D.-H. Chen et al.: A Novel Mutation in *FHL1* in a Family with X-linked Scapuloperoneal Myopathy: Phenotypic Spectrum and Structural Study of *FHL1* Mutations.

Supplementary Information

Molecular dynamics (MD) simulations were propagated with the Enzymix force field [1,2] in steps of 1 fs, and structures were saved at 20 ps intervals for analysis. The force field includes flexible, three-center water molecules and uses a multipole expansion to treat charge-charge interactions of all the atoms in the system. For better long-term stability, we added pseudopotentials to keep the protein's center of mass near the center of the water sphere and hold the total angular momentum of the water close to zero.

Zn atoms were assigned a charge of 1.5, and each of their ligands a net charge of -0.5. The van der Waals A and B parameters used for Zn were 17.1 and 1.0, respectively. The force constants ($k_b/2$ and $k_\theta/2$) used for Zn-ligand bond stretching and bending during the 15-ns equilibration period were 600 kcal mol^{-1} Å^{-2} and 200 kcal mol^{-1} radian^{-2}, respectively. Following the equilibration trajectories, a series of simulations of the wild-type protein were run for periods of 0.50 to 0.84 ns using various values of the force constants. This was done in order to find ranges of values that kept the mean bond lengths and angles close to those seen in the NMR structure of FHL1 and in crystal structures of other similar zinc proteins (Zn-S distance ≈ 2.32 Å, Zn-N ≈ 2.09 Å, and ligand-Zn-ligand bond angles ≈109.5^o; [3-6]). Using values at the low ends of these ranges should avoid overconstraining the structures of the zinc centers in models of the mutant proteins. Figure S1 shows that the distances and angles diverge from the observed values when the force constants are reduced below about 50 kcal mol^{-1} Å^{-2} for stretching and 50 kcal mol^{-1} radian^{-2} for bending. (We did not try to distinguish between force constants for Zn-S and Zn-N bonds.) These values were used for all subsequent simulations of both the wild-type and mutant proteins.

To calculate the root-mean-square deviations (RMSD) of the Cα carbons from the initial coordinates, protein structures were aligned as described by Kabsch [7]. The deviations for each model are expressed relative to the initial, low-temperature structure obtained by optimizing the sidechain of the mutant sidechian, embedding the protein in water, and minimizing the local energies of the system by a 4-ps MD simulation at 30 K. During the following 15-ns equilibration period at 310 K, the RMSD value for the wild-type protein leveled off at about 0.75 Å, while those for the mutants stabilized at values between 0.75 and 1.2 Å. The RMSD values remained essentially constant during subsequent 12.5-ns trajectories with relaxed Zn-ligand force constants (Figure S2). All the results presented in Figs. 4, 5, and S2-S8 pertain to these last trajectories.

Solvent-accessible surface area (SASA) was calculated with SERF v.2.0 as described by Flower [8], using the Shrake-Rupley [9] algorithm with 642 points per

atom and a solvent radius of 1.4 Å. Minor modifications of the program were made to include hydrogen atoms explicitly, allow rapid averaging over large numbers of structures saved during an MD trajectory, and distinguish between polar and nonpolar surface areas. For the latter purpose, Zn, O, N, S, hydrogens bound to these atoms, and carbonyl and carboxyl C atoms were considered to be polar, and all other atoms (aliphatic and aromatic C atoms and their bound hydrogens) nonpolar.

The Excess Nonpolar Surface (EXNS) within radius *R* of residue *J* was defined as

$$EXNS(J,R) = \left\langle \frac{1}{I} \sum_{i=1}^{I(j,R)} \gamma_i A_i \right\rangle_{j(J)} , \qquad (1)$$

where index *j* runs over all atoms (*j*) of residue *J*; $\langle \cdots \rangle_{j(J)}$ denotes an average over these atoms; index *i* runs over all protein atoms within distance *R* of *j*; *I*(*j*,*R*) is the total number of protein atoms in the sphere around *j*; A_i is the SASA of atom *i*; and γ_i = +1 if atom *i* is nonpolar as defined above, and -1.5 if *i* is polar. Note that *EXNS*(*J*,*R*) considers the polarity and SASA of atoms of neighboring residues as well as those of residue *J* itself, and has units of $Å^2$/atom. It is similar in spirit to the dimensionless Spatial Aggregation Propensity (SAP) defined by Chennamsetty et al. [10], but differs in that the contribution of atom *i* depends on the polarity and accessible surface of only that atom, not the overall hydrophobicity of the residue to which *i* belongs. The methylene atoms of a Lys residue, for example, are nonpolar and so contribute positively to EXNS of the Lys and nearby residues, even though the ε-amino group makes lysine hydrophilic overall. In the calculations presented here, we used *R* = 8 Å. The magnitude of the weighting factor -1.5 was chosen to make EXNS negative for ~75% of the residues of a typical globular protein, where nonpolar atoms tend to be sequestered from the solvent. As discussed by Chennamsetty et al. [10] for SAP, a cluster of residues with positive EXNS represents an exposed patch of nonpolar surface that could lead to aggregation or contribute to specific interactions with other proteins.

