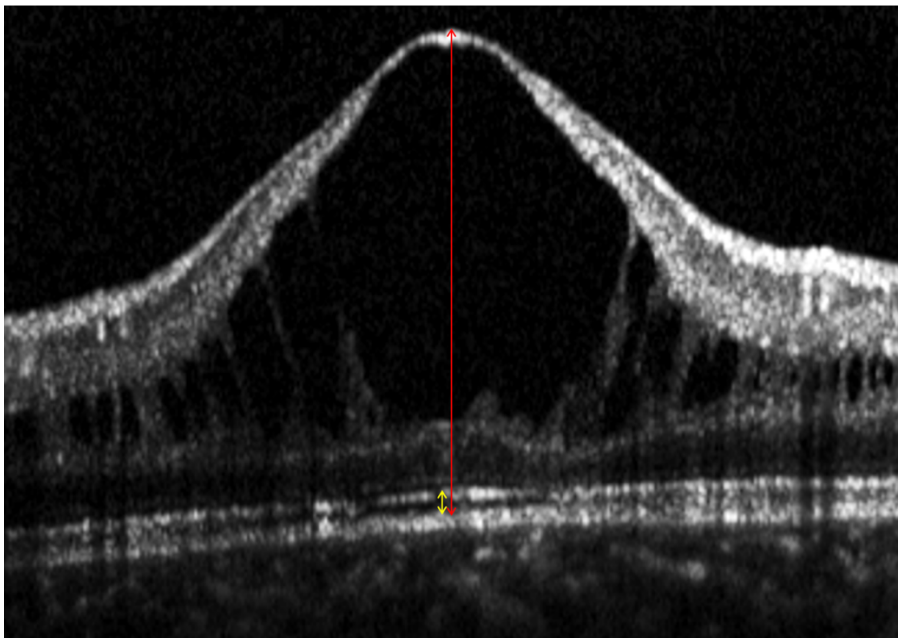
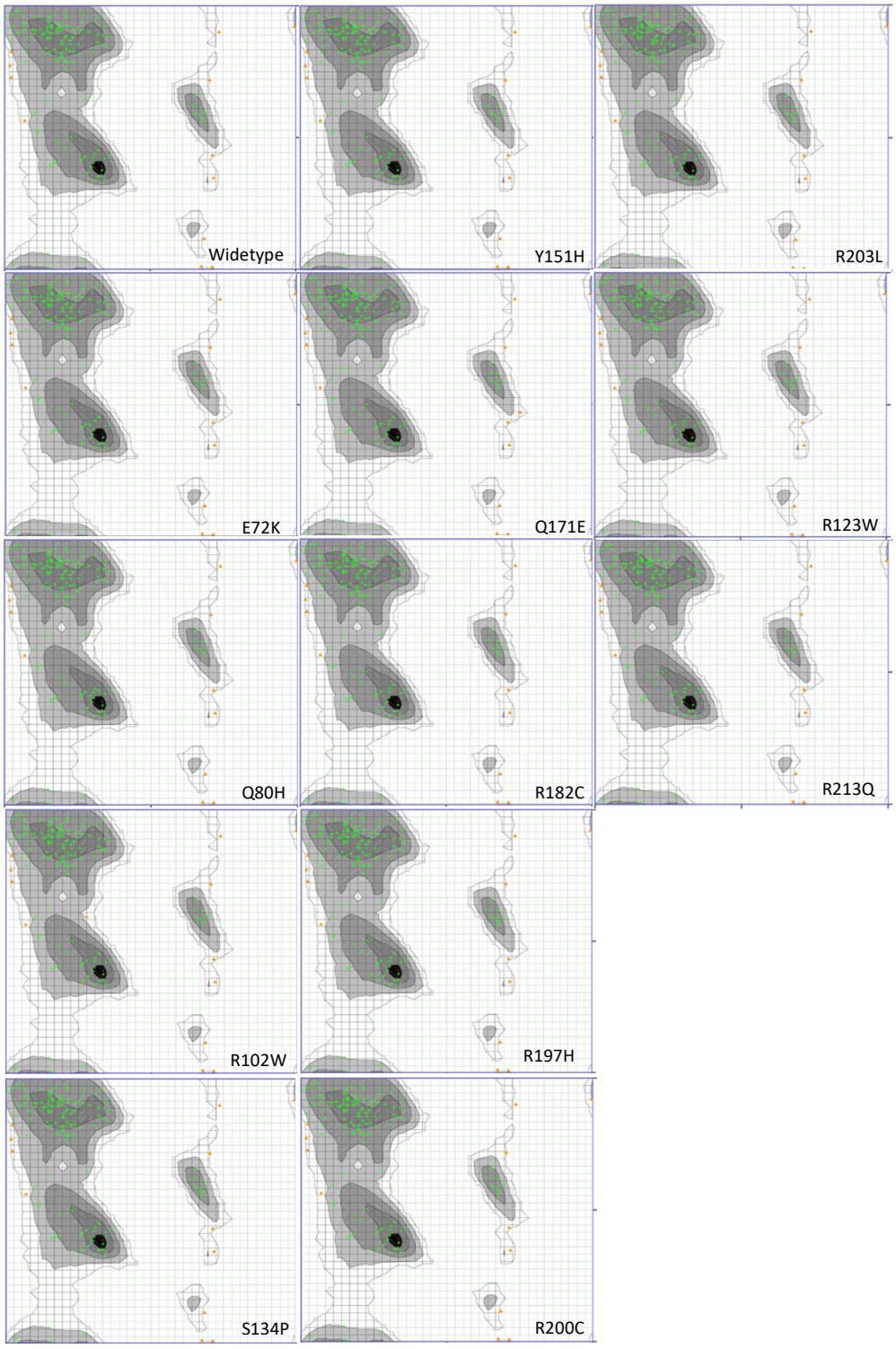
Supplementary File 1.

