RSC Advances

This article can be cited before page numbers have been issued, to do this please use: M. Lambrughi, M. Lucchini, M. Pignataro, M. Sola and C. A. Bortolotti, *RSC Adv.*, 2016, DOI: 10.1039/C6RA06312H.

ROYAL SOCIETY OF CHEMISTRY

www.rsc.org/advances

View Article Online
DOI: 10.1039/C6RA06312H

The dynamics of the β-propeller domain in Kelch protein KLHL40 changes upon nemaline myopathy associated mutation

Matteo Lambrughi[#], Matteo Lucchini[#], Marcello Pignataro[^], Marco Sola[#], Carlo Augusto Bortolotti[#§⊙]

[#] Department of Life Sciences, University of Modena and Reggio Emilia, Via Campi 103, 41125, Modena (Italy)

[^] Department of Chemical and Geological Sciences, University of Modena and Reggio Emilia, Via Campi 103, 41125, Modena (Italy)

[§] CNR-Nano Institute of Nanoscience, via Campi 213/A, 41125 Modena, Italy

[⊙] Corresponding author. Mail to: carloaugusto.bortolotti@unimore.it

Electronic Supplementary Information (ESI) available: Additional analyses performed on the MD sampling. RMSD, RMSF and analysis of secondary structure for both WT and E528K mutant simulations. Comparison between B-factors of X-ray structures of KLHL40 and KLHL41 and RMSF for WT KLHL40 from MD simulations. Principal Component analysis performed on the combined macro-trajectory separately on each blade.

View Article Online
DOI: 10.1039/C6RA06312H

Abstract

Evolutionarily widespread, functionally and structurally diverse and still largely unexplored, Kelch proteins, characterized by the presence of a conserved C-terminal β-propeller, are implicated in a number of diverse fundamental biological functions, including cytoskeletal arrangement, regulation of cell morphology and organization, and protein degradation. Mutations in the genes encoding for Kelch superfamily members are being discovered as the cause of several neuromuscular diseases and cancer. The E528K mutation in Kelch protein KLHL40, which regulates skeletal muscle myogenesis, has been identified as a frequent cause of severe autosomal-recessive nemaline myopathy (NM). We use all-atoms molecular dynamics simulations to characterize the dynamic behaviour of the β-propeller of the wild-type protein and identify correlated motions underlying the *in vivo* functionality. We also modelled the NM-associated mutation and we found that it does not lead to dramatic disruption of the β-propeller architecture; yet, residue 528 is a hub in the correlated motions of the domain, and mutation-induced local structural alterations are propagated to the whole protein, affecting its dynamics and physico-chemical properties, which are fundamental for in vivo interaction with partners. Our results indicate that rational design of drugs can be envisioned as a strategy for restoring the proper internal network of communication and resetting KLHL40 to its physiological state.

Keywords

Molecular dynamics; propeller; nemaline myopathy; mutation; Kelch; KBTBD5.

View Article Online
DOI: 10.1039/C6RA06312H

Introduction

The Kelch superfamily encompasses a large number of proteins, displaying high diversity in terms of structure, biological role and host, but all invariably featuring a Kelch-repeat domain, which is composed by a series of consecutive Kelch motifs[1,2]. The Kelch motif can be identified at the sequence level: it ranges between 44 and 56 aminoacids in length, and it exhibits low sequence identity across the diverse genes, apart from the conserved signature motif composed by four hydrophobic residues, followed by two Gly, one Tyr and one Trp.[3] At the structure level, each Kelch motif folds into a 4-stranded antiparallel β-sheet;[1,3] the Kelch repeats arrange themselves into a β-propeller, with each motif forming one of the "blades" of the propeller itself. The Kelch superfamily is therefore hallmarked by the presence of the β-propeller tertiary structure, but can be divided into three different subfamilies, namely KLHL, KBTBD and KLHDC, depending on the presence of different additional domains and on their organization into the mature protein structure:[1] in KLHL members, the C-terminal Kelch repeats (their number typically ranging from five to seven) are preceded by an N-terminal BACK domain and a BTB one,[4,5] which is instead absent in KBTBD. KLDHC subfamily members usually lack both BTB and BACK: in some cases, they only feature the Kelch propeller, whereas in other cases they might also be formed by other domains.

The Kelch superfamily in humans encompasses 63 proteins.[1] Although the molecular details of the role of each of these proteins are still completely lacking or incomplete for most of them, a common role of Kelch proteins as substrate adaptors in the ubiquitination process is now recognized.[6–10] The ubiquitin-proteasome system is responsible for the control of protein turnover: Kelch proteins play a crucial role in it, by organizing and adapting the multimolecular complex between the enzyme Cullin3-dependent E3

View Article Online
DOI: 10.1039/C6RA06312H

ubiquitination ligase and the substrate that needs to be ubiquitinated, and then degraded by the proteasome machinery.[11,12] In particular, the β-propeller serves as a substrate recognition domain, with each Kelch repeat domain binding a wide and diverse range of substrates.

Kelch proteins have been gaining increasing attention in the last few years, with a growing body of evidence indicating the emerging role in the development (and, in case of mutations or malfunctioning, to diseases of) the skeletal muscle.[1] One of the pathologies that have been associated with mutations in Kelch proteins is nemaline myopathy (NM), a rare form of congenital myopathy, characterized by muscle weakness and nemaline bodies in affected myofibers.[13,14] Many clinical studies have shown that the onset of this disorder can occur in infancy, childhood or adulthood, and that family history is consistent with an autosomal recessive or autosomal dominant inheritance of the disease, although many affected individuals represent single cases attributable to a *de novo* dominant mutation.[15,16] While the details of the pathogenesis still represent an open issue, to date mutations in nine genes have been associated with NM: six of them code for thin filament proteins (the most common disease-causing mutations having been found in the genes NEB[17] and ACTA1,[18–20] which encode for the proteins nebulin and actin), and this finding led to the hypothesis that mutations resulting in misfolded or poorly expressed thin filament proteins would hamper the ability of the myofiber to generate adequate force during contraction.[21] The three remaining genes belong to the Kelch superfamily, namely KBTBD13,[22] KLHL40 (also classified as KBTBD5),[13,23,24] and KLHL41,[25] and the mutations, almost invariably occurring in the β-propeller domain, were demonstrated to cause severe autosomal-recessive forms of NM.[13] It was reasonable to suggest that the three Kelch superfamily proteins would be involved into the regulation of thin filament proteins: interestingly, at variance with other Kelch proteins, involved in substrate

View Article Online
DOI: 10.1039/C6RA06312H

ubiquitination and degradation, KLHL40, a regulator of skeletal muscle myogenesis,[26] was recently suggested to have an opposite role, promoting stability of its physiological partners nebulin and leiomodin3.[23]

The lack of structural details on Kelch proteins and the number and diversity of potential physiological partners have limited the identification of molecular determinants either common to the whole superfamily or specific to each species. Here, we focus our attention on the β-propeller domain of the human KLHL40 member of the Kelch superfamily: using the crystal structure[7] as a starting point for our molecular modelling, we performed 500 nanoseconds-long, all-atom explicit solvent molecular dynamics (MD) simulations, to investigate the dynamic behaviour of the protein, which is known to determine the in vivo functionality of biomolecules.[27–32] We then constructed *in silico* the E528K mutant, which is associated with the most widely occurring NM-causing mutation in KLHL40, and investigated it by means of MD, comparing it with the wild-type (WT) species.

To the best of our knowledge, this is one of the first extensive investigations of the dynamics/function relationship in human Kelch proteins. Unveiling the conformational repertoire sampled by WT KLHL40 and identifying a network of communication between its different regions represented a crucial step for shedding light on the molecular mechanisms underlying Kelch proteins function. Moreover, it served as the reference system to understand how the disease causing mutation drives structural and dynamic changes in the system, in terms of alteration of physiologically relevant concerted motions, optimal communication pathways and modification of physico-chemical properties that are expected to drive the interaction with partners *in vivo*.

Methods

View Article Online
DOI: 10.1039/C6RA06312H

Simulations

The molecular dynamics simulations were performed using the structure of the Kelch domain of Human KLHL40 (PDB ID 4ASC) and introducing the E528K mutation *in silico*. MD simulations were performed using the 4.6.5 version of the GROMACS software (www.gromacs.org), implemented on a parallel architecture, using the force field Charmm22[33,34] with CMAP correction[35] was used. We performed all-atom, explicit-solvent, MD simulations, employing periodic boundary conditions and each simulation was conducted for 0.5 μs. The protein systems were soaked in a dodecahedral box of TIP3P water molecules[36] with all the protein atoms at a distance equal or greater than 1 nm from the box edges and charge-neutralizing counter ions (Na^+ and Cl^-) were added. Productive MD simulations were performed in the NVT ensemble at 300K using the Berendsen temperature coupling[37]. We employed full particle-mesh Ewald electrostatics[38] and the Van der Waals and Coulomb non-bonded interactions were truncated at 10 Å. We used periodic boundary conditions and a 2fs time step using the LINCS algorithm[39]. The systems were initially minimized using steepest descent algorithm over 10000 steps. The optimization step was followed by 50 ps of solvent equilibration at 300K, while restraining the protein atomic positions using a harmonic potential. Then the systems were annealed to 300 K, starting from 50 K, in a 100 ps run. The electrostatic properties of WT and E528K mutant were investigated using the APBS software.[40] The .pqr input file required to run APBS was prepared using PDB 2PQR5.[41,42] The electrostatic potential was obtained by solving the linearized Poisson–Boltzmann equation (LPBE) at 298.5 K, using dielectric constant values of 2 and 78.54 for protein (solute) and solvent, respectively.

