Structure network-based landscape of rhodopsin misfolding by mutations and algorithmic prediction of small chaperone action

Angelo Felline [a], Davide Schiroli [b], Antonella Comitato [b], Valeria Marigo [b,c], Francesca Fanelli [a,c,*]

[a] Department of Life Sciences, University of Modena and Reggio Emilia, via Campi 103, 41125 Modena, Italy
[b] Department of Life Sciences, University of Modena and Reggio Emilia, via Campi 287, 41125 Modena, Italy
[c] Center for Neuroscience and Neurotechnology, Italy

ARTICLE INFO

ABSTRACT

Failure of a protein to achieve its functional structural state and normal cellular location contributes to the etiology and pathology of heritable human conformational diseases. The autosomal dominant form of retinitis pigmentosa (adRP) is an incurable blindness largely linked to mutations of the membrane protein rod opsin. While the mechanisms underlying the noxious effects of the mutated protein are not completely understood, a common feature is the functional protein conformational loss. Here, the wild type and 39 adRP rod opsin mutants were subjected to mechanical unfolding simulations coupled to the graph theory-based protein structure network analysis.

A robust computational model was inferred and *in vitro* validated in its ability to predict endoplasmic reticulum retention of adRP mutants, a feature linked to the mutation-caused misfolding. The structure-based approach could also infer the structural determinants of small chaperone action on misfolded protein mutants with therapeutic implications.

The approach is exportable to conformational diseases linked to missense mutations in any membrane protein.

© 2021 Published by Elsevier B.V. on behalf of Research Network of Computational and Structural Biotechnology. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

A majority of the over 16,000 missense mutations linked to human diseases affect folding or trafficking, rather than specific protein function [1]. Many conformational disease-linked mutations occur in integral membrane proteins. Failure by a protein to achieve its functional structural state and normal cellular location is a common contributor to the etiology and pathology of heritable human conformational diseases. Yet, the molecular mechanisms by which single amino acid changes lead to disease phenotypes are largely unknown [1].

Defects in protein folding or trafficking are at the basis of the pathogenicity of some mutations associated with retinitis pigmentosa (RP), a group of debilitating hereditary diseases causing the progressive degeneration of photoreceptor cells [2] and visual impairment in an estimate of 1.5 million patients worldwide [3]. RP affects about 1 in 3000–7000 people, being the most common cause of inherited blindness in developed countries [3–5]. More

than 70 genes have been associated with RP, and it exhibits extreme heterogeneity in terms of severity and mode of inheritance. Most of these genes are expressed specifically in photoreceptor cells and some of them encode key components of the signaling pathways of the G protein coupled receptor (GPCR) rhodopsin, the visual pigment molecule of rod cells that generates a detectable electrical response following light capture. Rhodopsin is composed of a protein opsin and a chromophore, 11-*cis*-retinal, forming a protonated Schiff base (PSB) with the amino acid residue K296 (reviewed in Refs. [6–9]). Like all GPCRs, rod opsin, encoded by the *RHO* gene, is an up-down bundle of seven transmembrane helices (H) linked to three intracellular (I) and three extracellular (E) loops as well as to an extracellular N-term and an intracellular C-term. Properly folded rod opsin is transported from the inner segment to the membranous disks of the rod outer segment, where it plays its function. In those cellular compartments the extracellular regions of the protein are, indeed, intradiscal.

Most of the over 140 missense mutations found in the *RHO* gene are linked to the autosomal dominant form of RP (adRP) [10]. adRP *RHO* mutants have been subjected to extensive biochemical and cellular characterizations, leading to a classification into three major groups (I, IIa/II, and IIb/III) (reviewed in [10,11]), based on

---

* Corresponding author at: Department of Life Sciences, University of Modena and Reggio Emilia, via Campi 103, 41125 Modena, Italy.
E-mail address: fanelli@unimo.it (F. Fanelli).

rod opsin ability to reconstitute with 11-*cis*-retinal and to be retained into the endoplasmic reticulum (ER) [11]. In detail, class I mutants resemble wild type (WT) in ability to regenerate with 11-*cis*-retinal and subcellular localization. In contrast, class IIa (or II in ref. [11]) and class IIb (or III in ref. [11]) mutants show altered membrane localization in their opsin state associated with ability or inability, respectively, to regenerate with 11-*cis*-retinal *in vitro* (Table S1).

Another classification of *RHO* mutants linked to adRP and autosomal dominant congenital stationary night blindness (adCSNB), which can also be linked to mutations in *RHO*, is based on a number of variegate effects on rod opsin function (Classes 1–7) [5,12]. The considered effects that found this alternative classification include: post-Golgi trafficking and outer segment targeting (Class 1), ER retention and impairment in 11-*cis*-retinal binding (Class 2), vesicular trafficking and endocytosis (Class 3), post-translational modifications (Class 4), Gt activation (Class 5), constitutive activation (Class 6), and efficiency to dimerize (Class 7) [5,12] (Table S1).

The majority of the adRP rod opsin mutants are likely misfolded and retained in the ER (reviewed in Refs. [10–12]). As a conformational disease [1,13,14], adRP linked to *RHO* mutations is amenable to treatment with pharmacological chaperones, small molecules that bind specific sites within a protein's native or quasi-native structure, thereby shifting the folding equilibrium towards the native state. Indeed, pharmacological therapies for these diseases are based on either promoting correct folding, inhibiting aggregation, increasing degradation, or protecting from cell death [5,15].

Over the last years, we developed and implemented computational strategies to infer the structural determinants of adRP linked to *RHO* mutations [16–18]. In a recent study, computational experiments were combined with *in vitro* analyses to characterize the effects of 33 adRP *RHO* mutations on stability and subcellular localization of the protein in the absence and presence of the retinal chromophore [18]. Computational experiments consisted in thermal unfolding molecular dynamics (MD) simulations coupled with the protein structure network (PSN) analysis, a cutting-edge graph theory-based approach [19], which we implemented in the free software Wordom and in a webserver [20–22]. Indeed, concepts and methods borrowed from the graph theory are being increasingly used to study several aspects of structural biochemistry and medicinal chemistry [16,18,23–33]. The working hypothesis motivating previous and present computational experiments was that adRP mutants classified as misfolded indeed represent "quasi-native" states of the opsin protein. The integration of atomic and subcellular levels of analysis was accomplished by the linear correlation between indices of mutational impairment in structure network and in subcellular localization. In that respect, estimation of ER retention was based on the Pearson correlation coefficient (PCC) of opsin/isorhodopsin co-localization with the ER-resident calnexin. The graph-based index of structural perturbation served also to divide the mutants in four clusters, consistent with their differences in subcellular localization and responses to the chaperone activity of retinal [18].

The stability core of the protein inferred from the PSN analysis [18] consistently with earlier predictions [34] was targeted by virtual screening of over 300,000 anionic compounds leading to the discovery of a novel chaperone, the 13-*cis*-5,8-epoxy-retinoic acid (13-*cis*-5,8-ERA), able to bind bovine rod opsin with 14-fold lower EC_{50} (8.6 ± 0.2 nM) than 9-*cis*-retinal (123.9 ± 7.0 nM). Acting as a reversible orthosteric inhibitor of retinal binding, the compound proved more effective than 9-*cis*-retinal in mitigating ER retention and promoting membrane localization of three adRP *RHO* mutants, i.e. T17M, P23H, and E181K [18]. Other two studies reported on the discovery of small chaperones able to bind bovine rod opsin and to promote membrane localization of P23H opsin though with EC_{50}

comparable or worse than that of 9-*cis*-retinal [35,36]. Specifically: i) the discovered small chaperones are non-retinoids; ii) the compound YC-001, discovered by cell-based high-throughput screening, showed inverse agonist and non-competitive antagonist activities towards rod opsin [35]; and iii) the compound S-RS1 inferred from virtual and thermofluor screening was found to stabilize rod opsin by binding to the orthosteric side [36].

In the present investigation, the wild type (WT) and 39 adRP *RHO* mutants, comprising a set of 33 mutants with previously determined ER retention [18] (hereafter indicated as a training set) and a set of 6 mutants with unknown ER retention (hereafter indicated as new mutants because herein we determined their subcellular localization for the first time) were subjected to mechanical unfolding simulations (steered molecular dynamics (SMD) simulations) under the effects of small chaperones like 11-*cis*-retinal (i.e. rhodopsin state) and 13-*cis*-5,8-ERA (the latter probed only on the P23H mutant). SMD simulations were coupled to PSN analysis.

The advantage of using SMD over thermal unfolding simulations is that predictions of the stability points are more robust and the computational approach can be more easily automated. Here we present a computational approach useful to estimate responsiveness of adRP rod opsin mutants to small chaperones like 11-*cis*-retinal and to aid discovery/design of novel chaperones with therapeutic potential.

2. Materials and methods

2.1. Computational experiments

2.1.1. Molecular simulations

The molecular system and protocol employed for SMD simulations have been previously described [16]. Incidentally, 14 out of the 20 mutants previously studied also participate in the set of 39 mutants considered herein, each of them being subjected to a number of new replica simulations in the present study. In detail, the molecular system subjected to SMD simulations was a rhodopsin nonamer resulting from the addition of one three-rhodopsin layer to the semiempirical rhodopsin oligomer released in the protein databank (PDB) as 1N3M and replacement of each monomer with the complete 1U19 structure [37,38]. According to that model, two monomers of rhodopsin make contacts mediated by E2, I2, H4, and H5 (Fig. 1A). The semiempirical oligomeric model of rhodopsin was consistent with the paracrystalline organization of the pigment in disk membranes as inferred from atomic force microscopy [37,39]. The recent cryoEM structures of a cross-linked rhodopsin dimer and of a rhodopsin dimer reconstituted into nanodiscs from purified monomers show a dimer interface mediated by H1-H1 and H8-H8 contacts (PDB: 6OFJ [40]). An oligomer was inferred, made of rows of dimers held together by E2-I3 and H4-H6,H5 contacts (Fig. 1B). To probe also the new oligomeric model, we substituted the 1U19 structure deprived of the C-tail, which would generate steric clashes, onto each monomer in the new hexameric arrangement. In the present investigation both the previous nonameric and recent hexameric models of rhodopsin were employed as inputs of SMD simulations.