MD simulations setup

In the first step of the MD simulation, the protein molecules were solvated within a TIP3P cubic water model that employed explicit periodic conditions. To neutralize the charges of the systems, chloride or sodium ions were added. The second step of energy minimization aimed to remove any unfavorable contacts from solvation by performing a minimization procedure that consisted of 5,000 steps of steepest-decent-based and 10,000 steps of conjugated-gradient-based minimization until the energy of the system was reduced to less than 100 kJ/mol/nm. In the third step of equilibration, the systems were equilibrated at a constant temperature of 300 K using a two-step ensemble process of NVT (constant number, volume, and temperature) and NPT (constant number, volume, and temperature) for a duration of 100 ps. The NVT ensemble employed the Berendsen temperature coupling with no pressure coupling, while the pressure was isotopically controlled through a Parrinello-Rahman method at a reference pressure of 1 bar in the NPT ensemble. The calculation of long-range electrostatic interactions was performed using the Particle Mesh Ewald (PME) method. In the fourth step of production simulation, the systems underwent an unconstrained dynamics simulation with a 2 fs integration time step, utilizing the leapfrog method as the integrator algorithm. The structures generated during the 10 ns long simulation were recorded every 100 ps and outputted as trajectory files.

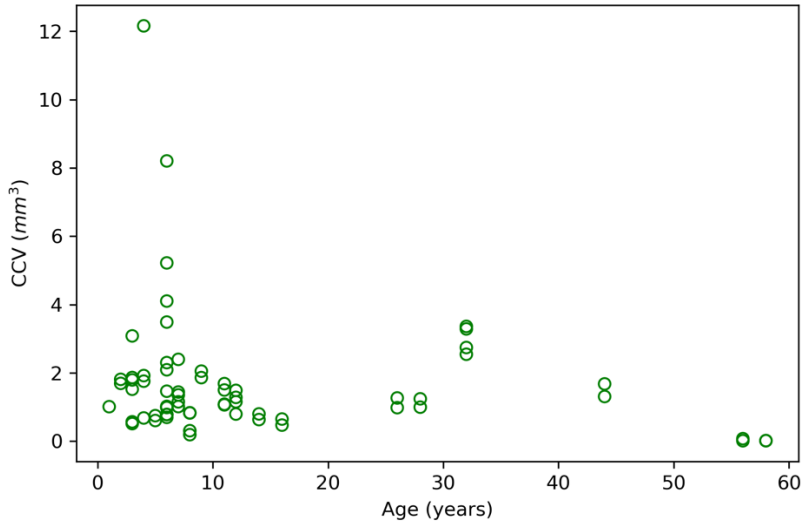


Supplementary Figure 1. Representative image showing the measurement of the central foveal thickness (CFT, red line) and the length of photoreceptor outer segment (PROS, yellow line) in OCT of patient with XLRS.

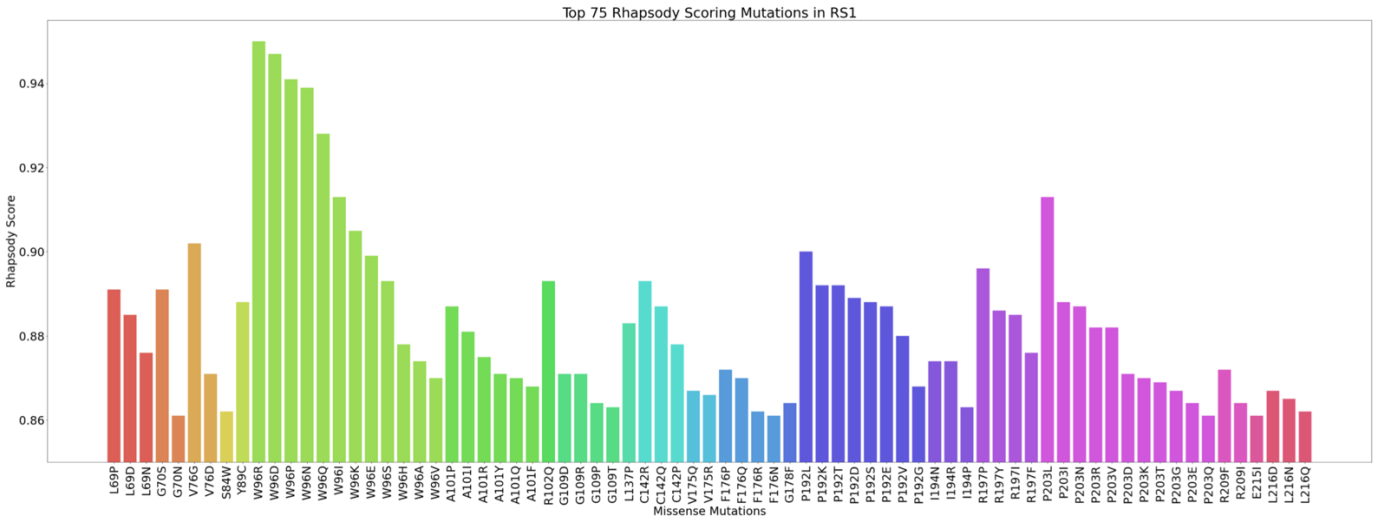


Supplementary Figure 2. Ramachandran plot for the structures of wildtype protein and mutant protein. The highly preferred observations, preferred observations, and questionable observations are shown as green crosses, brown triangles, red circles, respectively. The percentage of the highly

preferred observations, preferred observations, and questionable observations in Ramachandran plot for wildtype protein and mutant protein are summaries in Supplementary Table 1. We found that questionable observation is zero for all structures.



Supplementary Figure 3: Scatter plot of CCV against age for 37 individuals with XLRS. Each data point represents the CCV of an individual patient.



Supplementary Figure 4. Top Rhapsody scoring of residue substitutions variants. The 3D structure of the protein is formulated based on the homology model of RS1(PDB:3JD6). Wildtype residues of top-scoring (>0.86) are illustrated in the bar chart. The variants on same location are colored by identical color. Locations with more than four variants in the top 75 scores include: 96, 203, 192, 197, 101, 176, 109.