TimeScapes analysis

We used TimeScapes package, Version 1.3[43] to identify conformational changes and alterations in WT and mutant Kelch domain. To identifying time-dependent contact graphs

View Article Online
DOI: 10.1039/C6RA06312H

from distance matrices, TimeScapes employs a contact metric based on higher- order generalizations of Delaunay tetrahedralization, techniques (GMD graphs) and a contact metric based on distance cutoffs (Cutoff graphs). For the TimeScapes analysis, we used the same residue subset used for the PCA, considering all protein atoms. The contact metric used to separate contacts/noncontacts between side chains of two aminoacids and define the buffer region were set to 7.0 Å (contacts) and 8.0 Å (buffer) for the Cutoff graphs and 2 (contacts) and 3 (buffer) edge order for the GMD graphs, using parameters as suggested in the literature[43]. The half width median filter was set to a value of 15 ns.

Cross correlation analysis

We performed the calculation of cross correlation considering only the Cα atoms using the Bio3D package, version 1.1–4.[44] We calculated the cross-correlation coefficient, C_{ij}, for the displacement of all Cα atom pairs, i and j, in order to identify protein regions with correlated atomic motions.

$$C_{ij} = < \Delta r_i \cdot \Delta r_j > / \ (< \Delta r_i^{\,2} >< \Delta r_j^{\,2} >)^{1/2}$$

where Δr_i is the variation from the mean position of the i_{th} atom calculated from the structure in the simulations analyzed.[45]

Intramolecular interaction analysis and structural communications

We here employed PyInteraph[46] to investigate non-covalent intramolecular interactions and communication between different regions of the proteins and post-processed the data with xPyder.[47] PyInteraph was used to evaluate the occurrence and persistence of hydrogen bonds, hydrophobic interactions and electrostatic interactions between pairs of residues during the simulations. The threshold of persistence of a single interaction between a pair of atoms during simulation time that was set to 20%, to discard non-

View Article Online
DOI: 10.1039/C6RA06312H

significant and poorly populated interactions in the ensemble.[48] The software represents the protein as a simple, weighted graph, composed by nodes, the residues, connected by edges, the interactions, whose weights are defined by the persistence of the interaction in the MD (i.e. the number of frames in which the interaction was present, over the total number of frames). Residues connected by more than 3 edges to other residues[49] were considered as hubs of the networks. The connected components of the graph were also calculated: these are sub-network in which all the nodes are linked by at least one edges, but they have no interaction with the rest of the nodes in the global graph.

Results and discussion

The diversity at both sequence and structure level that can be observed within Kelch domains has been suggested to mediate different protein-protein interaction and cofactor specificity:[7] this is why it is important to describe the structure of KLHL40 by highlighting similarities and differences with respect to other members of the same superfamily. The six blades of the KLHL40 propeller are composed each by a four-stranded antiparallel β-sheet. (Figure1) The C-terminal βA strand completes the blade I, whose twisted structure makes it peculiar, if compared with that of other Kelch proteins like KEAP1, KLHL2 or KLHL7.[7] Moreover, KLHL40 has shorter BC loops in blade I (3 residues), and longer BC loops in blade II and VI (18 and 20 residues, respectively) than other Kelch proteins. The only protein closely paralleling KLHL40 from a structural point of view is KLHL41, another Kelch superfamily member that was found mutated in NM-affected patients.

View Article Online
DOI: 10.1039/C6RA06312H

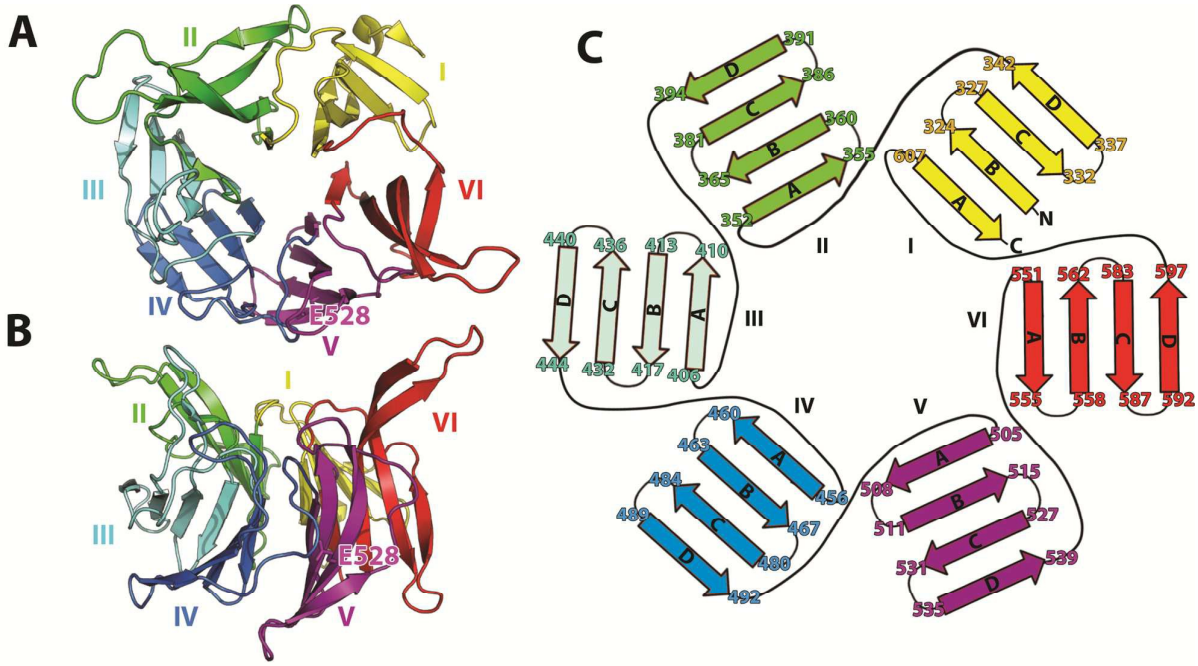


Figure 1. Structure of the human KLHL40 member of the Kelch superfamily. A) upper view and B) side view of the β-propeller, here showed as cartoon. The blades are highlighted with colors: blade I yellow, blade II green, blade III cyan, blade IV azure, blade V purple, blade VI red. The mutated residue Glu528 associated with nemaline myopathy is located in blade V. C) The structural architecture of KLHL40 Kelch domain. Each blade is composed by a 4-stranded antiparallel β-sheet, forming together the β-propeller domain; numbers of the first and last residue of each β-sheet are provided.

Fluctuations of solvent exposed BC loops dominate the dynamics of the β-propeller domain and drive conformational transitions between minima.

We first simulated WT KLHL40, using its crystal structure (PDB ID 4ASC)[7] as the starting point of our simulations. The Kelch domain of KLHL40 displays low root mean square displacement (RMSD), showing that the stability of the domain is maintained after the first 10 ns of simulation time (see Figure S1 in the Electronic Supplementary Information, ESI). Moreover, the analysis of secondary structures and their persistence shows that they are

View Article Online
DOI: 10.1039/C6RA06312H

strongly conserved during the simulation, also indicating the preservation of the β-propeller architecture (Figure S2, ESI).

We used principal component analysis (PCA) to identify the dominant collective motions at the nanoseconds timescale. The first three PCs capture more than 70% of the total of atomic positional fluctuations. This finding allows us to state that these PCs describe the majority of the fluctuations, and therefore they provide a good representation of the conformational space sampled during the whole simulation time. We projected the trajectory over the single PCs to identify the motions described by each of them. We firstly calculated PCA over all protein atoms: as expected, the terminal regions, and in particular the long and disordered C-terminal tail (residues 620-628) undergo large fluctuations, most likely unrelated to the protein functionality. We then repeated PCA without the C-terminal region, to focus on the dynamics of the rest of the Kelch domain (residues 314-620). To identify what protein regions provide the highest contribution to the protein motion with highest amplitude, we calculated the RMSF of the main chain atoms along each of the first 3 PCs. Residues displaying the highest RMSF are those that mostly contribute to each PC: they are represented with thicker radius, and coloured in red, in Figure 2A, while thinner, blue-coloured residues are those displaying the lower RMSF (and hence minimal contribution to the corresponding PC). The analysis points out that the first PCs are dominated by motions that principally involve residues in the surface exposed loops BC of blade VI, II and III (Figure 2A).