Protein structures were examined and rendered by VMD 1.8.7 [11].

References

1. Lee FS, Chu ZT, Warshel, A. Microscopic and semimicroscopic calculations of electrostatic energies in proteins by the POLARIS and ENZYMIX programs. J Comp. Chem 1993, 14: 161-185.

2. Parson WW, Warshel A. Calculations of electrostatic energies in proteins using microscopic, semimicroscopic and macroscopic models and free-eneergy

perturbation approaches. In: Aartsma TJ, Matysik J, editors. Biophysical Techniques in Photosynthesis. Dordrecht: Springer 2008: 401-420.

3. Alberts IL, Nadassy K, Wodak SJ. Analysis of zinc binding sites in protein crystal structures. Protein Sci 1998, 7: 1700-1716.

4. Dudev T, Lim C. Metal binding affinity and selectivity in metalloproteins: insights from computational studies. Annu Rev Biophys 2008, 37: 97-116.

5. Calimet N, Simonson T. Cys_xHis_y-Zn^{2+} interactions: Possibilities and limitations of a simple pairwise force field. J Mol Graph Model 2006, 24: 404-411.

6. Tamames B, Sousa SF, Tamames J, Fernandes PA, Ramos MJ. Analysis of zinc-ligand bond lengths in metalloproteins: trends and patterns. Proteins, 2007, 69: 466-475.

7. Kabsch W. A solution for the best rotation to relate two sets of vectors. Acta Cryst 1976, A34: 827-828.

8. Flower D. SERF: A Program for accessible surface area calculations. J Comput Graphics & Modelling 1997, 15: 238-244.

9. Shrake A, Rupley JA. Environment and exposure to solvent of protein atoms. Lysozyme and insulin. J Mol Biol 1974, 79: 351-372.

10. Chennamsetty N, Voynov V, Kayser V, Helk B. Trout BL. Design of therapeutic proteins with enhanced stability. Proc. Natl. Acad. Sci. USA 2009, 106: 11937-11942.