Pulling was applied to the central rhodopsin monomer of the nonameric assembly and to one monomer in the central row of the hexameric assembly (Fig. 1A,B). In SMD simulations, the force applied between two selected atoms is proportional to the difference between the distance of the two atoms and a linearly increasing length. As in previous SMD simulations of adRP rod opsin mutants [16], the force was applied between the Cα-atom of the first amino acid of rhodopsin (in the intradiscal side) and a fixed dummy atom, placed 35 nm from the putative membrane surface
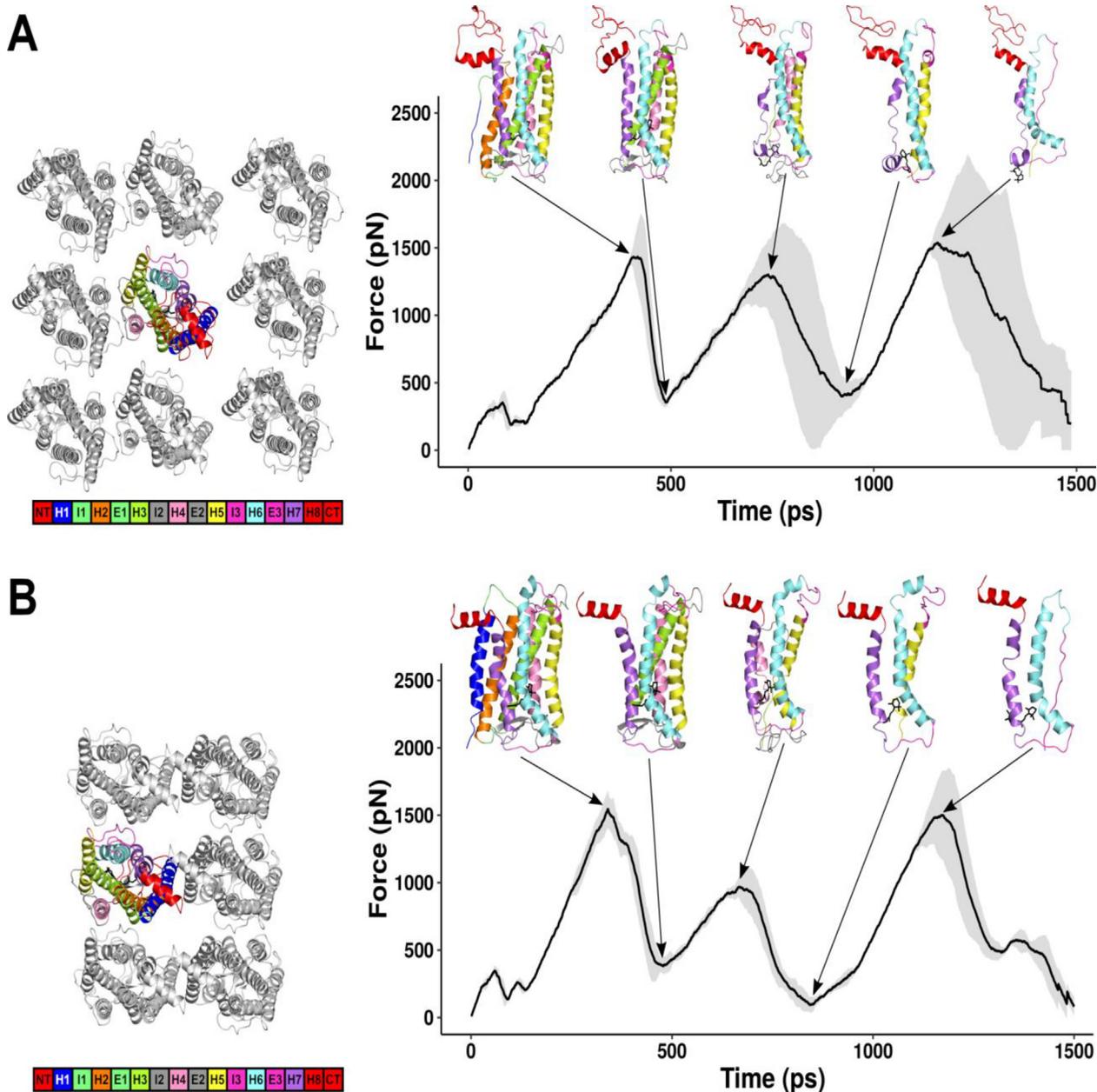
Fig. 1. SMD simulations of WT rhodopsin. (A) In the left panel, the dark rhodopsin nonamer subjected to SMD simulations is shown, each monomer being the crystal structure encoded as 1U19. Pulled rhodopsin is the central monomer, i.e. the one colored as follows: H1, H2, H3, H4, H5, H6, and H7 are blue, orange, green, pink, yellow, aquamarine, and violet, respectively. H8, the N- and C-termini are red, E1 and I1 are light green, E2 and I2 are gray, and E3 and I3 are magenta (see the color-legend bar). The 11-*cis*-retinal is shown as black sticks. The oligomer is seen from the intracellular side in a direction perpendicular to the membrane plane. In the right panel, the average force/time profile of the five most similar trajectories is shown with the standard deviation represented as gray areas. Cartoons of the frames corresponding to the three main force peaks and to the valleys after the first two main force peaks are shown and colored as described above. (B) The description of the left and right panels here is the same as the one of panel A but referred to SMD simulations of the WT form in the hexameric arrangement. In this case, pulled rhodopsin is one of the two monomers in the central row, i.e. the one colored according to the different receptor portions (see the color-legend bar). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

in the cytosolic side. A dummy atom had been already employed also in our previous simulations of bacterioopsin pulling to keep the force perpendicular to the putative membrane plane [41].

SMD simulations were carried out on the WT and 39 adRP mutant rod opsins bound to 11-*cis*-retinal and on P23H mutant rod opsin bound to the recently discovered [18] small chaperone 13-*cis*-5,8-ERA in its 5R8S and 5S8R trans enantiomeric forms. The orientation of the 5R8S enantiomer in the input complex was the one previously inferred from virtual compound screening [18], whereas that of 5S8R was obtained upon fitting on the 5R8S isomer. Simulations were run in the absence of the native disul-

phide bridge between C110 (in H3) and C187 (in E2). This choice, already justified in the previous study [16], is due to the fact that the presence of the native disulphide bridge between the N-terminal end of H3 and the C-terminal strand of E2 greatly reduces the system's degrees of freedom masking also the unfolding/pulling process of helices 4–7. This choice is also justified by the evidence that the bridge forms after packing of the TM helices and of the intradiscal regions. Indeed, previous *in vitro* experiments showed that the C110A/C187A mutant resembles WT rod opsin in the ground state and folds correctly to bind 11-*cis*-retinal [42]. The stability of the mutant to thermal decay was found reduced,

relative to the WT, by a $\Delta\Delta G$ of only -2.9 kcal/mol [42]. Consistently, in some cases, adRP mutations that affect packing of helices and intradiscal regions may impede the formation of such bridge also favoring possible non-native bridges (reviewed in ref. [12]). In summary, the disulphide bridge likely serves to stabilize the final fold, which, however, forms independently of the bridge. Therefore, to properly infer the structure-network elements that likely participate in folding and are perturbed by mutation, the disulphide bridge must be absent during the mechanical unfolding process.

Residues D83 and E122 were kept in their protonated forms, according to spectroscopic determinations [43] and to the finding that the E122L/L125R double mutant improves the WT-like features of the receptor, as far as protein expression, glycosylation and trafficking are concerned [44]. E122 was simulated both in its protonated and deprotonated states only in the L125R, C167R, and H211R mutants, in which the closeness of the substituting arginine may lower the pK_a of the amino acid, thus favoring a charge-reinforced H-bond between glutamate and mutated side chain. Selected replacing anionic amino acids, i.e. D87, D89, and D114, were simulated in both protonation states. For those mutants characterized by torsion angles in the replaced side chain, up to three rotamers were probed by MD, corresponding to the Dunbrack and Karplus (D&K) [45], Ponder and Richards (P&R) [46], and Sutcliffe (Sut) [47] rotamer libraries.

Energy minimizations and SMD simulations were carried out by means of the CHARMM force field (in all-atom mode) [48], using the GBSW implicit membrane/water model [49]. For the 11-*cis*-retinal protonated Schiff's base, we employed the all-atom topologies and parameters introduced in the version c32b2 of the force field. The compound 13-*cis*-5,8-ERA was parametrized according to the CHARMM General Force Field. With respect to the physical parameters representing the membrane in the GBSW model, the surface tension coefficient (representing the non-polar solvation energy) was set to 0.03 kcal/(mol·Å^2). Membrane thickness (centered at Z = 0) was set to 35.0 Å with a membrane smoothing length of 5.0 Å.

Collectively, the use of supramolecular rhodopsin and of implicit water/membrane models is in line with earlier mechanical unfolding simulations on bacterioopsin, which showed also the independence of the force/distance profiles from the details of the starting packing interactions between pulled and surrounded monomers [41]. The stretching, bending, and torsional degrees of freedom of the non-pulled monomers were limited to the side chains in direct contact with the pulled monomer, leading to an implicit solvent/protein cage model as already described [16,41]. This model prevents artifacts such as anomalous rotation of the protein during the pulling process carried out on rhodopsin (this study) or bacterioopsin monomers in an implicit membrane/water system [16,41]. This strategy thus avoids violations of experimental observations.