View Article Online
DOI: 10.1039/C6RA06312H

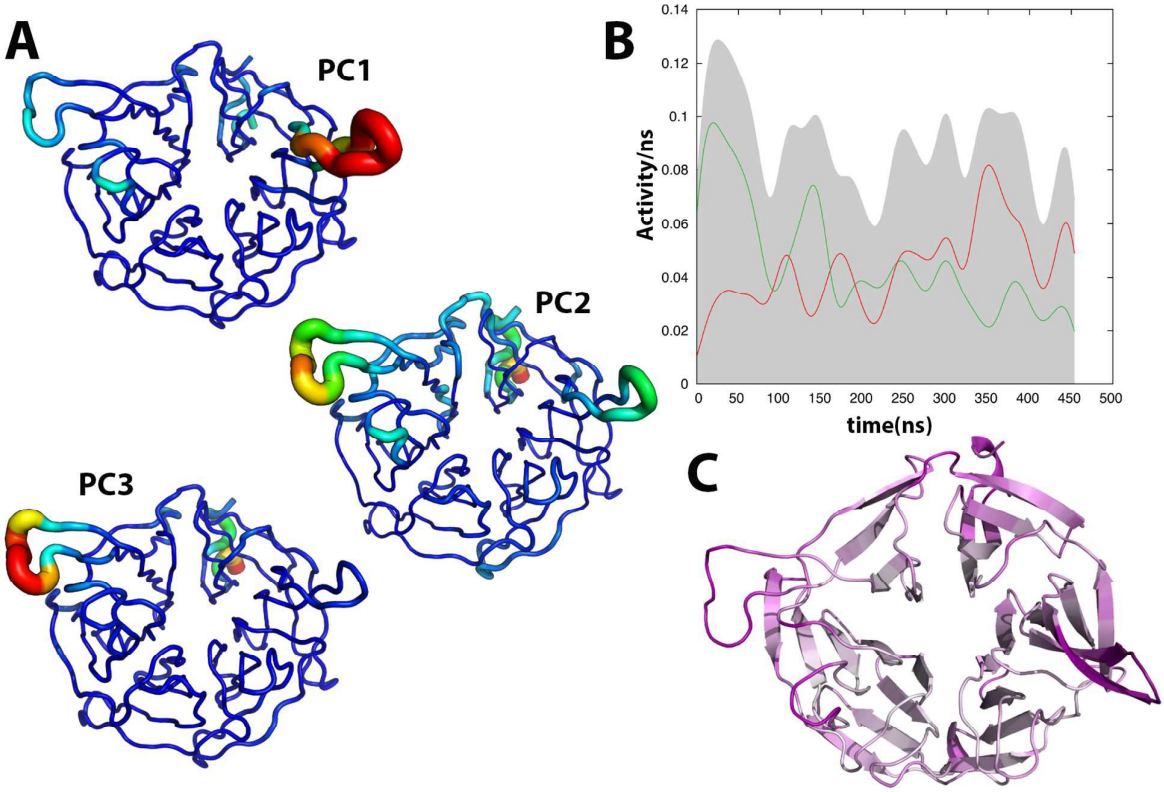


Figure 2. Dominant motions in the Kelch domain of KLHL40. A) Principal component analysis (PCA) to identify principal variable regions in the system. We calculated the projections of the simulations on each of the three first PCs and calculated the RMSF to identify the contribution of residues to the motions described. The values are represented as shade of colors, from blue (low RMSF values – high contribution to PC- poorly flexible regions) to red (high RMSF values – high contribution to PC – very flexible regions). B) TimeScapes analysis of time evolution of dominant motions. The grey profile show the global activity (i.e. broken and formed interactions) red and green profiles describe the activity associated with broken and formed contacts, respectively. C) The regions of the Kelch domain that provide major contributions to the protein activity measured by TimeScapes are highlighted with shade of color from white (small contribution) to purple (high contribution).

Moreover, other lower amplitude motions involve BC loops of blades IV and V, CD loop of blade II and CD loop of blade VI (Figure 2A). The rest of the core domain is less flexible, and we hypothesize that it is important to maintain the stability of the system. To gain an

View Article Online
DOI: 10.1039/C6RA06312H

additional characterization of the flexibility of the β-propeller, we calculated the RMSF along the whole trajectory. The RMSF profile (Figure S3, ESI) shows that the central β-sheets and loops in the core region of the domain have relatively low value of RMSF (mean values around 1-0.9 Å) further proving that the domain maintains its stability in the sampled conformational space. The RMSF analysis also supports the PCA results, i.e. that the surface exposed loops in the upper and side faces of the domain are the most flexible regions, with the largest motions taking place in the regions that provided the highest contributions to the first three PCs. Besides being indicative of the protein sites with higher flexibility, RMSF values can be used to assess the reliability of the sampling obtained via molecular modelling simulations, by comparison with B-factors extracted from the corresponding X-ray structure. A good agreement can be observed between average fluctuations calculated for the simulated MD ensemble and experimentally measured deviations, with a R^2=0.71. Moreover, also the comparison between the B-factor values of the Kelch domain of structurally close KLHL41 and fluctuations calculated from KLHL40 simulations show a similar behaviour. In the crystal structure of KLHL41 some loops are missing, suggesting that they are particularly flexible and can assume different conformations; these missing loops correspond to the region of the two long BC loops that have high atomic fluctuations in MD simulations.

The overall convergence between the evidences provided by PCA, RMSF and experimental data suggests that our MD conformational sampling is accurate and reliable in identifying the regions involved in atomic fluctuations of highest amplitude, thus providing a reasonable description of the conformational landscape accessible to the protein in solution.

Once the stability of our simulation was assessed, we turned our attention to the identification of the time evolution of the global motions of the propeller, searching for

View Article Online
DOI: 10.1039/C6RA06312H

stable conformations and structural transitions that might be physiologically relevant. This was done using the TimeScapes method,[43] which performs a time-dependent monitoring of broken and formed contacts between residues in MD trajectories: the total number of changes in contacts over time (defined as the "activity" of the system) is used as an indicator to identify and describe significant conformational events in proteins. The analysis of the activity profiles during simulation time provided by TimeScapes allows to identify basins of minimal activity, corresponding to putative stable states of the protein, and to cluster the trajectory in segments that have similar conformational states and patterns of intramolecular contacts between residues. We identified 9 segments during the evolution of the WT simulation (Figure 2B); the structure of the minima of each conformation segment have been identified and used as reference structures for further analysis. The protein portions displaying the highest number of broken and formed contacts, and therefore providing the major contributions to the transition events and to the protein activity, are the two loops BC of blades II and VI, and, to a lesser extent, from loops DA in blade I, BC in blade III and CD loops in blades V and VI (Figure 2C). TimeScapes analysis therefore complements PCA and RMSF-based conclusions, confirming the high mobility of the BC loops. Most of the solvent exposed, flexible regions displaying higher fluctuations are located in a protein region, whose corresponding site on Keap1 was assigned a key role in the interaction with the partner.[50] Keap1 is one of the few Kelch protein for which structure/function data are available.[51] In KLHL40 (and, most likely, in other β-propellers) BC loops fluctuations dominate the dynamics: they are likely to be crucial in defining the specific interaction with physiological targets, and changes in their structure or dynamics might severely alter the proper KLHL40 functionality, by modifying the interaction with physiological targets and thus interfering with the tight regulation of the protein degradation pathways.

View Article Online
DOI: 10.1039/C6RA06312H

The NM-associated E528K mutation does not lead to dramatic changes in the overall architecture of the β-propeller.

We then performed molecular dynamics simulations of the E528K mutant of the Kelch domain of KLHL40 and compared its behaviour with that of the WT form. As mentioned before, E528K is by far the most widely occurring NM-associated mutation in KLHL40. Glu528 is located on the C β-strand of blade V (Figure 1A); it is partially exposed to the solvent and only marginally involved in the region that seems to be crucial to the maintenance of the hydrophobic core of Kelch domain. We introduced the mutation *in silico* in the WT X-ray structure of Kelch domain of KLHL40 and collected overall 500 ns of explicit solvent full-atom MD simulations.

We initially calculated the residue dependent RMSF profile of the mutant along the simulation and compared it to that of the WT (Figure S3, ESI). The two profiles are very similar, a first evidence that the mutation does not dramatically perturb the overall stability and the dynamic signature of the Kelch domain. Further confirmation came from monitoring RMSD, radius of gyration, secondary structure content and persistence of structural elements during MD simulations: their values are similar for both mutant and WT (Figures S1 to S3, ESI), suggesting that, despite the mutation, the overall architecture of the β-propeller is maintained. Although the sampling methods and length of our simulations would not allow for observing complete unfolding events, we do not detect any structural change that would hint that the mutation is associated with events of unfolding, nor that extending further our sampling would allow for observation of severe loss of secondary and tertiary structure.

The NM-associated E528K mutation has long-range effects on the dynamics of solvent exposed BC and DA loops of the whole propeller.