11. Humphrey W, Dalke A, Schulten K. VMD - Visual molecular dynamics. J Mol Graph 1996, 14: 33-38.

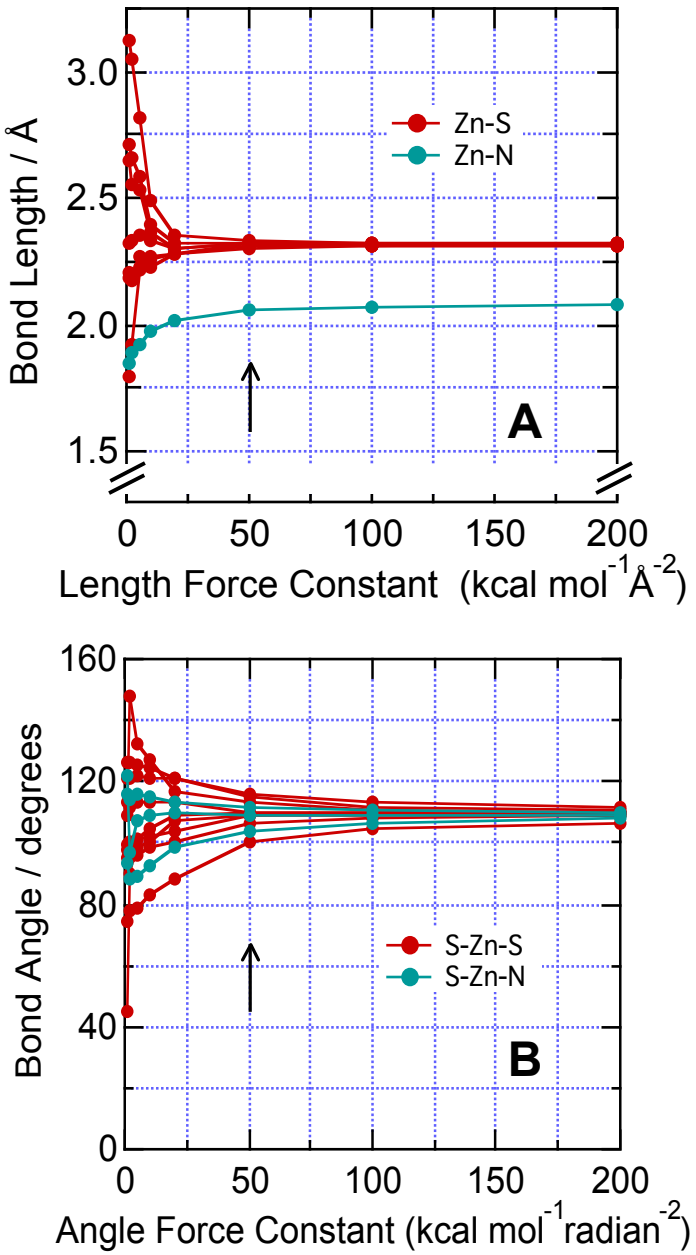

Figure S1. Mean lengths (A) and angles (B) of Zn-ligand bonds in the 2nd LIM domain of wild-type FHL1 as functions of the bond-stretching and -bending force constants. Each value represents an average of 25 or more protein structures saved at 20-ps intervals during an MD trajectory at 310 K, following 15 ns equilibration with stiffer force constants and, as required for convergence, 20 to 100 ps equilibration with the indicated values. In A, the force constant ($k_\theta/2$) for S-Zn-S and S-Zn-N angles was fixed at 50 kcal mol^{-1} radian^{-2}; in B, the force constant ($k_b/2$) for Zn-S and Zn-N bond lengths was 50 kcal mol^{-1} Å^{-2}. The energy minima for Zn-S and Zn-N bond lengths were set at 2.32 and 2.09 Å, respectively, and those for S-Zn-S and S-Zn-N angles, at 109.5°. The vertical arrows indicate the values used for subsequent simulations.

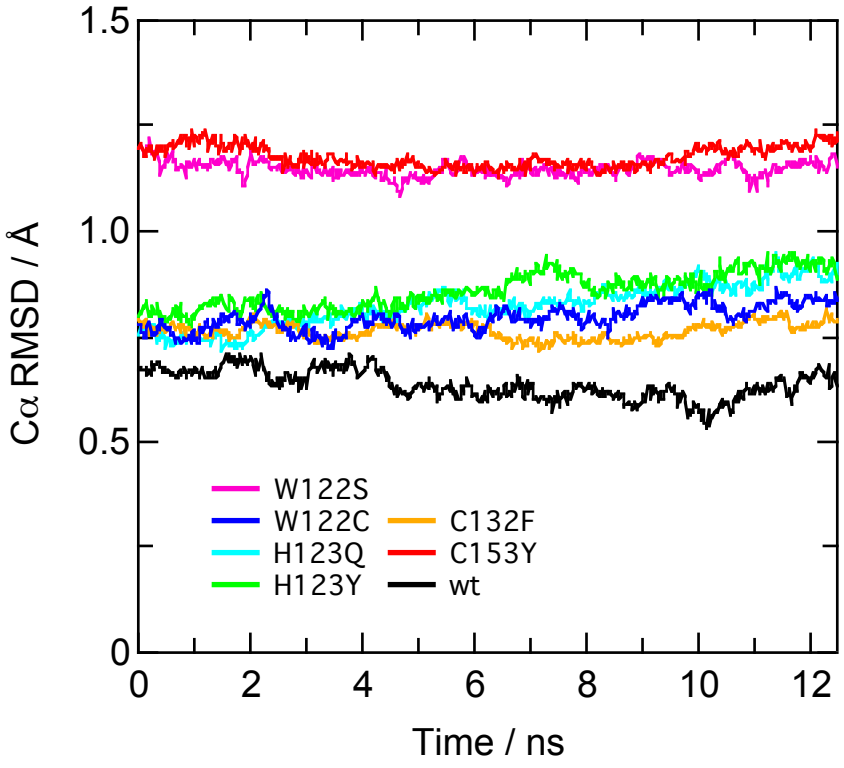

Figure S2. Root-mean-square displacements (RMSDs) of the Cα carbons of residues 99-157 of the 2nd LIM domain of wild type (wt) FHL1 and six mutants during MD trajectories at 310 K. The reference state for each protein is the initial, low-temperature structure obtained by optimizing the mutant sidechain, embedding the protein in water, and minimizing the local energies of the system by MD at 30 K. The trajectories represented here were preceded by 15-ns trajectories for equilibration of the protein. The force constants for stretching and bending of Mg-ligand bonds were adjusted 20 ps before the first time points plotted. Here and in the following figures, black traces represent wt FHL1, and magenta, blue, cyan, green, orange and red represent the W122S, W122C, H123Q, H123Y, C132F and C153Y mutants, respectively.