Prior to SMD simulations, the system was energy minimized by 1500 steps of steepest descent followed by adopted basis Newton-Raphson (ABNR) minimization, until the root mean square gradient was less than 0.001 kcal/mol Å. The minimized coordinates were heated to 300 K with 7.5 K rises every 2.5 ps per 100 ps by randomly assigning velocities from a Gaussian distribution. After heating, the system was allowed to equilibrate for 500 ps. An integration step of 0.002 ps was used. SMD consisted in Langevin MD simulations with a friction coefficient of 0.2 ps^{-1}. The pulling force (with a constant equal to 100 pN/nm) was applied at a constant speed of 0.1 nm/ps, similar to previous SMD simulations on bacterioopsin [41]. Considering the WT and 39 mutants in complex with 11-*cis*-retinal, and P23H in complex with the two enantiomers of 13-*cis*-5,8-ERA, the different rotameric states of replacing amino acids, at least 10-simulation-replica for each molecular specie,

and two different input oligomeric arrangements, more than 1700 SMD simulations were carried out. It is worth nothing that, lacking the C-terminus, the hexameric model could not be used for the P347S C-terminal mutant, which was simulated only in the nonameric arrangement.

Since the frames that follow the second main force peak contain essentially noise rather than information, WT and mutant trajectories were cut at the frame corresponding to the minimum after the second main peak and subjected to PSN analysis (Figs. S1 and S2).

2.1.2. PSN analysis

PSN analyses were carried out by means of the Wordom software [21] on the WT and mutated forms of rhodopsin. The PSN implemented in Wordom, based on the approach described in relevant papers by Vishveshwara and co-workers [24,50], is a product of graph theory applied to protein structures. A graph is defined by a set of points (nodes) and connections (edges) between them. In a protein structure graph (PSG), each amino acid is represented as a node and these nodes are connected by edges based on the strength of non-covalent interactions between nodes [19]. The strength of interaction between residues i and j (I_{ij}) is evaluated as a percentage given by the following equation:

$$I_{ij} = \frac{n_{ij}}{\sqrt{N_i N_j}} \times 100$$

where I_{ij} is the percentage interaction between residues i and j; n_{ij} is the number of atom–atom pairs between residues i and j within a distance cutoff (4.5 Å); N_i and N_j are normalization factors for residue types i and j, which take into account the differences in size of the side chains of the different residue types and their propensity to make the maximum number of contacts with other amino acid residues in protein structures. The normalization factors for the 20 amino acids in the Wordom implementation were taken from the work by Kannan and Vishveshwara [51], whereas the normalization value for retinal, i.e. 163.4, was computed as the average number of contacts made by the molecule in a dataset of 408 structures from the PDB. The normalization factors of the two isomers of 13-*cis*-5,8-ERA (i.e. 5S8R = 103.06 and 5R8S = 118.67) were computed as previously described [18]. Glycines, excluded from PSN analysis in previous applications, were included in this study.

Thus, I_{ij} are calculated for all node pairs. At a given interaction strength cutoff, I_{min}, any residue pair ij for which $I_{ij} \geq I_{min}$ is considered to be interacting and hence is connected. Node interconnectivity is finally used to highlight cluster-forming nodes, where a cluster is a set of connected amino acids in a graph. Node clustering procedure is such that nodes are iteratively assigned to a cluster if they could establish a link with at least one node in such cluster. A node not linkable to existing clusters initiates a novel cluster and so on until the node list is exhausted. Cluster size, defined as the number of nodes in the cluster, varies as a function of the I_{min}, and the size of the largest cluster is used to calculate the I_{critic} value, which is worth using as an I_{min} cutoff for the analysis [24]. An analysis of 200 different proteins showed that, irrespective of the protein size or fold, the size of the largest cluster in each of the proteins undergoes a transition at a particular I_{min} value. This I_{min} value at which the size of the largest cluster decreases dramatically (i.e., the midpoint of the transition) was indeed the $I_{critical}$ [24]. The latter is therefore defined as the I_{min} at which the size of the largest cluster is half the size of the largest cluster at $I_{min} = 0.0\%$. We set the I_{min} equal to the I_{critic} approximated to the second decimal place. The I_{min} finally employed for the analysis is the average over all the I_{min} computed on each trajectory frame.

The residues making four or more edges are referred to as hubs at that particular I_{min}. The four-link cutoff relates to the intrinsic limit in the possible number of non-covalent connections made

by an amino acid in protein structures due to steric constraints and it is close to its upper limit. Most amino acid hubs indeed make from 4 to 6 links.

All PSGs were computed by setting a frequency cutoff equal to 25%, meaning that only those links and hubs present in at least 25% of the total trajectory frames were taken into account.

PSGs then served to quantify the differences in structure network between WT and mutants through the NP_{SMD} structure perturbation index described in the Result section. For each mutant in each of the two oligomeric assemblies employed as inputs, the index was computed on the average structure network obtained by considering only the 5 most similar (in link frequency) PSGs from the simulated trajectories. The final score of each mutant is the average of the scores relative to the two different oligomeric forms.

2.1.3. Analysis of amino acid conservation

Conservation analysis was performed by using the ConSurf web server (http://consurf.tau.ac.il) [52]. In a typical ConSurf application, the query protein is first BLASTed against the UNIREF-90 database. Redundant homologous sequences are then removed using the CD-HIT clustering method. The resulting sequences are next aligned and the generated multiple sequence alignment (MSA) is used to reconstruct a phylogenetic tree. Given the tree and the MSA, the Rate4Site algorithm is used to calculate position-specific evolutionary rates under an empirical Bayesian methodology. The rates are normalized and grouped into nine conservation grades 1-through-9, where 1 includes the most rapidly evolving positions, 5 includes positions of intermediate rates, and 9 includes the most evolutionarily conserved positions [52]. We employed the sequences in the crystal structures of dark rhodopsin (PDB: 1U19) as a query in default conditions, but changing the minimal and maximal identity percentages to 20% and 80%, respectively.

2.2. In vitro experiments

2.2.1. cDNA constructs and site directed mutagenesis

Site directed mutagenesis performed on human *RHO* cDNA [53] in the pRc/CMV expression plasmid (pRc/CMV-RHO) with a Quik-Change® Site-Directed Mutagenesis Kit and primers containing the desired base change. All mutagenized cDNAs were checked by sequencing.

Primers used were as followed:

A164E (5′-CTGGGTCATGG**A**GCTGGCCTGCGCC-3′);
Y178N (5′-GGCTGGTCCAGG**A**ACATCCCCGAGG-3′);
G188E (5′-CAGTGCTCGTGTG**A**AATCGACTACTAC-3′);
P171E (5′-GCGCCGCACCC**GA**ACTCGCCGGCTG-3′);
G182S (5′-CATCCCCGAG**TC**CCTGCAGTG-3′);
P347S (5′-CAGGTGGCC**T**CGGCCTAA-3′).

2.2.2. Rod opsin expression in COS-7 cells.

$2x10^4$ COS-7 cells were seeded on coverslip slides and transfected with 300 ng of WT or mutant pRc/CMV-RHO constructs by calcium phosphate transfection as described previously [18]. Twenty-four hours after transfection the culture medium was changed with fresh complete DMEM with or without 10 μM 9-*cis*-retinal or the two-isomer *trans*-mixture of 13-*cis*-5,8-ERA. We used 9-*cis*-retinal and not 11-*cis*-retinal because it is chemically more stable and we previously demonstrated that has similar effects on rod opsin expressed in the COS-7 system [18]. No significant differences in transfection efficiencies were observed with different constructs (WT or mutants). Data for ER- and plasma membrane-localization in the presence and absence of 9-*cis*-retinal or the 13-*cis*-5,8-ERA racemic mixture derived from at least three biological replicates.

2.2.3. Immunohistochemistry.

Forty-eight hours after transfection cells were fixed with 2 % paraformaldehyde in PBS for ten minutes. For co-localization analysis fixed cells were permeabilized and blocked in PBS with 3% BSA and 0.1 % TritonX-100 for 1 h at room temperature, washed five times with PBS and incubated with the monoclonal mouse anti-rhodopsin antibody 1D4 (epitope at the C-terminal of rod opsin; Sigma, #R5403; 1:1000) together with the polyclonal rabbit anti-calnexin antibody (CLNX, H-70: sc-11397, Santa Cruz; 1:25) for 1 h at room temperature. To analyze the membrane localization of rhodopsin, fixed cells were blocked in PBS with 3% BSA without detergent to avoid permeabilization of the cell membrane and incubated with the primary mouse anti-rhodopsin antibody RetP1 (epitope at the N-terminal of rod opsin; Abcam, ab3267; 1:10000). As secondary antibodies we used Alexa Flour 568 goat anti-mouse IgG (Life Technologies; 1:1000) and Oregon-Green 488 goat anti-rabbit IgG (Invitrogen; 1:1000) incubated with 0.1 μg/mL DAPI for nuclear staining. Slides were mounted with Mowiol 4–88 (Sigma) and analyzed with a Zeiss Axio Imager A2 microscope.

2.2.4. In vitro experiments: Digital image analysis.

Image processing, including measurements of the PCC and cell-counting was conducted using the open source ImageJ v1.48 software. Co-localization of the 1D4 and CLNX immunofluorescence was evaluated by calculating the PCC, which determines the relative fluorescence intensities of the green Oregon-Green 488 (labeling CLNX) and red Alexa Flour 568 (labeling rod opsin) in the same groups of pixels in a region of interest (ROI) [54]. The PCC-formula for an image consisting of a red and green channel is given below:

$$PCC = \frac{\sum_i \left(R_i - \bar{R}\right) \times \left(G_i - \bar{G}\right)}{\sqrt{\left(R_i - \bar{R}\right) \times \left(G_i - \bar{G}\right)^2}}$$

R_i and G_i are the intensity values of pixels and $\bar{R}$ and $\bar{G}$ the mean intensities of the red and green channels across the entire ROI respectively.