View Article Online
DOI: 10.1039/C6RA06312H

We used the TimeScapes method, using the same approach applied for the WT species, to investigate the time evolution of global motions of the mutant and identify major amplitude conformational transitions (Figure 3A). We identified 5 segments during the time evolution of the simulation, separated by events of high activity, i.e. corresponding to significant rearrangements in the intramolecular contacts. Similarly to what was observed for the WT, the activity of the Kelch domain of the E528K mutant mostly concerns the two BC loops of blade II and VI (Figure 3B). Moreover, the Glu-to-Lys substitution alters the flexibility and enhances the amplitude of the motions of a few selected regions, all located in the three terminal blades, that did not display high activity in the WT: the DA loop of blade IV (that directly faces the mutation site), BC loop of blade V and the DA loop of blade VI.

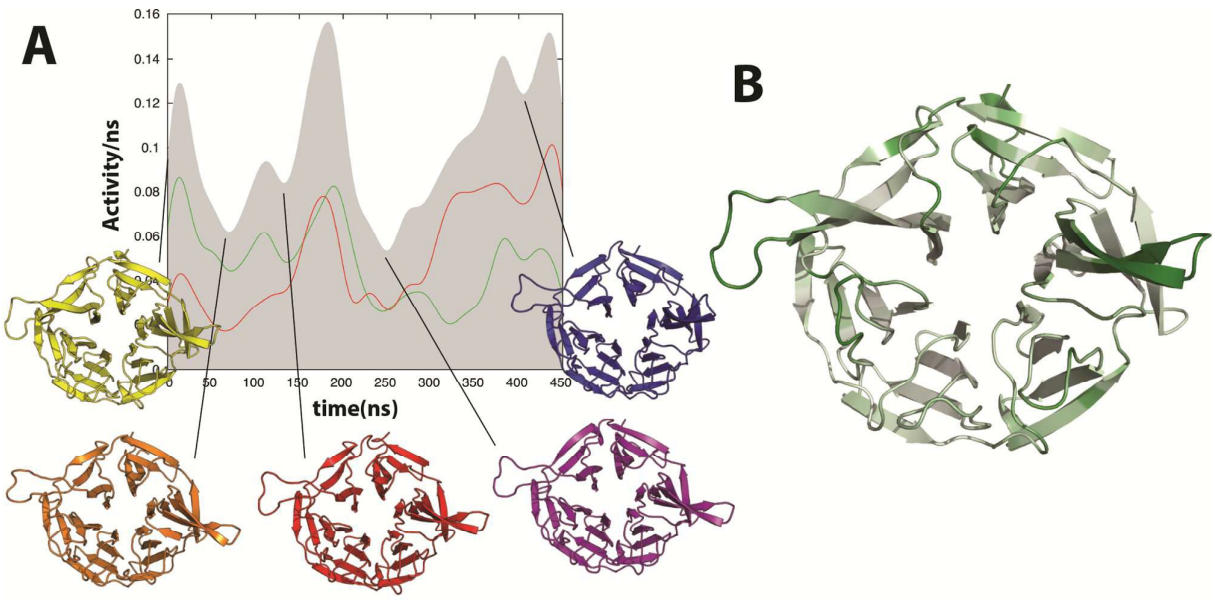


Figure 3. TimeScapes analysis of time evolution of dominant motions in the E528K mutant of Kelch domain of KLHL40 A) The grey profile show the global activity (i.e. broken and formed interactions) red and green profiles describe the activity associated with broken and formed contacts, respectively. The structure of each minima is represented as colored cartoon. B) The regions of the Kelch domain that provide major contributions to the protein activity measured by

View Article Online
DOI: 10.1039/C6RA06312H

TimeScapes are highlighted with shade of color from white (small contribution) to green (high contribution).

To further investigate structural alterations in the E528K mutant, we performed PCA, calculated over the main-chain atoms of the proteins, on a “macro-trajectory” obtained by concatenating those of the wild type and the mutant forming a single MD ensemble. The first three PCs account together for more than 60% of the global motions of the domain: therefore, they describe the higher magnitude large-scale motions of the propeller and they can be used as essential space for the representation of the differential fluctuations of WT KLHL40 and of its mutant. The contour plot, obtained by projecting the macro-trajectory on the essential plane defined by the first and second PC, shows that the WT and mutant species populate two conformational basins with minimal overlap (Figure 4A): this means that they have significantly different motions, contributing each to one of the identified two states. To identify with higher precision what protein regions have different dynamics upon mutation, we calculated the RMSF of main chain atoms along each of the first 3 PCs of the macrotrajectory. In this way we could point out the residues that mostly contribute to each PC: they are represented with thicker radius, and with colour coding ranging from blue (low contribution) to red (high), in Figure 4B. The first PC describes motions mainly associated with fluctuations in the BC loops of blade VI, but it also comprises motions of loop BC and DA of blade V, loop DA of blade IV, loop BC of blade III and loop DA of blade VI. The PCA analysis points out that these are the regions of the Kelch domain that undergo major changes in their flexibility between mutant and WT during MD simulations.

View Article Online
DOI: 10.1039/C6RA06312H

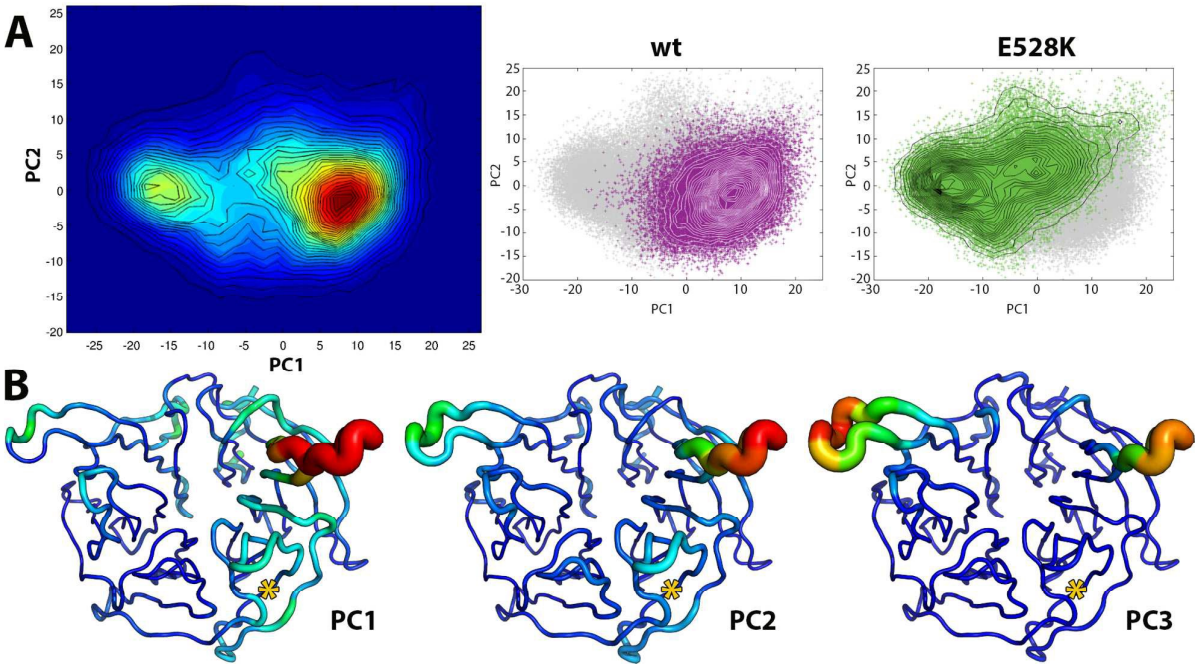


Figure 4. Dominant motions of the Kelch domain are altered by the E528K mutation. A) Bidimensional probability density distribution of the values of PC1 and PC2 (red: high probability; blue: low probability). The main different motions between WT and mutant are described by PC1. Right panel. Bidimensional plot of the contribution of each simulation (purple: WT; mutant: green) to the ensemble defined by PC1 and PC2. B) RMSF of main chain atoms calculated along each of the first 3 PCs of the macrotrajectory. The values are represented as shade of colors, from blue (low values- low flexible regions) to red (high values – high flexible regions). The mutated residue is indicated by a yellow star.

The high-magnitude motions of loop BC of blade VI can possibly hide differences, induced by mutations, in the dynamics of other regions of the Kelch domain: we then performed PCA considering separately each of the six blades of the Kelch domain. This allows us to look for changes, in the dynamics of β-propeller, investigating blade by blade, and see what subdomains are mostly affected by the mutation. The corresponding probability density contour plots and projection of the trajectory along the first PCs are shown in the SI: the main outcome of this analysis is that the fluctuations of blades II and III are similar in WT and mutant, while different conformational basins can be observed for blade I, blade

View Article Online
DOI: 10.1039/C6RA06312H

IV, blade V and blade VI, suggesting that the introduction of the mutations in the blade V has not only local effects but also long-range downfalls on the flexibility of other distant protein sites.