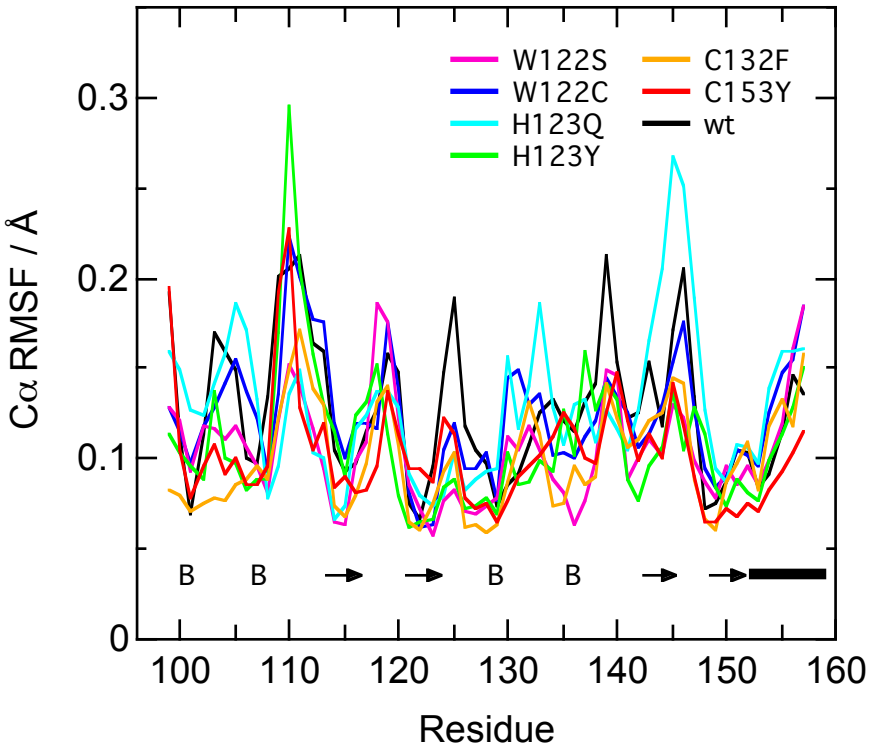

Figure S3. Root-mean-square fluctuations of the Cα carbons from their mean positions as functions of the residue number, averaged over 12.5-ns MD trjectories of wt FHL1 and six mutants at 310 K. Secondary structural elements are indicated below the traces: B = bridge β; arrow = extended β; bar = α-helix. The mutations have relatively little effect on the structural fluctuations of the protein.