The channels of single RGB images were split into grayscale pictures and the red and green channels used for analysis. An ROI in the red channel was created by free hand drawing around the fluorescence signal labeling RHO in single cells and used for calculating PCC (see Ref. [54] for details). PCC values can range from 1 to −1. Whereas a value of 1 represents perfect correlation, value −1 represents perfect but inverted correlation and values near zero represent distributions of fluorescent signals that are uncorrelated with one another [54].

The percentage of cells expressing WT or mutant rod opsin was determined by counting the number of 1D4[+] cells versus the number of DAPI-stained cells. To evaluate the percentage of cells with WT or mutant rod opsin at the cell surface the same procedure was performed with images taken from cells immunostained with RetP1 without membrane permeabilization. For each experiment, the percentage of RetP1[+] divided by the percentage of 1D4[+] cells (RetP1/1D4) was calculated. The mean PCC- and RetP1/1D4-values of at least three independent experiments were used for statistical analysis by Student's unpaired *t* test. All data are presented as mean values ± standard errors of the means (SEMs).

For the six mutants analyzed in this study, we analyzed 30 cells per mutant in the presence or absence of 9-*cis*-retinal or the two-isomer *trans*-mixture of 13-*cis*-5,8-ERA.

3. Results

3.1. A computational model predicts the effects of small chaperones like 11-cis-retinal on the subcellular localization of misfolded adRP RHO mutants

In the present investigation, those 33 adRP mutants with previously determined ER retention, i.e. accounted for by the PCC index for the co-localization of rod opsin and the ER-resident calnexin (Table S1, see Methods) [18], were used as a training set for algorithmic prediction of ER retention of new adRP mutants under the effect of small chaperones like 11-cis-retinal. The computational model relied on SMD simulations coupled to PSN analysis. SMD simulations on six adRP RHO mutants in their rhodopsin state served to validate the model (Table S1).

Pulling was applied to a rhodopsin monomer in the central row of the nonameric and hexameric assemblies (see Methods, Fig. 1A,B). As already found in a previous study [16], mechanical unfolding of WT and mutant rhodopsin is strongly dictated by the membrane topology of the protein (Fig. 1B,A and Figs. S1 and S2). Indeed, both the regular spacing between the three main force peaks and the relative heights (i.e. local maxima of the force) are consistent with the up-down architecture of the seven-helix bundle and the direction of pulling. In this respect, during the extraction from the helix-bundle (i.e. mechanical unfolding from the N-terminus), odd helices are pulled in a direction that allows them to unfold by helical stretching. On the contrary, even helices cannot stretch during pulling because the pulling direction is opposite to the helical stretch direction; thus, even helices have to flip by 180° before leaving the helix bundle, and this causes the main resistance to pulling. This occurs independently of the receptor form (i.e. WT or mutant) and the oligomeric model (i.e. nonameric or hexameric) used as an input of simulations (Fig. 1).

The first main peak corresponds to a structural state, in which the entire H1 is completely or minimally unfolded, for the nonameric and hexameric models respectively, whereas the remaining α-helices retain their original conformation. The only difference between the unfolding simulations of the two oligomeric models is that, differently from the nonameric arrangement, in the hexameric arrangement H1 contributes to the first main force peak, as it participates in the dimer interface characterized by H1-H1 contacts (Fig. 1B). The valley after the first main peak is characterized by the pulling of the H1-I1-H2 regions out from the bundle. At the second main peak, also the entire H3 and, only for the nonameric model, the central portion of H7 (originally in 3.10 conformation) are unfolded, whereas H4, H5, H6 and H8 almost retain their original α-helical structure. The valley after such main peak is characterized by pulling of the unfolded H3-I2-H4 regions out from the bundle. The protein is almost entirely unfolded at the third main peak, where only H6, H7 (in part), and H8 retain their α-helix conformation (Fig. 1A,B).

Almost all native interactions are lost after the first two main force peaks, therefore, PSGs were essentially based on the trajectory frames that precede the second main valley in the force/time profile, where about 60% of the rhodopsin structure is unfolded and pulled in the extracellular side (Fig. 1). Collectively, the dimer interface in the hexameric model makes H1 and H7 more stable and resistant to misfolding than the dimer interface in the nonameric model (Fig. 1).

The extensive PSN analysis of more than 1500 SMD trajectories of WT and 33 adRP training-set mutants highlighted the native (i.e. characterizing the WT form) stable (i.e. persisting in at least 25% of the trajectory frames) hubs as the proper framework to infer the structural perturbation induced by adRP rod opsin mutations. Indeed, WT hubs and hub-links found

in ≥ 25% of frames turned out to mark the folded state of rhodopsin (Fig. 2A-C).

Comparison of the native hub-links shows substantial similarities between the two oligomeric models. Indeed, 63% of nodes are shared by the two networks (Fig. 2A). The receptor portions holding the majority of native hubs are H3, H5, H6, H7, and, above all, E2 (Fig. 2B-C). Such hubs include sites susceptible to spontaneous pathogenic mutation (i.e. M44, R135, P171, Y178, P180, E181, D190, M207, H211, and K296), 50% of which lie in E2 (Fig. 2A-C). Remarkably, 51% (19 out of 37) of native hubs falling in the helices are highly evolutionary conserved amino acids (i.e. according to ConSurf see Methods). The chromophore 11-cis-retinal is the most persistent hub (Fig. 2), which in turn is linked to seven stable hubs: E113 (in H3, ConSurf color: 5 (H3, 5)), E122 (H3, 6), I189 (E2, 5) in the nonameric model, replaced by F208 (H5, 5) in the hexameric model, M207 (H5, 4), W265 (H6, 8), Y268 (H6, 6), and K297 (H7, 6) (Fig. 2C, D). Only one of such hubs, W265, is evolutionary conserved (i.e. holds a ConSurf value ≥ 7). In addition to those hubs, the 11-cis-retinal is also linked to A117, T118 (both in H3), and F212 (H5) (Fig. 2C,D).

The ten nodes persistently linked to 11-cis-retinal are hereafter defined as retinal binding pocket (RBP). In general, those RBP links involving hubs tend to be persistent and stronger (according to the interaction strength) than the others (Fig. 2C,D).

For the training set of 33 adRP RHO mutants (Table S1 and Fig. 3A,B), a number of index formulations quantifying mutation-induced perturbations of the native stable structure network of rhodopsin were automatically tested, which resulted in linear correlations with the in vitro determined PCC index of ER retention. Those correlations are such that the higher the network-perturbation index, the higher the weakening (i.e. reduction in frequency compared to the WT) of the native hubs and their links by mutation, the higher ER retention of the rhodopsin mutant is.

The best performer index in that respect (NP$_{SMD}$, Fig. 3B) is formulated here below:

$$NP_{SMD} = \frac{\frac{Mut_{Pert}^{Hubs}}{WT_{Tot}^{Hubs}} + \frac{Mut_{Pert}^{HLinks}}{WT_{Tot}^{HLinks}}}{2} 100$$

where WT_{Tot}^{Hubs} and WT_{Tot}^{HLinks} are the total number of native hubs and their links in the WT average network, respectively, while WT_{Pert}^{Hubs} and WT_{Pert}^{HLinks} are the total number of native hubs and their links with a reduced frequency in the mutant average network. The NP$_{SMD}$ score employed in the correlative model shown in Fig. 3 is the average of the scores relative to the nonameric and hexameric models (see also Methods).

In summary, extensive SMD simulations and PSN analysis on WT rhodopsin and a training set of 33 adRP-linked RHO mutants with known degree of ER retention [18] led to the building of a computational structural model relying on the linear correlation between the NP$_{SMD}$ index of structure perturbation and the PCC index of ER retention.

The model highlighted native hubs and their links as the targets of the structural perturbations by misfolding mutations.

3.2. In vitro experiments validate algorithmic predictions

The linear regression model between NP$_{SMD}$ and PCC (see Fig. 3 legend) was used to predict the effects of 11-cis-retinal on ER retention of the six new mutants not included in the training set (i.e. A164E, P171E, Y178N, G182S, G188E, and P347S) but simulated and analyzed like the training set (Table S1).