The E528K mutation has short-range effects in the H-bond network and electrostatic interactions

Besides these long-range effects of the mutation, local effects on the flexibility of the protein region in the proximity of the mutation (i.e. the whole blade V, which contains the substituted residue 528) could be observed. To better investigate the local alterations associated with E528K mutations we used the PyInteraph suite of tools,[46] which allows calculating networks of weak intramolecular interactions from conformational ensembles of proteins, by monitoring distances and geometries of atomic contacts. Such interactions are represented by a weighted graph, composed by nodes (residues) and edges (interactions occurring between the residues). We compared WT and mutant in terms of occurrence and persistence of three intramolecular interactions: hydrogen bonds, hydrophobic and electrostatic interactions. The results are mapped in Figure 5.

View Article Online
DOI: 10.1039/C6RA06312H

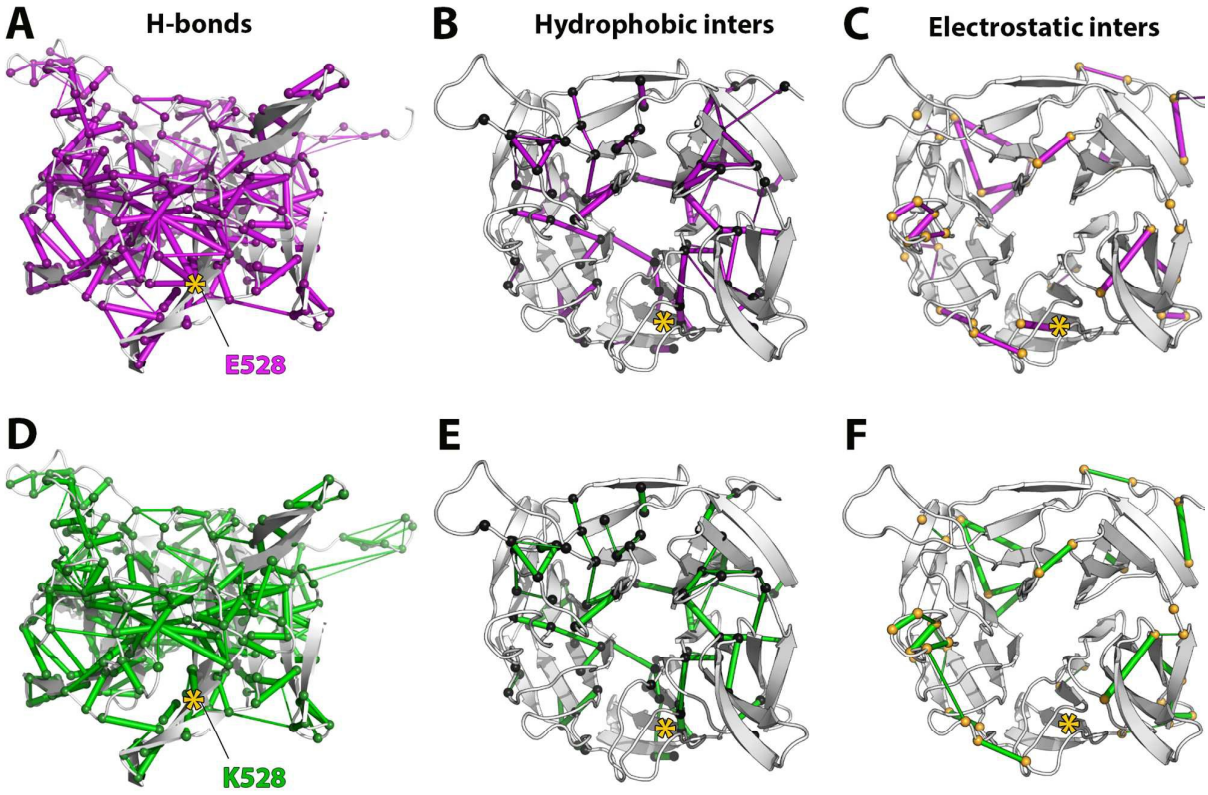


Figure 5. Networks of non covalent intramolecular interactions in the Kelch domain of WT (top, purple) and E528K mutant (bottom, green) of KLHL40. Hydrogen bonds are displayed in a lateral protein view; A: WT, D: mutant. Hydrophobic (B: WT, E: mutant) and electrostatic interactions (C: WT, F: mutant) are displayed from a top view. Only interactions with a minimum persistence of 20% during simulations were considered. Contacts are represented as sticks and the Cα of the residues involved in the interactions as spheres. The thickness of the sticks is proportional to their persistence during simulation time. The mutated residue is indicated by a yellow star.

The WT form of the Kelch domain is characterized by large network of highly persistent H-bonds (Figure 5A) that comprises all the blades and is important for the structural stability of the secondary structures of β-sheets and of the β-propeller architecture and functionality. Glu528 is involved in a network of hydrogen bonds principally with Arg500, Thr498, Gln184 and Met496. After the introduction of the mutation, the large majority of the high persistence hydrogen bonds are maintained, but all the hydrogen bonds involving

View Article Online
DOI: 10.1039/C6RA06312H

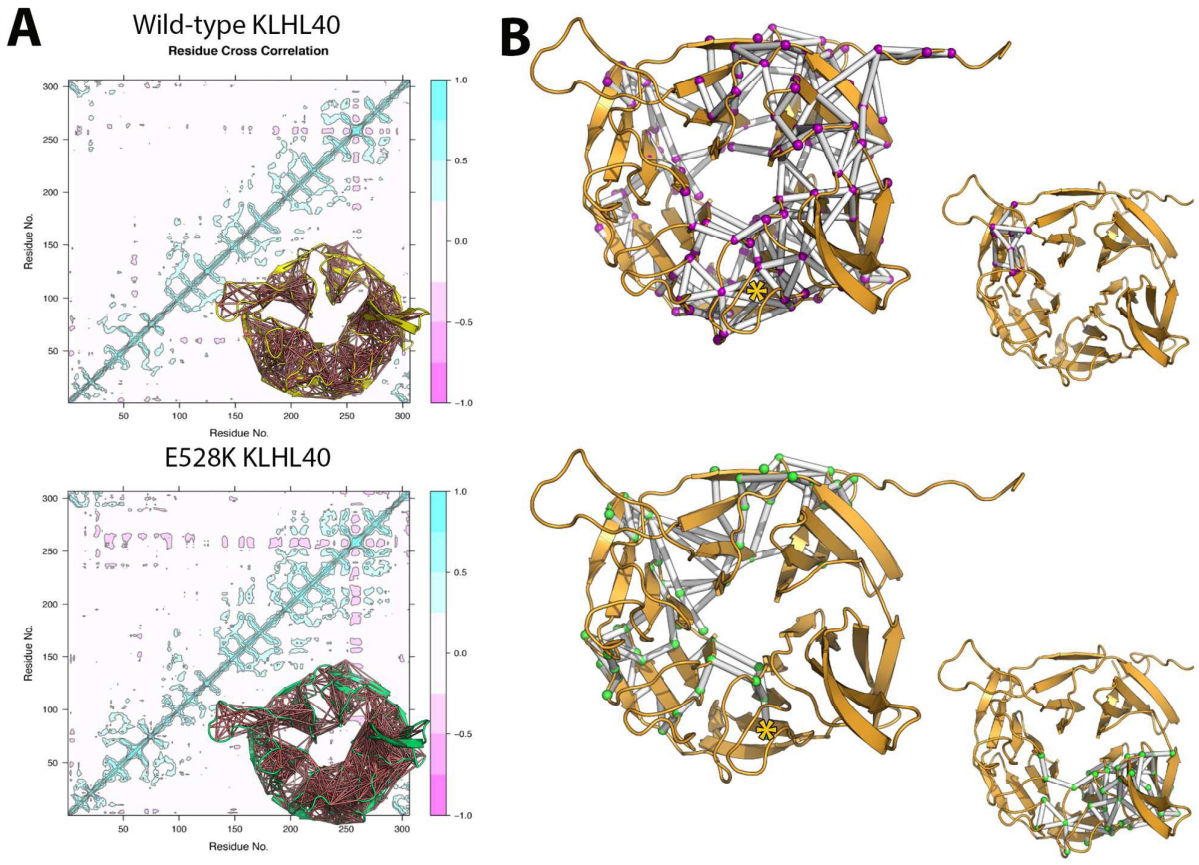
the residue 528 are lost (Figure 5D). We then turned our attention to hydrophobic interactions, which generally play an important role in the stabilization of the protein core and in the maintenance of the proper 3D structure.[52,53] Our analysis shows a wide and highly stable network of hydrophobic interactions that are conserved both in the WT and in the mutant (Figure 5B and 5E, respectively). As for the electrostatic interactions, we observed several persistent salt bridges in both WT and mutant simulations (Figure 5C and 5F, respectively); the analysis of the WT also shows highly persistent (70% of simulation time) salt bridge between Glu528 and Arg500, which disappears upon mutation.