Supplementary Table 1. Results of Ramachandran plot for wildtype protein and mutant protein

Variants	Highly preferred observations	Preferred observations	Questionable observations
Wildtype	91.837%	8.163%	0.000%
E72K	91.837%	8.163%	0.000%
Q80H	91.837%	8.163%	0.000%
R102W	91.156%	8.844%	0.000%
S134P	91.781%	8.219%	0.000%
Y151H	91.837%	8.163%	0.000%
Q171E	91.216%	8.784%	0.000%
R182C	91.837%	8.163%	0.000%
R197H	91.837%	8.163%	0.000%
R200C	91.837%	8.163%	0.000%
P203L	91.892%	8.108%	0.000%
R213W	91.837%	8.163%	0.000%
R213Q	91.837%	8.163%	0.000%

Supplementary Table 3. In silico analysis of missense variants

Protein	p.E72K	p.Q80H	p.R102W	p.S134P	p.Y151H	p.G171E	p.R182C	p.R197H	p.R200C	p.P203L	p.R213W	p.R213Q
SIFT	0.264(T)	0.133(T)	0.005(D)	0.007(D)	0(D)	0.002(D)	0.003(D)	0(D)	0(D)	0.023(D)	0(D)	0(D)
Polyphen2 HDIV	1(PD)	1(PD)	1(PD)	1(PD)	0.998(PD)	1(PD)	1(PD)	1(PD)	1(PD)	1(PD)	1(PD)	1(PD)
Polyphen2 HVAR	0.994(PD)	0.999(PD)	1(PD)	0.999(PD)	0.948(PD)	1(PD)	0.999(PD)	0.999(PD)	0.999(PD)	0.999(PD)	1(PD)	0.998(PD)
Mutation Taster	1(DC)	1(DC)	1(DCA)	1(DC)	1(DC)	1(DC)	1(DC)	1(DC)	1(DC)	1(DCA)	1(DC)	1(DC)
Mutation Assessor	2.22(M)	3.08(M)	3.715(H)	3.3(M)	3.875(M)	2.295(M)	3.075(M)	3.65(H)	3.925(H)	3.755(H)	4.53(H)	3.98(H)
FATHMM	-5.4(D)	-5.54(D)	-4.99(D)	-4.91(D)	-5.71(D)	-5.31(D)	-4.81(D)	-6.56(D)	-6.75(D)	-5.61(D)	-5.04(D)	-5.06(D)
PROVEAN	-3.65(D)	-3.9(D)	-7.77(D)	-3.06(D)	-4.47(D)	-4.36(D)	-3.86(D)	-4.88(D)	-7.87(D)	-8.47(D)	-7.87(D)	-3.93(D)
VEST3	0.948(D)	0.82(D)	0.992(D)	0.993(D)	0.969(D)	0.901(D)	0.936(D)	0.977(D)	0.997(D)	0.984(D)	0.992(D)	0.984(D)
MetaSVM	1.113(D)	1.079(D)	1.079(D)	1.098(D)	1.014(D)	1.097(D)	1.096(D)	0.967(D)	0.979(D)	1.026(D)	1.052(D)	1.043(D)
MetaLR	0.972(D)	0.971(D)	0.972(D)	0.967(D)	0.986(D)	0.971(D)	0.964(D)	0.992(D)	0.993(D)	0.986(D)	0.979(D)	0.982(D)
M CAP	0.715(D)	0.647(D)	0.824(D)	0.939(D)	0.977(D)	0.826(D)	0.892(D)	0.966(D)	0.945(D)	0.973(D)	0.95(D)	0.97(D)
CADD	34(D)	24.3(D)	32(D)	27.6(D)	26.8(D)	31(D)	30(D)	35(D)	34(D)	33(D)	34(D)	35(D)
DANN	0.999(D)	0.997(D)	0.999(D)	0.999(D)	0.999(D)	0.998(D)	0.999(D)	1(D)	0.999(D)	0.999(D)	0.999(D)	1(D)
FATHMM MKL	0.98(D)	0.963(D)	0.909(D)	0.981(D)	0.981(D)	0.972(D)	0.917(D)	0.981(D)	0.96(D)	0.99(D)	0.927(D)	0.981(D)
Geno Canyon	1(D)	1(D)	0.999(T)	1(D)	1(D)	1(D)	0.996(T)	1(D)	1(D)	1(D)	1(D)	1(D)
REVEL	0.921(D)	0.848(D)	0.958(D)	0.952(D)	0.983(D)	0.819(D)	0.918(D)	0.984(D)	0.944(D)	0.981(D)	0.967(D)	0.985(D)
ReVe	0.985(D)	0.927(D)	0.999(D)	0.999(D)	0.996(D)	0.954(D)	0.981(D)	0.998(D)	0.999(D)	0.999(D)	0.999(D)	0.999(D)
GERP	5.43(C)	4.37(C)	4.56(C)	4.91(C)	4.91(C)	4.91(C)	2.54(C)	5.63(C)	3.7(C)	5.63(C)	4.68(C)	5.63(C)
phyloP	7.565(C)	4.718(C)	2.996(C)	8.652(C)	8.652(C)	7.32(C)	2.774(C)	7.568(C)	5.299(C)	9.602(C)	3.172(C)	7.568(C)
phastCons	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)	1(C)
SiPhy	18.318(C)	14.092(C)	12.706(C)	13.9(C)	13.9(C)	17.523(C)	13.903(C)	18.674(C)	12.473(C)	18.674(C)	14.368(C)	18.674(C)