Alternative replacements of A164, P171, Y178, and G188 were already present in the training set (i.e. A164V, P171Q, Y178C, and G188R). We specifically chose those adRP RHO mutants to
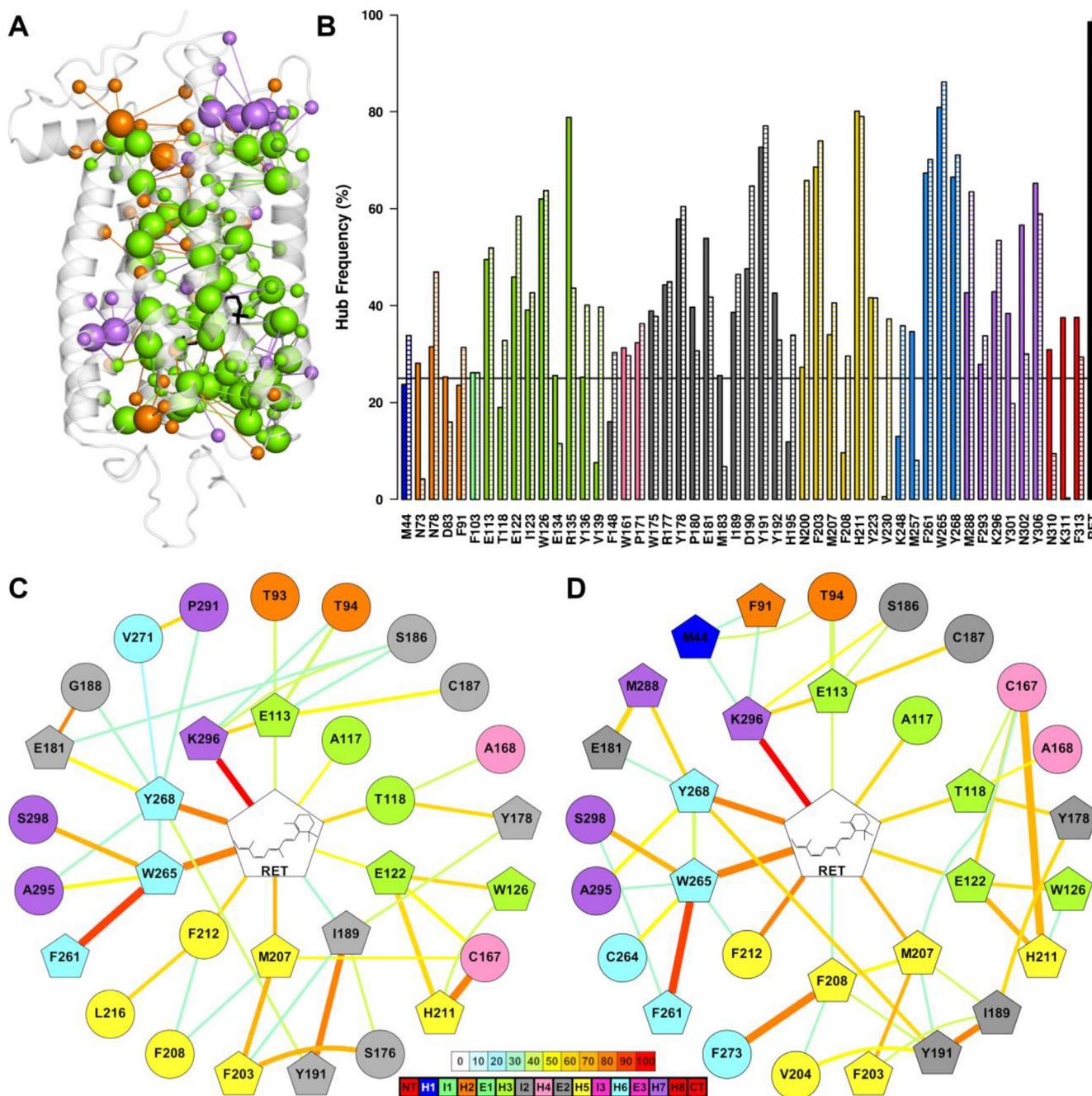
Fig. 2. Native stable hubs and their links from SMD simulations of dark rhodopsin. (A) The difference PSG made by the native stable hubs and hub-links in the nonameric and hexameric models is shown on the crystal structure of dark rhodopsin (PDB: 1U19). Common nodes or links are green, nodes or links peculiar to the nonameric model are orange, and nodes or links peculiar to the hexameric model are violet. The big spheres (centered on the Cα-atoms) are the hubs, whereas the small spheres are the nodes linked to such hubs. (B) The histograms represent the frequency of each native hub and are colored according to their position (see the related color-legend bar). Filled histograms refer to the nonameric model whereas dashed histograms refer to the hexameric model. (C) The native 2D-network of the retinal binding pocket (RBP), i.e. network made by the nodes linked to 11-*cis*-retinal (herein named RET), concerning the nonameric model is shown. Hubs are represented as pentagons. Nodes are colored according to their location (see the related color-bar legend), whereas links are colored according to their frequency (see the related color-bar legend). Only stable links are shown (i.e. holding a frequency ≥ 25%). Link width is proportional to the link strength (i.e. the average interaction strength between the linked nodes). The orientation of the ligand in the central image is unrelated to the orientation in the 3D structure. (D) The native 2D-network of the RBP concerning the hexameric model is shown. See the explanation in panel (C). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

challenge the computational model in its ability to predict the effects of non-conservative substitutions at the same site. Moreover, A164E, P171E, and Y178N were poorly biochemically characterized so far. Indeed, P171E lacks any biochemical characterization, whereas A164E and Y178N are ascribed to the misfolded Class 2 by Athanasiou and co-workers [12] (Table S1). As for the other three new mutants, G182S and G188E were widely

reported as retained in the ER and unable to reconstitute with 9-*cis*-retinal (i.e. ascribed to class IIa/II [10,11] and Class 2 by Athanasiou and co-workers[12]), whereas P347S was ascribed to the WT-like class I (or 1) (Table S1 [10–12]). Thus, the criteria for selecting the test-set mutants did not include the mutation site and the choice was justified by the fact that the correlative model indeed is based on adRP mutants variedly distributed in those rod
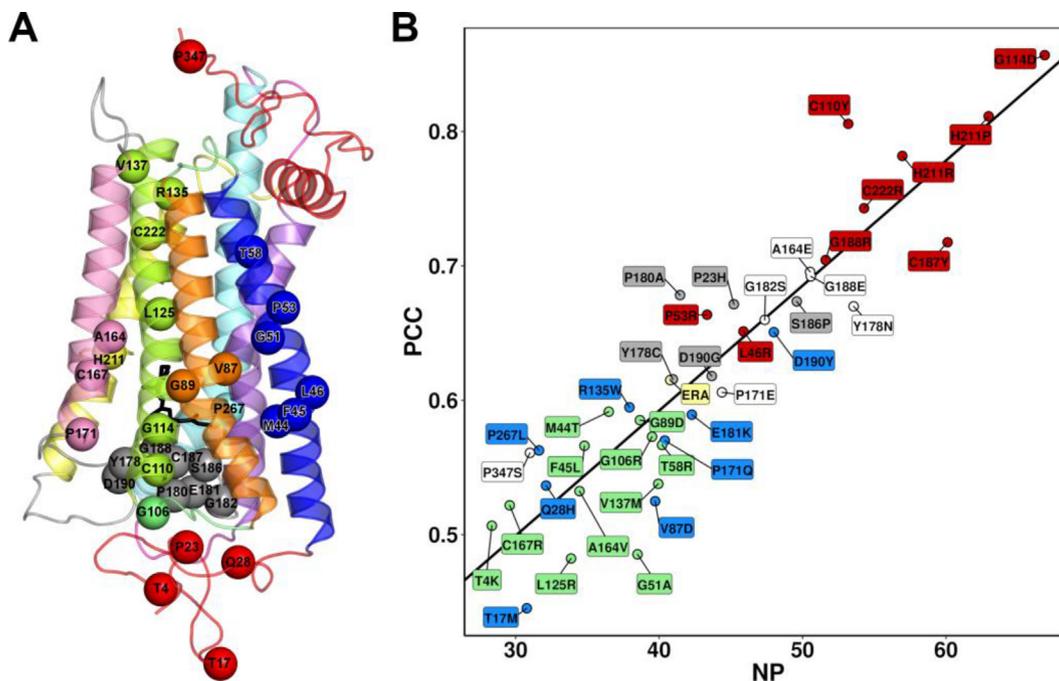
Fig. 3. Linear correlation between the indices of network perturbation (NP_SMD) and ER retention (PCC). (A) The mutation sites on the structure of dark rhodopsin are represented as spheres centered on the Cα-atoms on the crystal structure of dark rhodopsin (PDB: 1U19). (B) The scatter plot shows the linear correlation between the NP_SMD accounting for mutational effects on the native structure network and the PCC index accounting for ER retention. The colors of circles reflect the structural cluster assignment for each mutant [18] (i.e. clusters 1, 2, 3, and 4 are, respectively, green, blue, gray, and red). The linear regression equation is PCC = 0.0097NP_SMD + 0.2035, r = 0.9191; p = 2.026157e-16. Such equation was used to predict the PCC of the six novel mutants A164E, P171E, Y178N, G182S, G188E, and P347S, shown on the plot as well (white circles) as well as of P23H-bound to the small chaperone 13-*cis*-5,8-ERA (yellow circle, see below in the text). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

opsin regions that host almost all adRP rod opsin mutations discovered so far (i.e. all helices but H7, N- and C-termini, as well as E2). This suggests that the performance of the computational model is independent of the replacing amino acid location.

To validate predictions by the computational model and deepen the characterization of the adRP mutants, we analyzed *in vitro* the subcellular localization of the six new mutants, following the protocol previously described [18]. In brief, WT and mutant rod opsins were transiently expressed in COS-7 cells in the absence and presence of 9-*cis*-retinal. The ER localization of opsin and *iso*-rhodopsin were quantified by the PCC index. As already stated in our previous study [18] while *in silico* experiments on the rhodopsin state employed the natural chromophore, 11-*cis*-retinal, *in vitro* experiments employed 9-*cis*-retinal. The ability of 9-*cis*- and 11-*cis*-retinal to promote proper folding and trafficking has been previously demonstrated for the adRP mutants P23H, T17M, and Q28H [11,15,55–58] and the fact that the 9-*cis*-isomer is chemically more stable than the 11-*cis*-isomer while showing similar behaviors persuaded us to use 9-*cis*-retinal in *in vitro* studies [58,59].

Plasma membrane localization was assessed in the absence of cell permeabilization by the RetP1 antibody, recognizing the extracellular N-terminus of the protein. Samples treated in parallel in the same conditions were used to define the percentage of transfected cells in each experiment, based on the analysis with an antibody, called 1D4, detecting all opsin in the cells not only the one at the cell surface. The ratio between these two values (i.e. RetP1/1D4) quantified the number of cells showing plasma membrane localization of rod opsin. As already seen in our previous study [18], all mutants and even the WT form display some level of ER retention which is expected because we employed transient transfection, which means that cells start to express rod opsin few hours after DNA uptake and therefore a lot of protein is produced.

Thus, the protein, highly synthesized by the cell, will be detected in the ER. In these conditions, a PCC equal to 0.49 for WT rhodopsin suggests that a high amount is at the plasma membrane and cells are still synthesizing the protein in the ER. Indeed, this also happens in photoreceptors, in which the ER in the inner segment is normally stained.