The E528K mutation alters the communication between the propeller blades

So far, we have seen that the E528K NM-causing mutation leaves the overall fold of the propeller substantially unaltered, but it leads to a few, though important, structural local modifications (alteration of H-bonds network and electrostatic interactions), resulting in changes in the dynamic features of the molecule even in regions that are rather far from the mutation site. These findings suggest that the effects of the mutation somehow propagate themselves along the Kelch domain, most likely affecting the internal communication between protein subnetworks, which is known to be crucial to the allosteric regulation of a protein activity and, more generally, to the proper fulfilling of the physiological task of each biomolecule[54–56]. To shed light on the effects of structural local changes on the concerted motions of the whole domain, we performed network analysis using the tools included in both PyInteraph and Bio3D suites.[44,57] Both methods are inspired by the graph theory and they are based respectively on the analysis of distance-based contacts between atoms in proteins and cross correlation matrix (CMM). Figure 6A displays the CMM calculated via Bio3D; in the CMM, blue and pink spots identify correlated (i.e. moving along the same direction) and anti-correlated (moving along opposite directions) regions, respectively. Together with each matrix, the CMM values

View Article Online
DOI: 10.1039/C6RA06312H

have also been plotted on the corresponding three-dimensional structure, to help the reader grasping the effects of the mutation on the internal correlations. In the WT, there are five 5 areas of significant positive correlation (roughly spanning residues 350-400, 410-460, 450-510, 500-550, 540-600) that correspond to the blades II to VI of the domain. There is dense communication and high correlation *within* the same blade, each behaving like a small subdomain. In the E528K, although five areas of positive correlation can still be identified, one can also see changes in the correlation within blade V and, most importantly, significant correlation *between* blades, and in particular between blade V, that carries the mutation, and blades IV and VI. This result suggests that the mutation alters the coupling of the motions in the blade V, which loses its identity as a small subdomain, altering also its communications with other blades.



View Article Online
DOI: 10.1039/C6RA06312H

Figure 6. Long range conformational changes associated with E258K mutation. A) Cross Correlation Matrix (CMM), calculated with Bio3D, for WT (top) and E528K (bottom); turquoise: residues with highly correlated motions; pink: residues with anticorrelated motions. In the 3D reprsentation superimposed on the plot, correlated residues in the range of 0.25-1.0 values are represented as red sticks. Residues numbering is 1 to 309, corresponding to residues 314 to 622 in the mature protein. B) Network analysis of the weak intramolecular interactions (hydrogen bonds, hydrophobic and electrostatic interactions) for WT (top) and E528K (bottom), calculated using the PyInteraph suite. The two biggest populated sub-networks for the WT (top, purple spheres) and E528K mutant (bottom, green spheres) are shown. The interactions involved are represented as white sticks. The mutated residue is indicated by a yellow star.

To investigate how far the perturbations induced by the mutation can affect the network of intramolecular contacts, we calculated the sub-networks (or connected components) in the protein structure graph (Figure 6B). These are groups of residues, in which all the nodes are linked by at least one edge with the other nodes in the sub-network, but no connections exist between the nodes of the connected component and the rest of the graph. This analysis allows identifying protein regions that have strong internal communications. Although WT and mutant Kelch domain in the respective MD ensembles have a similar number of sub-networks (11 major sub-networks were identified for both species) their composition is very different. In the WT the main sub-network (i.e. the sub-network composed by the highest number of nodes) comprises residues from all the blades and describes contacts all over the beta propeller. These results suggest the existence of a global network of structural communication between the blades inside the Kelch domain. On the contrary, in the mutant the network of contacts between the blades is split in two major sub-networks: the main one comprises residues from blade I, II, III and IV and only communicates with few residues in blade V and VI, including the mutated residue 528. The second main sub-network in the mutant Kelch comprises residues from

View Article Online
DOI: 10.1039/C6RA06312H

blade V and VI, with only a few connections with blade III and I. These evidences point out that the mutation strongly affects the local contacts, causing alteration in the network of communications within blades V and VI and between these two sub-domains and the whole propeller: the propeller loses its nature of a whole, complex and unique sub-network.

Globally, the correlation and network analyses suggest that in the WT Kelch domain there is dense communication and coupled motions within each blade and concerted motions and contacts between the blades of most likely functional importance. The mutation modifies the contacts involving the whole blade V, and also long-range communication and tight coupling with other blades.

Glu528 is a hub in the correlated motions of the whole propeller and crucial to its functionality

We then turned our attention to the identification of the hub residues, which are defined as highly interconnected residues in protein structures (i.e. residues connected by more than three edges to their neighbours). Hub residues have been suggested to be central in the structural communications inside the protein and important for the maintenance of the secondary structure and protein fold.[48,58] Therefore, significant changes in the number and location of hub residues in a protein, e.g. induced by a mutation, is likely to underlie dramatic changes in the protein concerted motions, that are crucial for a biomolecule to fulfil its physiological task. Our calculations show that the total number of hub residues is similar in WT and E528K variant (97 and 93, respectively), but some important differences can be observed (Figure 7).

View Article Online
DOI: 10.1039/C6RA06312H

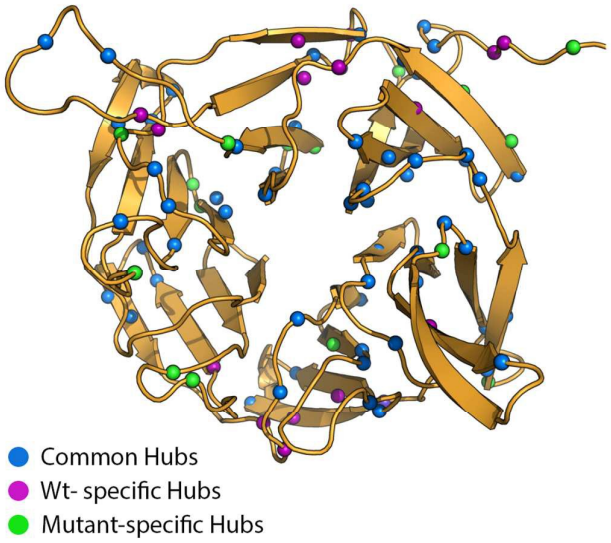


Figure 7. Hub residues in the Kelch domain of KLHL40. The hubs are displayed as coloured spheres; blue: hub residues that are common to the WT and mutant; green: residues that are predicted as hubs only in the E528K mutant; purple: residues that are predicted as hubs only in the the WT.

Probably, the most important one is that Glu528 is identified as a hub in the WT but not in E528K: this implies that the NM-associated mutation is located on a site of fundamental importance to the protein dynamics. Moreover, four more residues located close to the mutation site in blade V and IV (Ala494, Thr498, Met496 and Trp537) are also predicted as hubs only in the WT.

The mutation induces changes in the surface electrostatic potential, with potential downfalls on the interaction with physiological partners.

Once we had identified local and long-range downfalls of the NM-causing mutation, we tried to correlate alterations at the structure/dynamic level to changes in physico-chemical properties that would impact on the protein functionality. The main role of the Kelch propeller is to bind a number of different (in most cases, yet to be identified) substrates to mediate (or, in some cases, prevent from) their degradation. The surface properties of the

View Article Online
DOI: 10.1039/C6RA06312H

Kelch proteins with available structural data are known to be highly diverse:[7] one of the features displaying large differences between family members is the surface electrostatic potential, that is known to play a major role in a number of intermolecular interactions in biological systems.[59–64] For Kelch proteins, the high variability of surface electrostatic potential was suggested to be the key to the wide diversity of substrates with which they interact. We calculated the surface electrostatic potential for representative structures of both the WT and the E528K mutant (Figure 8): the surfaces of two species display markedly different electrostatic properties.

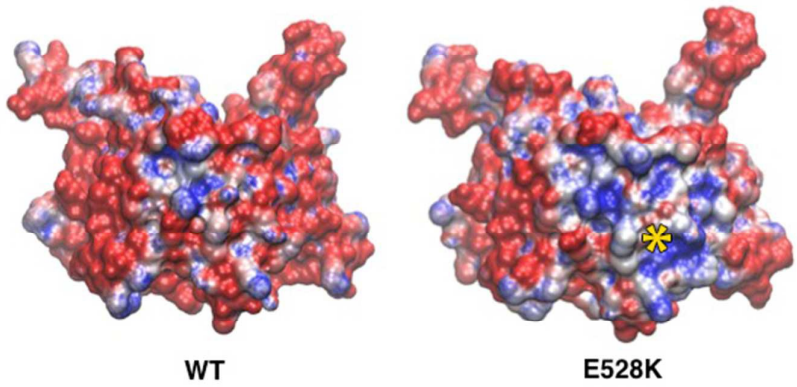


Figure 8. Electrostatic potential mapped on the molecular surface of WT (left) and E528K mutant (right). Potentials less than –5 kT/e are coloured in red, and those greater than +5 kT/e are depicted in blue. The mutated residue is indicated by a yellow star.

As could be expected, based on the mutation of a negatively charged residue with a positive one, there is an increase of the region of positive potential, but these differences extend way beyond the mutation site, indicating a significant rearrangement of the surface charges. This mutation-induced change is most likely to be a key determinant in the altered interaction of KLHL40 with its partners.