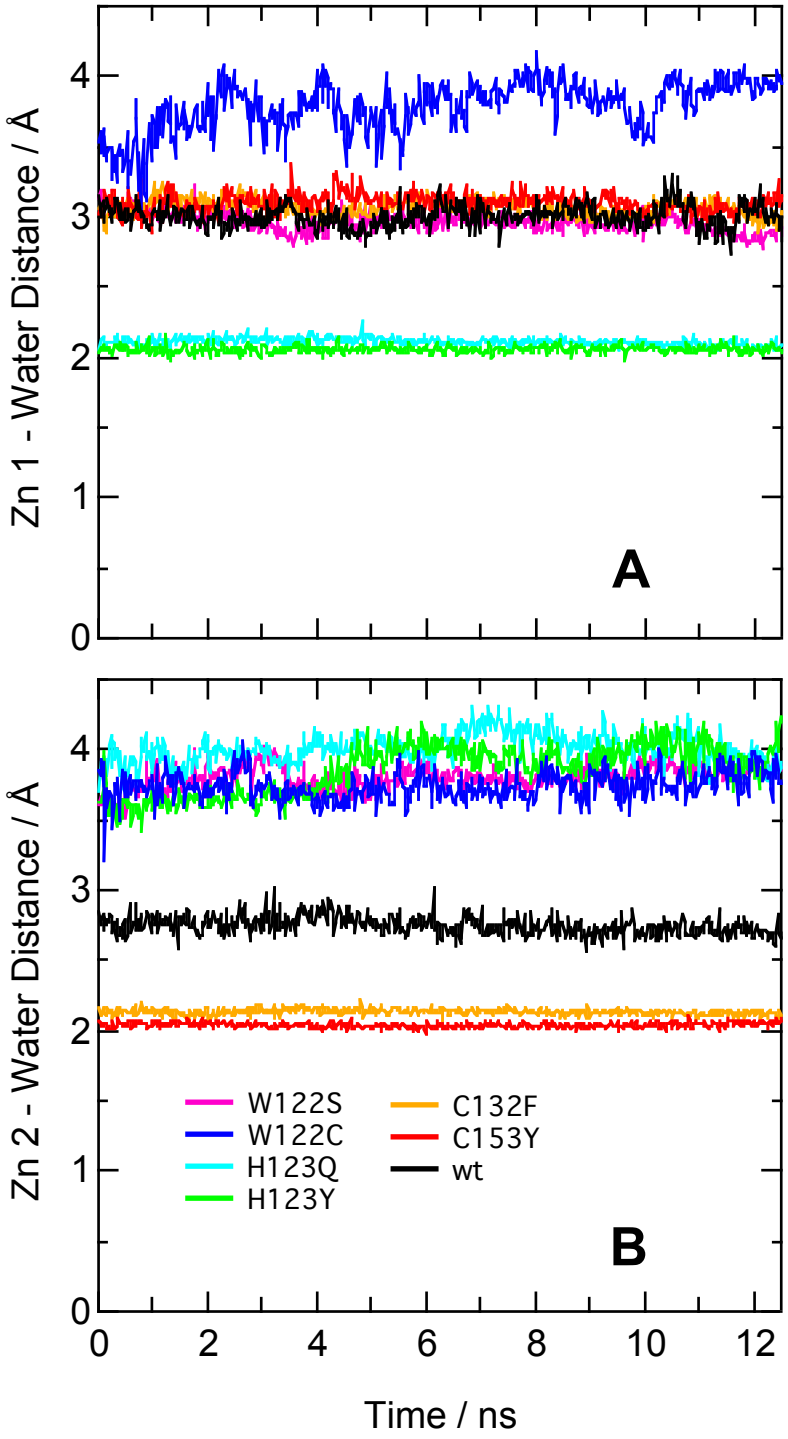

Figure S4. Distances to the nearest water O atom from Zn1 (A) and Zn2 (B) in wt FHL1 and six mutants during MD simulations at 310 K. Water replaces the missing ligand of Zn1 in the H123Q and H123Y mutants, and the missing ligand of Zn2 in C132F and C132Y. (All these movements occurred during the first 0.5 ns of the equilibration trajectories that preceded the trajectories shown.) The mean distance from OE1 of Q123 to Zn1 in the H123Q mutant was 2.31±0.03 Å.