Expression of the protein in a cellular system showed that, in the absence of 9-*cis*-retinal, five mutant opsins were retained in the ER and did not reach the plasma membrane. P347S was only slightly more retained in the ER than the WT (Fig. 4A). Quantification of these observations by RetP1/1D4 ratios and PCC demonstrated that 9-*cis*-retinal was able to mitigate ER retention of P171E to a higher extent than it does for A164E, Y178N, G182S, and G188E (Fig. 4B,C, Tables 1 and S1).

Remarkably, the computational model was able to predict the PCC index of the six new mutants, the largest difference between predicted and determined PCC values being 0.057 (Fig. 3B and Tables 1 and S1).

On the basis of these results, a protocol for predicting the mis-localization (i.e. ER retention) of adRP rod opsin mutants on the basis of the network perturbation score was established, which is made of the steps illustrated in Chart 1 (see also Methods). Those steps consist in: a) running 10 SMD simulations per system (i.e. two different oligomeric models of WT and mutants the latter considered also in different rotameric states, when applicable); b) building a PSG per SMD trajectory; c) for each system, building an average PSG over the five most similar (i.e. in link frequencies) PSGs; d) computing the NP_SMD score as an average of the scores achieved by the two oligomeric models; and e) building a linear correlation between network perturbation score and mis-localization index (e.g. the PCC of rod opsin-calnexin co-localization) to predict the effect of new mutations on the subcellular localization of the target protein. If the replacing amino
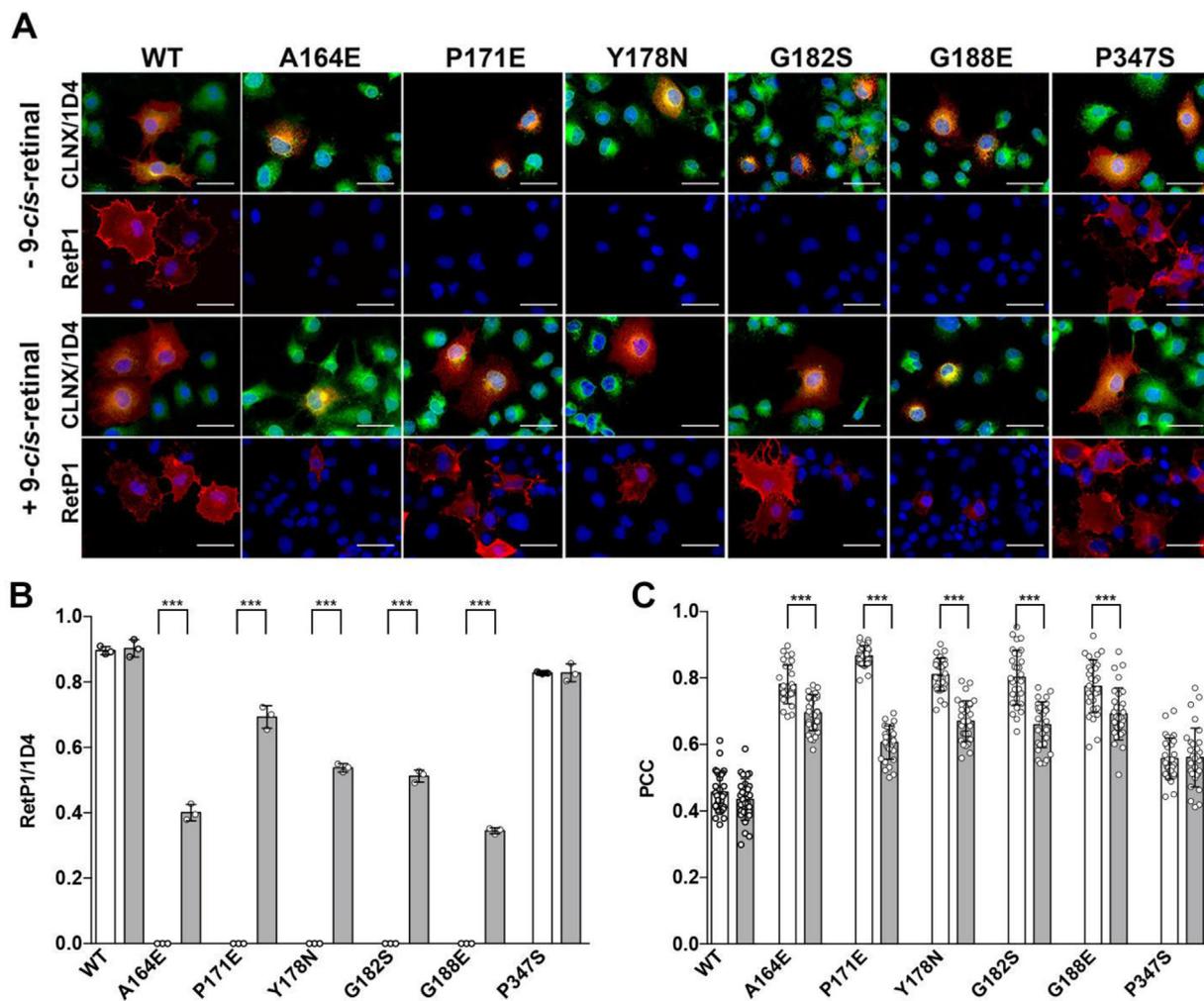
Fig. 4. Subcellular localization of the new opsin mutants in the absence and presence of 9-*cis*-retinal. (A) COS-7 cells were transfected with either WT opsin or mutant opsin. Cells were not treated (upper 2 rows, − 9-*cis*-retinal) or treated (lower 2 rows, + 9-*cis*-retinal) with 10 μM 9-*cis*-retinal. Immunofluorescence detected opsin (red) using 2 different antibodies. In the first and third rows cells were permeabilized and the 1D4 anti-opsin antibody, which binds an intracellular epitope, was used for the analysis of intracellular opsin. ER localization was assessed by co-staining with an antibody binding calnexin (green), an ER-resident protein. Nuclei were stained in blue. Transfected cells expressing opsin (labeled in red) were analyzed for co-localization of opsin with calnexin (CLNX/1D4). Co-localization (yellow merge of red and green) represents opsin protein in the ER, which in all transfected cells is detectable because cells highly produce the protein after genetic modification with transfected DNA. Nevertheless, our previous study demonstrated that if a mutation induces ER-retention the co-localization is significantly higher [18]. We measured ER-retention by image analysis (see Methods) using the PCC index, which quantifies co-localization of opsin (red) with calnexin (green). In (C) histograms of PCC from cells cultured in the absence (white bars) and presence (grey bars) of 10 μM 9-*cis*-retinal are shown. Student's *t*-test was applied to compare data from samples treated or not treated with 9-*cis*-retinal. *** P ≤ 0.001. In the second and forth rows of panel A cells were not permeabilized and the RetP1 anti-opsin antibody (red), which binds an extracellular epitope, was used for the analysis of opsin localized at the plasma membrane of transfected cells. Mutations, which cause ER-retention, prevent the protein from being properly localized at the cell surface and cannot be detected with the RetP1 antibody. Nuclei were stained in blue. The effects of the mutations on plasma membrane localization were quantified by counting cells stained by RetP1 and comparing the number of total cells transfected in each experiment (i.e. cells expressing opsin irrespective of its localization, thus stained after membrane permeabilization with 1D4 antibody). (B) Histograms of the RetP1/1D4 values representing the ratio of cells with rod opsin properly localized at the plasma membrane in the absence (white bars) and presence (grey bars) of 10 μM 9-*cis*-retinal are shown. The scale bars in the microphotographs are 50 μm. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

acid exists in different rotameric states, all of them have to be simulated and the average NP_{SMD} score considered for PCC predictions.

In summary, *in vitro* characterizations of six adRP mutants, not included in the training set, validated the ability of the computational model to predict their ER-retention under the effects of 9-*cis*/11-*cis*-retinal. These results laid the groundwork of a computational protocol for predicting the mis-localization of misfolded mutants on the basis of a network perturbation score.

3.3. The computational model provides the structural determinants of retinal responsiveness by adRP RHO mutants

In a recent study, the network perturbation indices computed for the opsin and rhodopsin forms of each adRP mutant and the dif-ferences between them were used to cluster the mutants, leading to four groups [18]. On average, clusters 1 and 4 show, respectively, the lowest and highest destabilization of the native structure network independent of the presence of retinal, whereas cluster 2 resembles cluster 4 in the absence of retinal and cluster 1 in the presence of retinal. In other words, for cluster 2, the presence of retinal hinders structural impairments by the mutations. Mutants of cluster 3 falls in between those of clusters 2 and 4, as far as the structural chaperone effect of the 11-*cis*-retinal is concerned [18].

Here, the ability of the computational protocol to predict ER retention of the new adRP *RHO* mutants was exploited to assign those mutants to one of the four clusters, with no need to simulate also the opsin forms.

Table 1
Predicted and determined values of the PCC index of ER-retention.