Conclusions

View Article Online
DOI: 10.1039/C6RA06312H

In this paper, we have investigated the dynamics of the β-propeller of WT human protein KLHL40, and we have unveiled the concerted motions and internal communication networks underlying its physiological activity. We compared the WT dynamic signature with that of the E528K variant: we found that the substitution does not lead to complete unfolding, but it alters a number of interactions among residues in close proximity to the mutation site. Moreover, long-range effects on the protein dynamics can be observed, meaning that the structural modifications induced by the mutation propagate along the structure unbalancing the dense communication within and between blades and affecting the concerted motions that are crucial for the protein dynamics. Our results unveil the complex network of interactions, conferring stability and functionality to the β-propeller, which is most likely general to other members of the Kelch superfamily. Moreover, as we found that upon NM-causing mutation there is no dramatic loss of protein architecture or stability, our molecular investigation might constitute a starting point for the rational design of drugs targeting KLHL40 malfunctioning under pathological conditions, to reset the system to its proper dynamics/function relationship. In particular, the comparison of the communication pathways under physiological conditions (WT) with those altered by the mutation will allow for the identification of crucial sites to be addressed, after careful evaluation of their druggability. Rational design of small oligopeptides or drugs might be then carried on extracting representative structures sampled and selected within our simulations.

Acknowledgements

The authors acknowledge the CINECA award under the ISCRA initiative (ISCRA C Projects DYNAMO and ReCAP) for the availability of high performance computing resources and support. The authors thank Dr. Gianina Ravenscroft and Prof. Nigel Laing at Harry Perkins Institute of Medical Research for fruitful discussions.

View Article Online
DOI: 10.1039/C6RA06312H

References

1 V. A. Gupta and A. H. Beggs, *Skelet. Muscle*, 2014, **4**, 11.

2 S. Prag and J. C. Adams, *BMC Bioinformatics*, 2003, **4**, 42.

3 J. Adams, R. Kelso and L. Cooley, *Trends Cell Biol.*, 2000, **10**, 17–24.

4 P. J. Stogios and G. G. Privé, *Trends Biochem. Sci.*, 2004, **29**, 634–7.

5 B. S. Dhanoa, T. Cogliati, A. G. Satish, E. a Bruford and J. S. Friedman, *Hum. Genomics*, 2013, **7**, 13.

6 Y.-R. Lee, W.-C. Yuan, H.-C. Ho, C.-H. Chen, H.-M. Shih and R.-H. Chen, *EMBO J.*, 2010, **29**, 1748–61.

7 P. Canning, C. D. O. Cooper, T. Krojer, J. W. Murray, A. C. W. Pike, A. Chaikuad, T. Keates, C. Thangaratnarajah, V. Hojzan, V. Ayinampudi, B. D. Marsden, O. Gileadi, S. Knapp, F. von Delft and A. N. Bullock, *J. Biol. Chem.*, 2013, **288**, 7803–14.

8 S. Shibata, J. Zhang, J. Puthumana, K. L. Stone and R. P. Lifton, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, 7838–43.

9 P. Rondou, G. Haegeman, P. Vanhoenacker and K. Van Craenenbroeck, *J. Biol. Chem.*, 2008, **283**, 11083–11096.

10 M. Wakabayashi, T. Mori, K. Isobe, E. Sohara, K. Susa, Y. Araki, M. Chiga, E. Kikuchi, N. Nomura, Y. Mori, H. Matsuo, T. Murata, S. Nomura, T. Asano, H. Kawaguchi, S. Nonoyama, T. Rai, S. Sasaki and S. Uchida, *Cell Rep.*, 2013, **3**, 858–868.

11 M. Zhuang, M. F. Calabrese, J. Liu, M. B. Waddell, A. Nourse, M. Hammel, D. J.

View Article Online
DOI: 10.1039/C6RA06312H

Miller, H. Walden, D. M. Duda, S. N. Seyedin, T. Hoggard, J. W. Harper, K. P. White and B. a Schulman, *Mol. Cell*, 2009, **36**, 39–50.

12 A. Boizot, Y. Talmat-Amar, D. Morrogh, N. L. Kuntz, C. Halbert, B. Chabrol, H. Houlden, T. Stojkovic, B. a Schulman, B. Rautenstrauss and P. Bomont, *Acta Neuropathol. Commun.*, 2014, **2**, 47.

13 G. Ravenscroft, S. Miyatake, V.-L. Lehtokari, E. J. Todd, P. Vornanen, K. S. Yau, Y. K. Hayashi, N. Miyake, Y. Tsurusaki, H. Doi, H. Saitsu, H. Osaka, S. Yamashita, T. Ohya, Y. Sakamoto, E. Koshimizu, S. Imamura, M. Yamashita, K. Ogata, M. Shiina, R. J. Bryson-Richardson, R. Vaz, O. Ceyhan, C. a Brownstein, L. C. Swanson, S. Monnot, N. B. Romero, H. Amthor, N. Kresoje, P. Sivadorai, C. Kiraly-Borri, G. Haliloglu, B. Talim, D. Orhan, G. Kale, A. K. Charles, V. a Fabian, M. R. Davis, M. Lammens, C. a Sewry, A. Manzur, F. Muntoni, N. F. Clarke, K. N. North, E. Bertini, Y. Nevo, E. Willichowski, I. E. Silberg, H. Topaloglu, A. H. Beggs, R. J. N. Allcock, I. Nishino, C. Wallgren-Pettersson, N. Matsumoto and N. G. Laing, *Am. J. Hum. Genet.*, 2013, **93**, 6–18.

14 N. Sambuughin, K. S. Yau, M. Olivé, R. M. Duff, M. Bayarsaikhan, S. Lu, L. Gonzalez-Mera, P. Sivadorai, K. J. Nowak, G. Ravenscroft, F. L. Mastaglia, K. N. North, B. Ilkovski, H. Kremer, M. Lammens, B. G. M. van Engelen, V. Fabian, P. Lamont, M. R. Davis, N. G. Laing and L. G. Goldfarb, *Am. J. Hum. Genet.*, 2010, **87**, 842–7.

15 C. W. Ockeloen, H. J. Gilhuis, R. Pfundt, E. J. Kamsteeg, P. B. Agrawal, A. H. Beggs, A. Dara Hama-Amin, A. Diekstra, N. V. a M. Knoers, M. Lammens and N. van Alfen, *Neuromuscul. Disord.*, 2012, **22**, 632–9.

16 P. B. Agrawal, R. S. Greenleaf, K. K. Tomczak, V.-L. Lehtokari, C. Wallgren-

View Article Online
DOI: 10.1039/C6RA06312H

Pettersson, W. Wallefeld, N. G. Laing, B. T. Darras, S. K. Maciver, P. R. Dormitzer and A. H. Beggs, *Am. J. Hum. Genet.*, 2007, **80**, 162–7.

17 V.-L. Lehtokari, K. Pelin, M. Sandbacka, S. Ranta, K. Donner, F. Muntoni, C. Sewry, C. Angelini, K. Bushby, P. Van den Bergh, S. Iannaccone, N. G. Laing and C. Wallgren-Pettersson, *Hum. Mutat.*, 2006, **27**, 946–56.

18 C. Wallgren-Pettersson, K. Pelin, K. J. Nowak, F. Muntoni, N. B. Romero, H. H. Goebel, K. N. North, A. H. Beggs and N. G. Laing, *Neuromuscul. Disord.*, 2004, **14**, 461–70.

19 K. J. Nowak, D. Wattanasirichaigoon, H. H. Goebel, M. Wilce, K. Pelin, K. Donner, R. L. Jacob, C. Hübner, K. Oexle, J. R. Anderson, C. M. Verity, K. N. North, S. T. Iannaccone, C. R. Müller, P. Nürnberg, F. Muntoni, C. Sewry, I. Hughes, R. Sutphen, a G. Lacson, K. J. Swoboda, J. Vigneron, C. Wallgren-Pettersson, A. H. Beggs and N. G. Laing, *Nat. Genet.*, 1999, **23**, 208–12.

20 N. G. Laing, D. E. Dye, C. Wallgren-pettersson, G. Richard, S. Lillis, T. L. Winder, H. Lochmüller, C. Graziano, S. Mitrani-rosenbaum, D. Tuomey, J. C. Sparrow, A. H. Beggs and K. J. Nowak, *Hum. Mutat.*, 2009, **30**, 1267–1277.

21 J. R. Nance, J. J. Dowling, E. M. Gibbs and C. G. Bönnemann, *Curr. Neurol. Neurosci. Rep.*, 2012, **12**, 165–174.

22 N. Sambuughin, K. S. Yau, M. Olivé, R. M. Duff, M. Bayarsaikhan, S. Lu, L. Gonzalez-Mera, P. Sivadorai, K. J. Nowak, G. Ravenscroft, F. L. Mastaglia, K. N. North, B. Ilkovski, H. Kremer, M. Lammens, B. G. M. van Engelen, V. Fabian, P. Lamont, M. R. Davis, N. G. Laing and L. G. Goldfarb, *Am. J. Hum. Genet.*, 2010, **87**, 842–7.