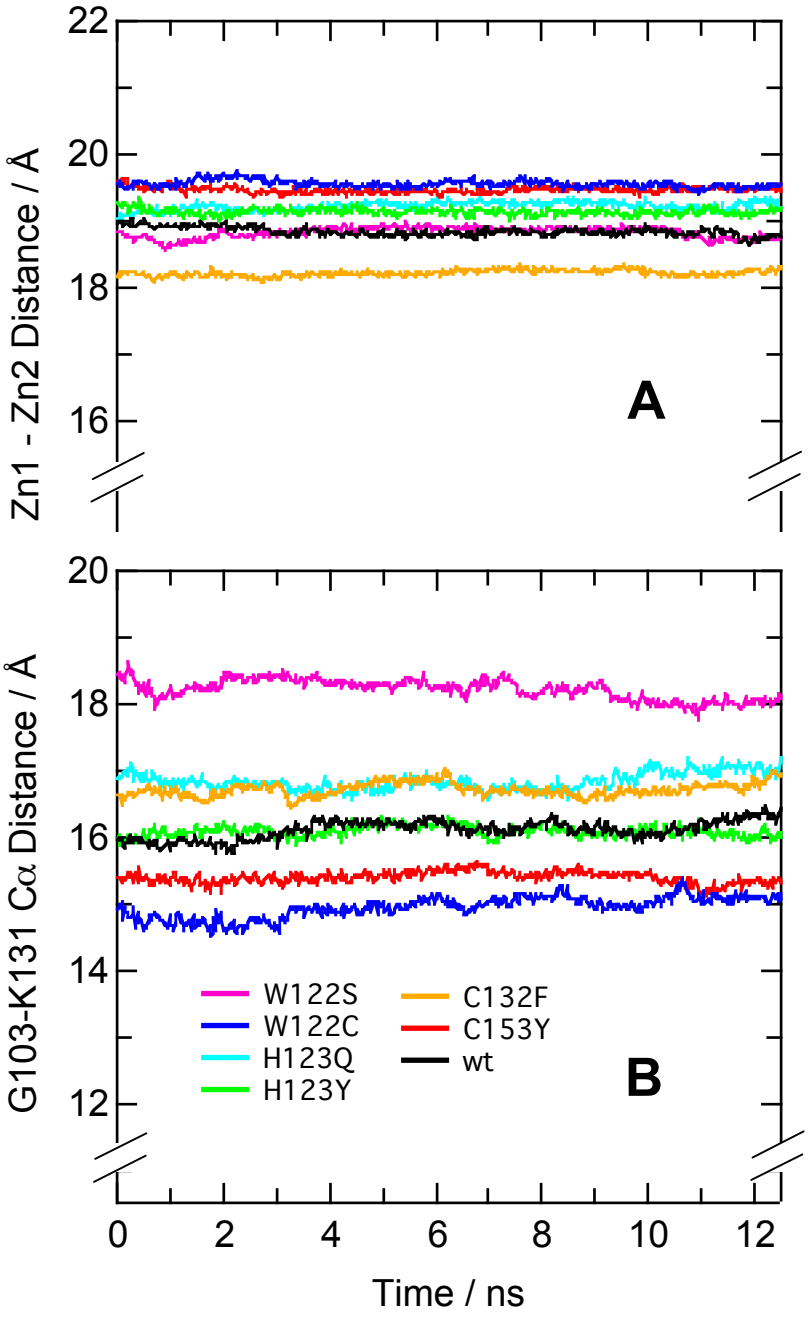

Figure S5. Distance between the two Zn atoms (A) and between the Cα atoms of residues at the ends of the two β-hairpins (G103 and K131, B) in wt FHL1 and six mutants during MD trajectories at 310 K.

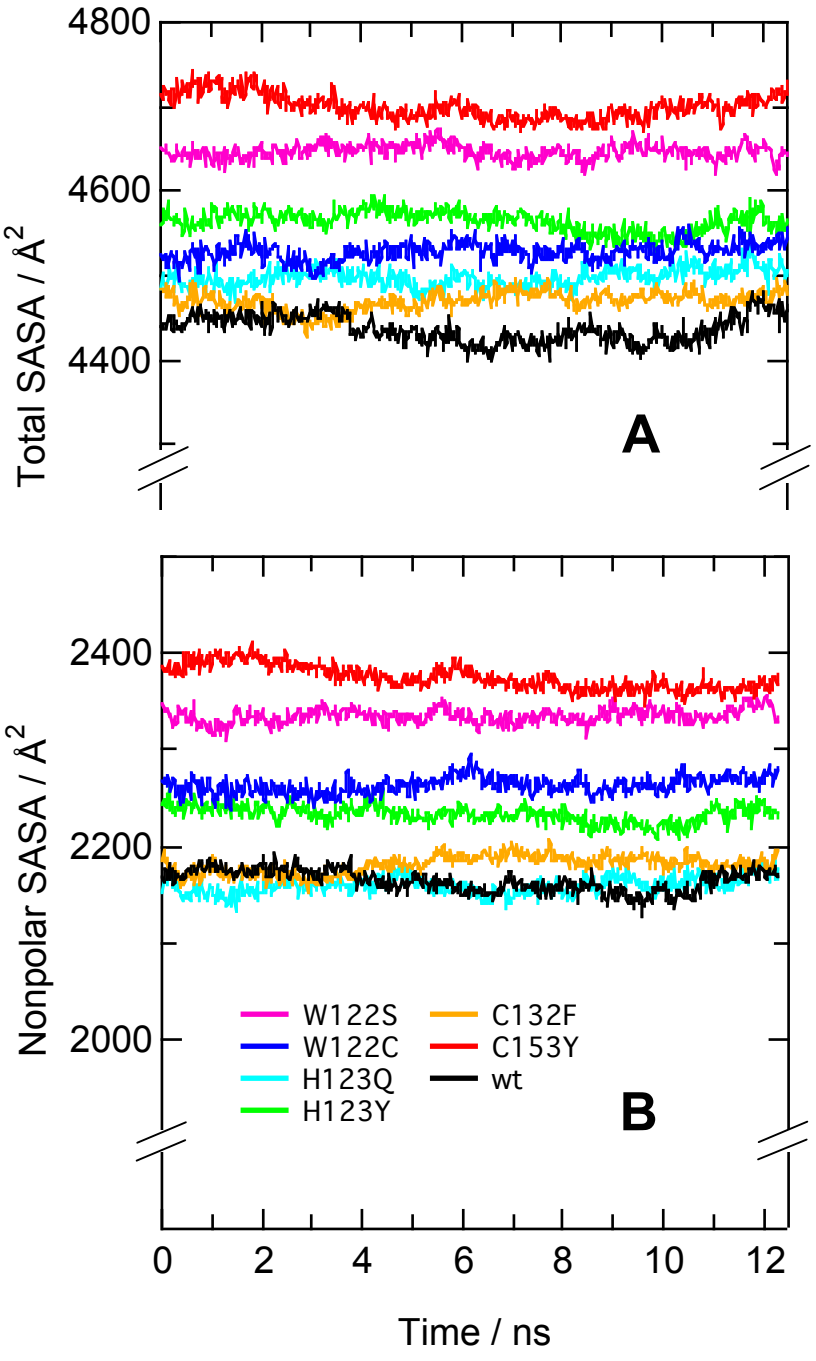

Figure S6. Total (A) and non-polar (B) solvent-accessible surface area (SASA) of residues 99-157 during MD simulations of wt FHL1 and six mutants at 310 K.