D: Damaging; T: Tolerable; PD: Probably damaging; DC: Disease causing; DCA: Disease causing automatic; M: Medium; H: High; C: Conserved

Supplementary Table 4. In silico analysis of small deletion, insertion, nonsense variants

Variant	protein change	Mutation Taster	PhyloP	PhastCons
c.96delC	p.W33Gfs*93	1 (DC)	2.898 (C)	1 (C)
c.223G>T	p.E75X	2 (DC)	1.391 (C)	0.973 (C)
c.267T>A	p.Y89X	3 (DC)	1.339 (C)	1 (C)
c.335G>A	p.W112X	4 (DC)	5.324 (C)	1 (C)
c.336G>A	p.W112X	5 (DC)	3.719 (C)	1 (C)
c.336_337insT	p.L113Sfs*8	6 (DC)	3.719 (C)	1 (C)
c.489delG	p.W163X	7 (DC)	0.553 (C)	1 (C)
c.579delC	p.I194Sfs*43	8 (DC)	4.501 (C)	1 (C)

DC: Disease causing; C: Conserved

Supplementary Table 5: In silico analysis of splicing variants.

Variants	SpliceAI			HSF Pro
	Δ type	Δ score	pre-mRNA position	Predict Impact
c.52G>A	Acceptor Loss	0		• Broken WT Donor Site : Alteration of the WT Donor site, most probably affecting splicing (HSF)
	Donor Loss	0.31	0 bp	
	Acceptor Gain	0		
	Donor Gain	0		
c.53-1G>A	Acceptor Loss	0.69	-1 bp	• Broken WT Acceptor Site : Alteration of the WT Acceptor site, most probably affecting splicing (HSF) • Broken WT Acceptor Site : Alteration of the WT Acceptor site, most probably affecting splicing (MaxEnt)
	Donor Loss	0.32	-26 bp	
	Acceptor Gain	0.02	-27 bp	
	Donor Gain	0		
c.327-2A>G	Acceptor Loss	0.99	-2 bp	• Broken WT Acceptor Site : Alteration of the WT Acceptor site, most probably affecting splicing (HSF) • Broken WT Acceptor Site : Alteration of the WT Acceptor site, most probably affecting splicing (MaxEnt)
	Donor Loss	0		
	Acceptor Gain	0.36	44 bp	
	Donor Gain	0.01	14 bp	