View Article Online
DOI: 10.1039/C6RA06312H

23 A. Garg, J. O. Rourke, C. Long, J. Doering, G. Ravenscroft, S. Bezprozvannaya, B. R. Nelson, N. Beetz, L. Li, S. Chen, N. G. Laing, R. W. Grange, R. Bassel-Duby and E. N. Olson, *J. Clin. Invest.*, 2014, **124**, 3529–3539.

24 E. J. Todd, K. S. Yau, R. Ong, J. Slee, G. McGillivray, C. P. Barnett, G. Haliloglu, B. Talim, Z. Akcoren, A. Kariminejad, A. Cairns, N. F. Clarke, M.-L. Freckmann, N. B. Romero, D. Williams, C. A. Sewry, A. Colley, M. M. Ryan, C. Kiraly-Borri, P. Sivadorai, R. J. N. Allcock, D. Beeson, S. Maxwell, M. R. Davis, N. G. Laing and G. Ravenscroft, *Orphanet J. Rare Dis.*, 2015, **10**, 148.

25 V. A. Gupta, G. Ravenscroft, R. Shaheen, E. J. Todd, L. C. Swanson, M. Shiina, K. Ogata, C. Hsu, N. F. Clarke, B. T. Darras, M. A. Farrar, A. Hashem, N. D. Manton, F. Muntoni, K. N. North, S. A. Sandaradura, I. Nishino, Y. K. Hayashi, C. A. Sewry, E. M. Thompson, K. S. Yau, C. A. Brownstein, T. W. Yu, R. J. N. Allcock, M. R. Davis, C. Wallgren-Pettersson, N. Matsumoto, F. S. Alkuraya, N. G. Laing and A. H. Beggs, *Am. J. Hum. Genet.*, 2013, **93**, 1108–1117.

26 W. Gong, R. M. Gohla, K. M. Bowlin, N. Koyano-Nakagawa, D. J. Garry and X. Shi, *J. Biol. Chem.*, 2015, **290**, 15350–15361.

27 K. Henzler-Wildman and D. Kern, *Nature*, 2007, **450**, 964–72.

28 L. Paltrinieri, M. Borsari, G. Battistuzzi, M. Sola, C. Dennison, B. L. de Groot, S. Corni and C. A. Bortolotti, *Biochemistry*, 2013, **52**, 7397–7404.

29 C. A. Bortolotti, A. Amadei, M. Aschi, M. Borsari, S. Corni, M. Sola and I. Daidone, *J. Am. Chem. Soc.*, 2012, **134**, 13670–8.

30 L. Paltrinieri, M. Borsari, A. Ranieri, G. Battistuzzi, S. Corni and C. A. Bortolotti, *J. Phys. Chem. Lett.*, 2013, **4**, 710–715.

View Article Online
DOI: 10.1039/C6RA06312H

31 M. Karplus and J. Kuriyan, *Proc. Natl. Acad. Sci. USA*, 2005, **102**, 6679–85.

32 Q. Zhao, *Rev. Theor. Sci.*, 2013, **1**, 1–19.

33 R. B. Best, X. Zhu, J. Shim, P. E. M. Lopes, J. Mittal, M. Feig and A. D. MacKerell, *J. Chem. Theory Comput.*, 2012, **8**, 3257–3273.

34 A. D. MacKerell, D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiórkiewicz-Kuczera, D. Yin and M. Karplus, *J. Phys. Chem. B*, 1998, **102**, 3586–3616.

35 A. D. Mackerell, M. Feig and C. L. Brooks, *J. Comput. Chem.*, 2004, **25**, 1400–15.

36 W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein, *J. Chem. Phys.*, 1983, **79**.

37 H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak, *J. Chem. Phys.*, 1984, **81**.

38 T. Darden, D. York and L. Pedersen, *J. Chem. Phys.*, 1993, **98**.

39 B. Hess, H. Bekker, H. J. C. Berendsen and J. G. E. M. Fraaije, *J. Comput. Chem.*, 1997, **18**, 1463–1472.

40 N. A. Baker, D. Sept, S. Joseph, M. J. Holst and J. A. McCammon, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 10037–41.

41 T. J. Dolinsky, J. E. Nielsen, J. A. McCammon and N. A. Baker, *Nucleic Acids Res.* , 2004, **32** , W665–W667.

View Article Online
DOI: 10.1039/C6RA06312H

42 T. J. Dolinsky, P. Czodrowski, H. Li, J. E. Nielsen, J. H. Jensen, G. Klebe and N. A. Baker, *Nucleic Acids Res.*, 2007, **35**, W522–W525.

43 W. Wriggers, K. a. Stafford, Y. Shan, S. Piana-Agostinetti, P. Maragakis, K. Lindorff-Larsen, P. J. Miller, M. P. Eastwood, R. O. Dror and D. E. Shaw, *J. Chem. Theory Comput.*, 2009, **5**, 2595–2605.

44 B. J. Grant, A. P. C. Rodrigues, K. M. ElSawy, J. A. McCammon and L. S. D. Caves, *Bioinformatics*, 2006, **22**, 2695–2696.

45 T. Ichiye and M. Karplus, *Proteins*, 1991, **11**, 205–17.

46 M. Tiberti, G. Invernizzi, M. Lambrughi, Y. Inbar, G. Schreiber and E. Papaleo, *J. Chem. Inf. Model.*, 2014, **54**, 1537–1551.

47 M. Pasi, M. Tiberti, A. Arrigoni and E. Papaleo, *J. Chem. Inf. Model.*, 2012, **52**, 1865–1874.

48 S. Vishveshwara and A. G. and P. Hansia, *Curr. Protein Pept. Sci.*, 2009, 10, 146–160.

49 K. V Brinda and S. Vishveshwara, *Biophys. J.*, 2005, **89**, 4159–4170.

50 T. N. Do, W.-Y. Choy and M. Karttunen, *J. Chem. Theory Comput.*, 2016, **12**, 395–404.

51 P. Canning, F. J. Sorrell and A. N. Bullock, *Free Radic. Biol. Med.*, 2015, **88**, 1–7.

52 M. Munson, S. Balasubramanian, K. G. Fleming, A. D. Nagi, R. O'Brien, J. M. Sturtevant and L. Regan, *Protein Sci.*, 1996, **5**, 1584–93.

53 N. Koga, R. Tatsumi-Koga, G. Liu, R. Xiao, T. B. Acton, G. T. Montelione and D. Baker, *Nature*, 2012, **491**, 222–227.

View Article Online
DOI: 10.1039/C6RA06312H

54 A. Pandini, A. Fornili, F. Fraternali and J. Kleinjung, *FASEB J.*, 2012, **26**, 868–881.

55 C. Böde, I. A. Kovács, M. S. Szalay, R. Palotai, T. Korcsmáros and P. Csermely, *FEBS Lett.*, 2007, **581**, 2776–2782.

56 N. V Dokholyan, *Chem. Rev.*, 2016.

57 L. Skjærven, X.-Q. Yao, G. Scarabelli and B. J. Grant, *BMC Bioinformatics*, 2014, **15**, 399.

58 K. Angelova, A. Felline, M. Lee, M. Patel, D. Puett and F. Fanelli, *Cell. Mol. Life Sci.*, 2010, **68**, 1227–1239.

59 D. Sept, a H. Elcock and J. a McCammon, *J. Mol. Biol.*, 1999, **294**, 1181–9.

60 K. A. Sharp and B. Honig, *Annu. Rev. Biophys. Biophys. Chem.*, 1990, **19**, 301–332.

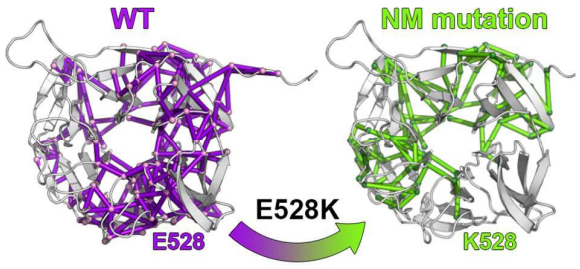
61 S. Han, S. Yin, H. Yi, S. Qiu and Z. Cao, 2010, 3118–3125.

62 F. B. Sheinerman, R. Norel and B. Honig, *Curr. Opin. Struct. Biol.*, 2000, **10**, 153–9.

63 I. Daidone, L. Paltrinieri, A. Amadei, G. Battistuzzi, M. Sola, M. Borsari and C. A. Bortolotti, *J. Phys. Chem. B*, 2014, **118**, 7554–7560.

64 Z. Serber and J. E. Ferrell Jr., *Cell*, 2016, **128**, 441–444.

View Article Online
DOI: 10.1039/C6RA06312H



The nemaline myopathy-associated E528K mutation in the KLHL40 alters the communication between the Kelch propeller blades