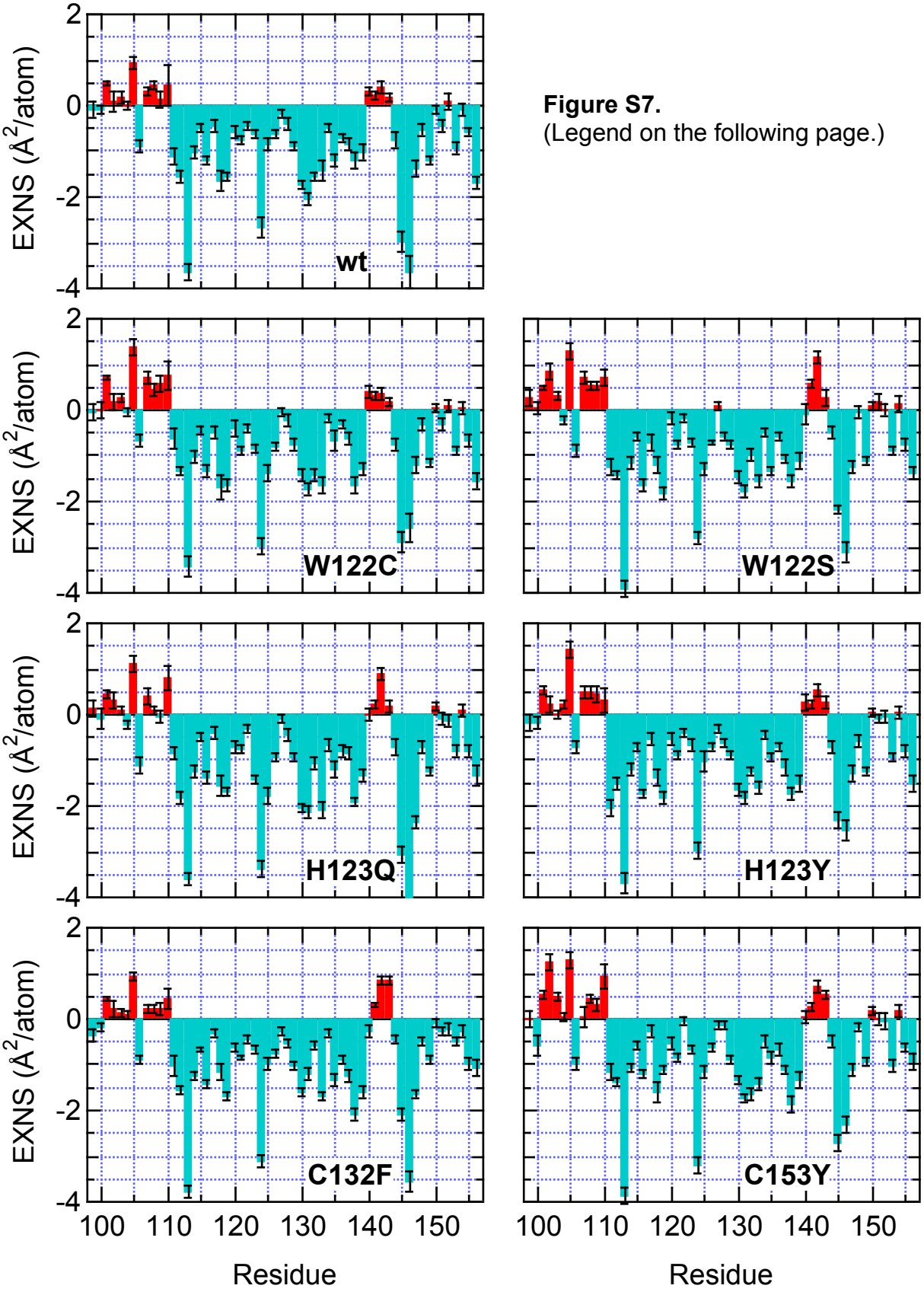

Figure S7.
(Legend on the following page.)

