinically relevant variants, *Nucleic Acids Res* **44**:D862–D868, 2016.
36. Getov I, Petukh M, Alexov E, SAAFEC: Predicting the effect of single point mutations on protein folding free energy using a knowledge-modified MM/PBSA approach, *Int J Mol Sci* **17**:512, 2016.
37. Capriotti E, Fariselli P, Casadio R., A neural-network-based method for predicting protein stability changes upon single point mutations, *Bioinformatics* **20**(Suppl 1):i63–68, 2004.
38. Petukh M, Dai L, Alexov E, SAAMBE: Webserver to predict the charge of binding free energy caused by amino acids mutations, *Int J Mol Sci* **17**:547, 2016.
39. Rodrigues CHM, Myung Y, Pires DEV, Ascher DB, mCSM-PPI2: predicting the effects of mutations on protein-protein interactions, *Nucleic Acids Res* **47**:W338–W344, 2019.
40. Ioannidis NM *et al.*, REVEL: An ensemble method for predicting the pathogenicity of rare missense variants, *Am J Hum Genet* **99**:877–885, 2016.
41. Chun S, Fay JC, Identification of deleterious mutations within three human genomes, *Genome Res* **19**:1553–1561, 2009.
42. Quang D, Chen Y, Xie X, DANN: A deep learning approach for annotating the pathogenicity of genetic variants, *Bioinformatics* **31**:761–763, 2015.
43. Wu J, Li Y, Jiang R, Integrating multiple genomic data to predict disease-causing nonsynonymous single nucleotide variants in exome sequencing studies, *PLoS Genet* **10**:e1004237, 2014.
44. El-Gebali S *et al.*, The Pfam protein families database in 2019, *Nucleic Acids Res* **47**:D427–D432, 2019.

45. Ashenberg O, Gong LI, Bloom JD, Mutational effects on stability are largely conserved during protein evolution, *Proc Natl Acad Sci USA* **110**:21071–21073, 2013.
46. Blom N, Gammeltoft S, Brunak S, Sequence and structure-based prediction of eukaryotic protein phosphorylation sites, *J Mol Biol* **294**:1351–1362, 1999.
47. Lee JR *et al.*, De novo mutations in the motor domain of KIF1A cause cognitive impairment, spastic paraparesis, axonal neuropathy, and cerebellar atrophy, *Hum Mutat* **36**:69–78, 2015.
48. Xue Y *et al.*, GPS 2.0, a tool to predict kinase-specific phosphorylation sites in hierarchy, *Mol Cell Proteomics*, **7**:1598–1608, 2008.
49. Li A, Xue Y, Jin C, Wang M, Yao X, Prediction of nepsilon-acetylation on internal lysines implemented in Bayesian discriminant method, *Biochem Biophys Res Commun* **350**:818–824, 2006.
50. Deng W *et al.*, GPS-PAIL: Prediction of lysine acetyltransferase-specific modification sites from protein sequences, *Sci Rep* **6**:39787, 2016.
51. Blom N, Sicheritz-Pontén T, Gupta R, Gammeltoft S, Brunak S, Prediction of post-translational glycosylation and phosphorylation of proteins from the amino acid sequence, *Proteomics* **4**:1633–1649, 2004.
52. Cooper DN, Ball EV, Krawczak M, The human gene mutation database, *Nucleic Acids Res* **26**:285–287, 1998.
53. Liu Z *et al.*, GPS-YNO2: computational prediction of tyrosine nitration sites in proteins, *Mol Biosyst* **7**:1197–1204, 2011.