Mut	NP$_{SMD}$	PCC$_{pred}$[a]	PCC[a]	ΔPCC[a]
A164E	50.523	0.693	0.696	−0.002
P171E	44.381	0.634	0.606	0.028
Y178N	53.557	0.723	0.670	0.053
G182S	47.364	0.663	0.660	0.003
G188E	50.578	0.694	0.692	0.002
P347S[b]	30.998	0.504	0.561	−0.057
P23H-ERA[c]	40.779	0.599	0.615	−0.016
P23H-RET[d]	45.208	–	0.671	–

[a] Predicted (PCC$_{pred}$) and determined (PCC) values and difference between them (ΔPCC).
[b] For the P347S C-terminal mutant, NP$_{SMD}$ score and PCC$_{pred}$ concern only the nonameric input model as the hexameric model lacks the C-terminus.
[c] P23H bound to the 13-*cis*-5,8-ERA mixture of the 5S8R and 5R8S enantiomers.
[d] The NP^{SMD} and PCC of P23H rhodopsin has been reported here to allow comparison with the 13-*cis*-5,8-ERA-bound form of the same mutant.
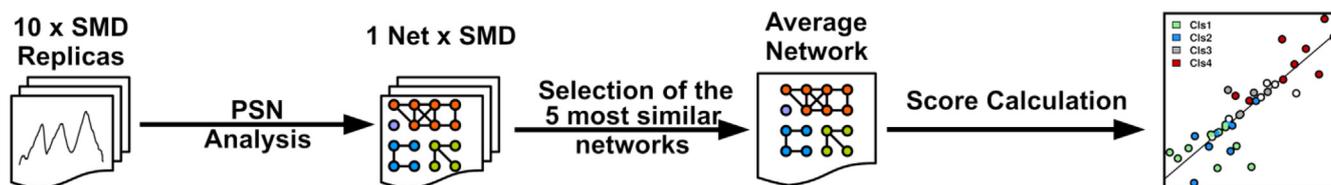
Chart 1. Computational protocol for predicting ER retention by adRP mutation.

Since mutant clustering strongly relied on the structural chaperone effects of 11-*cis*-retinal, focus was put on perturbations in the stable RBP native hubs including 11-*cis*-retinal (RET, Fig. 2D). The percentage of perturbed native RBP hub-links turned out to be an effective signature of mutant clusters. In this context, link perturbation by mutation means reduction in frequency compared to the WT and it is computed according to the following formula:

$$\text{Link}_{pert} = \frac{\text{WT}_{freq} - \text{Mut}_{freq}}{\text{WT}_{freq}} 100$$

where Link$_{pert}$ stands for link perturbation and WT$_{freq}$ and Mut$_{freq}$ indicate the average frequency of the considered link over the five SMD trajectories of the WT and mutant forms, respectively. Any Link$_{pert} \leq 0$ is set to zero. Thus, native hub-link perturbation per hub is the percentage of native hub-links with reduced frequency compared to the WT.

The percentage of perturbed RBP hub-links per native hub and the degree of sequence conservation of the amino acid behaving as a hub resulted as good coordinates for rod opsin misfolding by mutation (Fig. 5A). It is worth noting that stable native hubs from SMD simulations on both the nonameric and hexameric models contributed to the population surface. RBP native hubs hold low sequence conservation (i.e. between 4 and 6). Basins move from the left (low hub-link perturbation) to the right (high hub-link perturbation) going from clusters 1 to cluster 4. The model barely distinguishes the behavior of cluster-1 mutants from that of cluster-2 mutants, in line with previous observation that they behave almost the same in the rhodopsin/*iso*-rhodopsin states [18]. Cluster-3 mutants stay in between those of clusters 1/2 and 4.

On these bases, we used the perturbation of native RBP hub-links as a way to assign the six new mutants to one of the four clusters (Fig. 5B). Mutant assignment was based on the average difference in the perturbation of the RBP native hubs between each new mutant and the four clusters of mutants, according to the following formula:

$$\Delta_{m,c} = \frac{\sum_h^H |m_h - c_h|}{H}$$ where m is one of the new six mutants, c is one of the four clusters of mutants, h is one of the native hubs in the

RBP, H is the total number of stable native hubs in the RBP shell, m_h is the perturbation of hub h in new mutant m and c_h is the average perturbation of the same hub calculated over all the members of cluster c. The lower $\Delta_{m,c}$ the closer a mutant is to the mutants of a given cluster.

On these bases, the following assignments were done: P347S to cluster 1, Y178N and G188E to cluster 4 and the remaining mutants, A164E, P171E, and G182S, to cluster 3 (Fig. 5A and 5B). These assignments are quite in line with those based on the PCC values of the opsin and rhodopsin forms and the relative ΔPCC (Table S1). The only discrepancy is represented by P171E that *in vitro* data would assign to cluster 2 according to the ΔPCC between opsin and *iso*-rhodopsin (Fig. 4C and 5B).

Remarkably, structure-based cluster assignment done herein for Y178N and G188E are in line with previous assignment of such mutant to the misfolded Class 2 [12].

Thus, following the integration of the six new mutants in the four clusters, the analysis of link perturbation was extended to all 51 native stable hubs inferred from SMD simulations on the nonameric and hexameric arrangements (Fig. 6A,B). For the members of clusters 1 and 2, on average, only a minority of native hubs (12% and 20%) undergoes significant perturbation in their links (i.e. a reduction in frequency for at least 50% of their links; Fig. 6A,B). In contrast, for the members of cluster 4, 55% of hubs hold frequency reduction in at least 50% of their links. Remarkably, perturbed hubs in cluster 4 include 11-*cis*-retinal. This strengthens the hypothesis that the inability of 11-*cis*-retinal to exert a chaperone effect on the mutants of cluster 4 is likely due to the inability of the compound to properly bind and establish native connections with the protein. The members of cluster 3 stay in between those of clusters 2 and 4 as 33% of native hubs show perturbation in at least 50% of their links (Fig. 6A,B).

The distributions of percentage of perturbed hub-links per hub and the degree of sequence conservation of the amino acid behaving as a hub confirmed as good fingerprints of rhodopsin misfolding by mutation (Fig. 6C). Indeed, for the members of clusters 1 and 2, the population of native hubs holding at least 50% of links weakened is low independently of hub conservation. In this respect, the
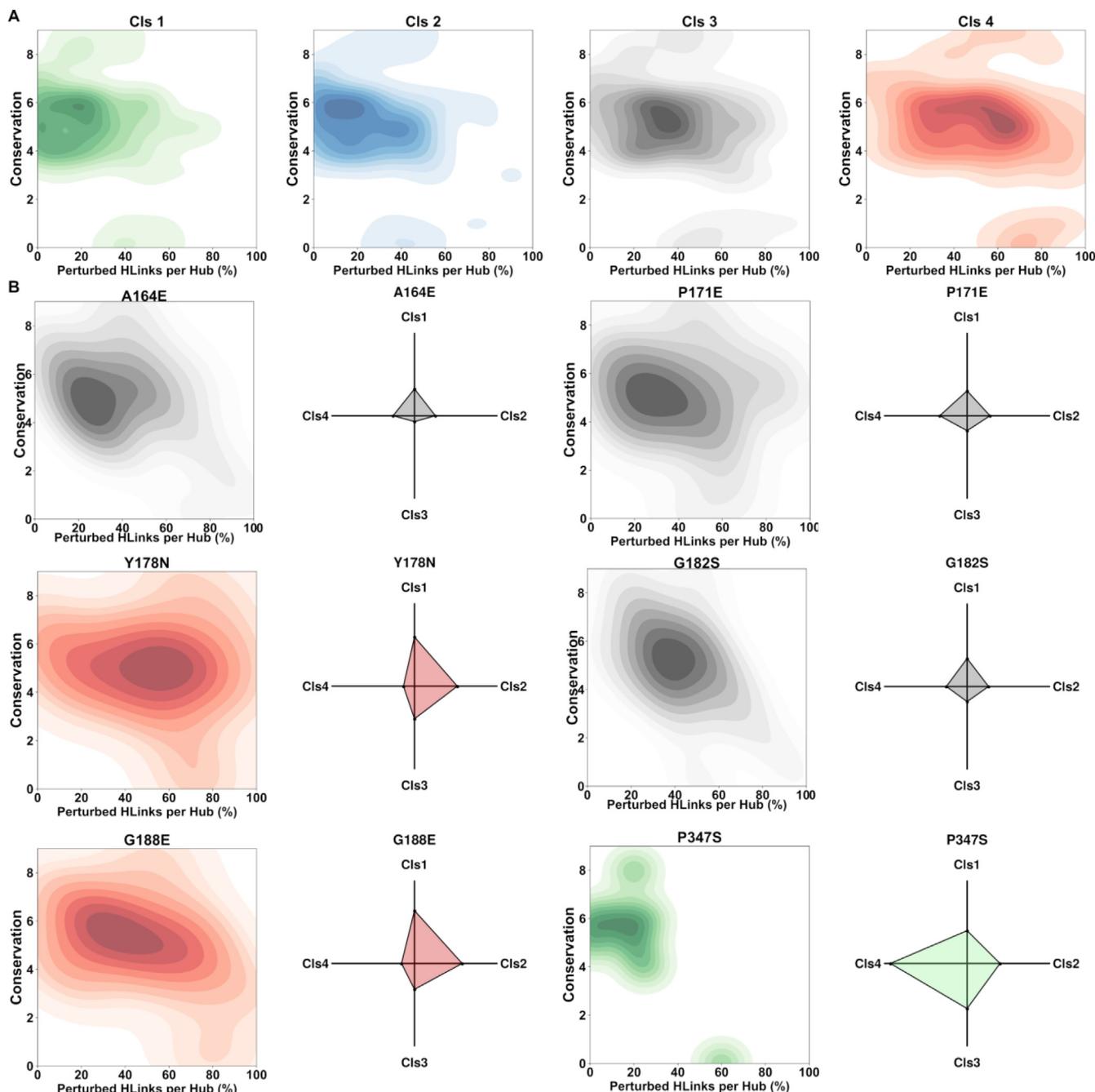
Fig. 5. Classification of the six new mutants. (A) The distribution of native stable RBP hubs in all mutants in a cluster and in each new mutant is shown according to two coordinates: the percentage of perturbed links per hub and the conservation grade (according to ConSurf) of the amino acid residue behaving as a hub. (B) The radar charts show the average difference in the perturbation of the stable native RBP hubs between each new mutant and the four clusters of mutants. The lower the difference, the closer the mutant is to one of the four clusters.

most populated basins involve native hubs with almost absent link perturbation (i.e. 0–20%) and either low (3–6) or high (7–9) degrees of sequence conservation (Fig. 6C).