Figure S7. Excess nonpolar surfaces (EXNS) of wt FHL1 and six mutants. EXNS was evaluated by equation 1 with R = 8 Å and was averaged over 12.5-ns trajectories at 310 K. The error bars are the standard deviations from the means. Red color indicates residues with positive EXNS; cyan, negative values. Positive EXNS values occur in two clusters of residues that are close together space (see Fig. 5). The mutant proteins tend to have higher EXNS and/or more residues with positive values relative to the wt protein.

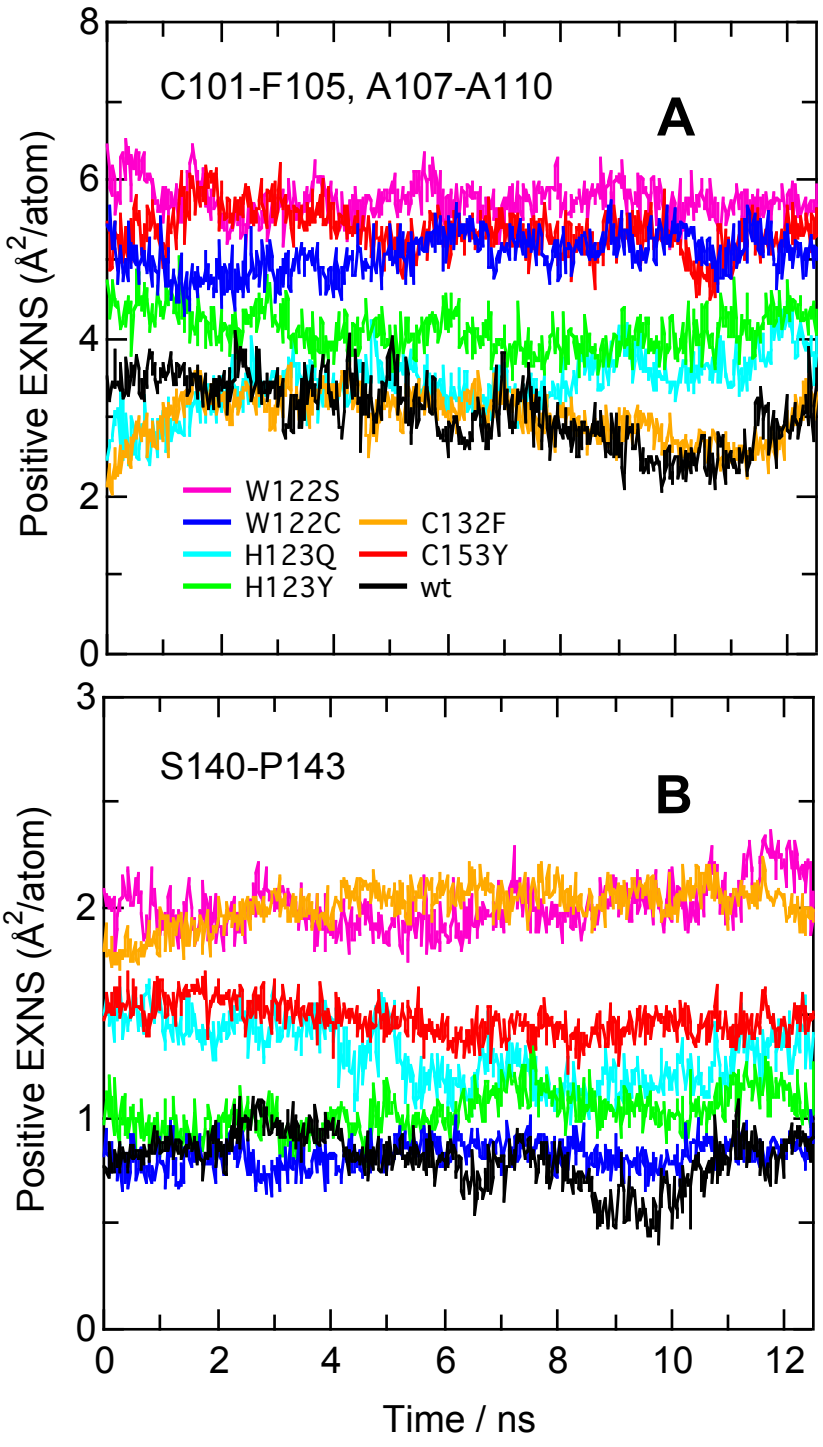

Figure S8. Summed positive excess nonpolar surfaces (EXNS) of atoms in and surrounding residues C101-F105 and A107-A110 (A) and S140-P143 (B) in WT FHL1 and six mutants during MD simulations at 310 K. EXNS was evaluated by equation 1 with R = 8 Å.