In contrast, cluster-4 mutants behave in a clearly different way as high populated basins of hubs holding low (4–6) sequence conservation can be found at very high (80–100), high (60–80%), or moderate (40–60%) percentages of perturbed links (Fig. 6C). High sequence conservation associates with moderate link perturbation (Fig. 6C). Again, the behavior of cluster-3 mutants stays in between those of clusters 2 and 4.

For the mutants of cluster 4, the basin at high link perturbation (60–80%) and low conservation (ConSurf 4–5) concerns, with a few

exceptions, amino acid residues essentially located in RBP or close to it: T118 (H3), E122 (H3), W126 (H3), V138 (H3), F148 (I2), R177 (E2), Y178 (E2), E181 (E2), M183 (E2), I189 (E2), Y191 (E2), Y192 (E2), N200 (H5), F203 (H5), M207 (H5), K248 (H6), F293 (H7), Y301 (H7), and K311 (H8) (Fig. 6C).

The other basins at moderate link perturbations (40–60%) and high amino acid conservation (ConSurf 7–9) involve hubs corresponding to functionally important amino acids: e.g. D83 (H2), the ERY motif in H3, W161 (H4), Y223 (H5), M257 (H6), W265 (H6), as well as N302 (H7) and Y306 (H7) of the NPxxY motif.

Thus, the population of hubs with variedly perturbed links and sequence conservation distinguishes the clusters of mutants that,
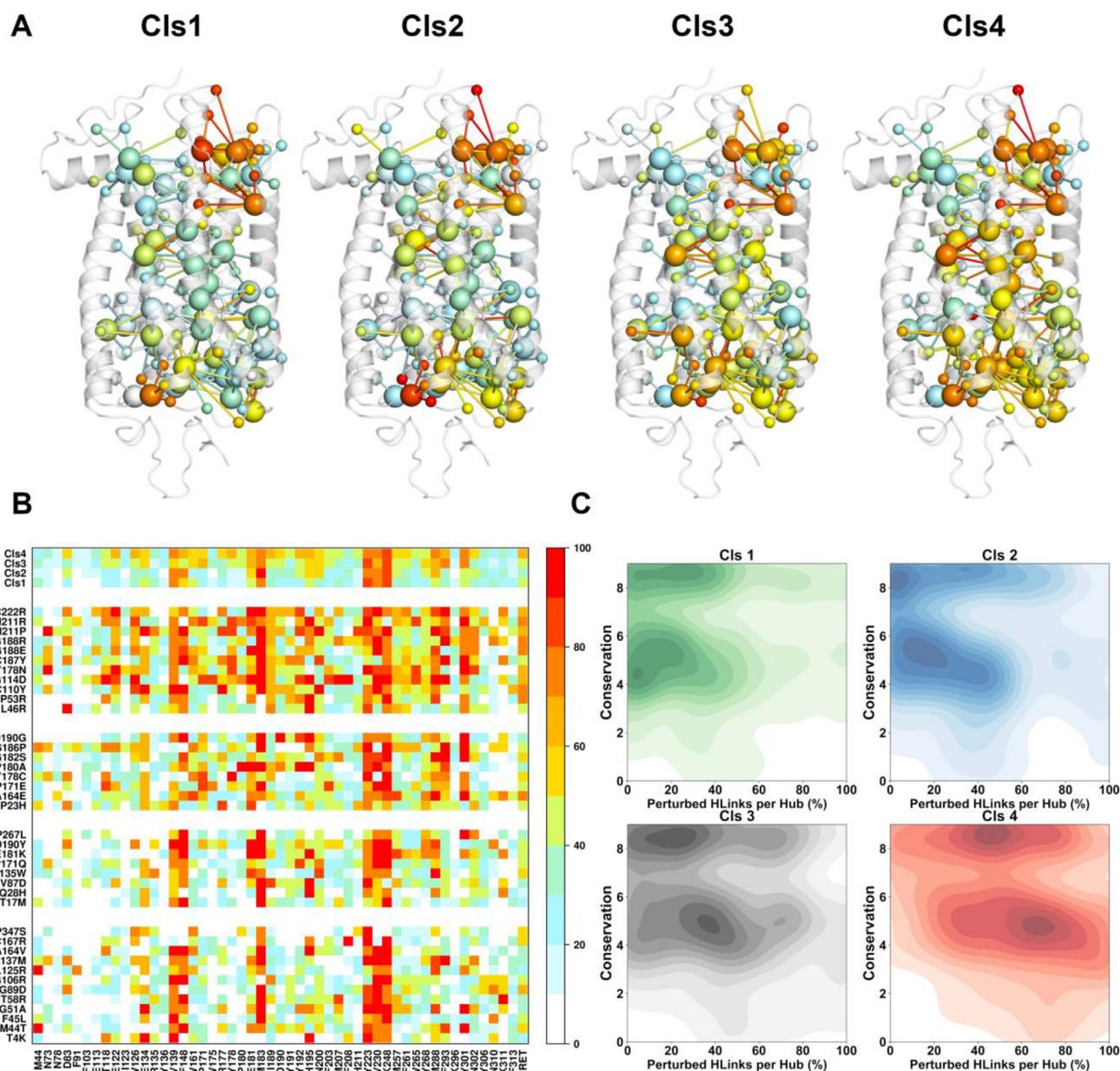
Fig. 6. Structural effects of adRP mutations. (A) The native hub-links and their connected nodes are shown on the crystal structure of dark rhodopsin (PDB: 1U19). The big spheres (centered on the Cα-atoms) are the 51 native stable hubs, whereas the small spheres are the nodes linked to such hubs. Each of the four structures shows the average perturbation, expressed by coloring (see the color scale in B), of the native hub-links over the mutants of each cluster in the two different oligomeric arrangements. In detail, the color of each hub represents the percentage of perturbed hub-links (i.e. with reduced frequency with respect to the WT) (see the color-legend bar on the right of the matrix). The color of each link represents the percentage of mutants in the cluster bearing such link perturbed. Each non-hub node is colored similarly to its link. (B) Rows refer to the molecular systems, 39 adRP-linked mutants divided in four clusters according to the previous [18] and present studies. Cls stands for cluster. Columns refer to the native stable hubs. The color of each matrix element represents the percentage of native hub-links perturbed as an effect of mutation in the two different oligomeric models. For Cls1, Cls2, Cls3, and Cls4 (first four rows from the top), the color of each matrix element represents the link perturbation averaged over the mutants of a cluster in the two oligomeric models. (C) The distribution of native stable hubs in all mutants in a cluster is shown according to two coordinates: the percentage of perturbed native links per hub and the conservation grade (according to ConSurf) of the amino acid residue behaving as a hub. Stable native hubs from SMD simulations on all the mutants in a cluster in both the nonameric and hexameric models contributed to the population surface.

in the presence of 11-*cis*-retinal, hold almost normal subcellular localization (i.e. the members of both clusters 1 and 2) from those that are still retained in the ER (i.e. the members of cluster 4). In that respect, ER retention couples with hub weakening in RBP. Remarkably, perturbation of native RBP hub-links could be used to assign the new mutants to one of the four clusters with no need to simulate the opsin form, an important progress compared to the previous study [18].

3.4. The computational model predicts the chaperone effect of 13-cis-5,8-ERA

In the previous study [18] we discovered a reversible orthosteric inhibitor of retinal binding, 13-*cis*-5,8-ERA, more effective than 9-*cis*-retinal in mitigating ER retention of three adRP mutants including P23H, the adRP mutant most studied *in vivo* [60–63]. For P23H, the small chaperone, tested at 10 µM concentration as a

mixture of four optical isomers (i.e. 5S8R, 5R8S, 5R8R, and 5S8S), showed a PCC equal to 0.640, slightly lower than that of 9-*cis*-retinal (i.e. 0.670, Table S1) [18]. In this study, we tested *in vitro* the *trans*-mixture of two enantiomers (5S8R and 5R8S), finding a further reduction in the PCC of P23H (i.e. to 0.615). In this regard, the *trans*-mixture was made on demand in an attempt to improve the chaperone effect compared to the four-isomer mixture.

To further challenge the ability of the computational protocol to predict the chaperone behaviour of small compounds, SMD simulations were carried out on the two enantiomers of 13-*cis*-5,8-ERA bound to the P23H opsin mutant in the two oligomeric assemblies, followed by PSN analysis.

The docking mode of the two enantiomers is characterized by the charge-reinforced H-bond between the carboxylate of 13-*cis*-5,8-ERA and both Y268 (in H6) and K296 (in H7) (Fig. S3A-F). For both enantiomers, the salt bridge between the ligand carboxylate and K296 persists up to H7 unfolding and pulling (i.e. till the end of pulling), mimicking, at least in part, the covalent protonated Schiff-base in rhodopsin (Fig. S3B,C,E,F). The olefinic chain of the ligand makes van der Waals interactions with T118 (in H3), I189 (in E2), and Y268 (in H6). Major divergences among the two isomers occur at the bicyclic moiety of the ligand that, especially for the 5R8S isomer, occupies the same site as the β-ionone ring of 11-*cis*-retinal, making van der Waals interactions with M207, F208, F210 (all in H5), and W265 (in H6).

Collectively, while occupying the same binding site, the two enantiomers of 13-*cis*-5,8-ERA and 11-*cis*-retinal display divergent interactions.

The average NP$_{SMD}$ score relative the two isomers in the two different oligomeric models was used to predict the PCC of the racemic mixture by the linear correlation equation shown in Fig. 3. Remarkably, the correlative model correctly predicted the chaperone action of the two-isomer mixture (i.e. the predicted PCC values is indeed 0.615, Table 1).

In line with the PCC index of ER retention, the distributions of hub-link perturbation highlighted the racemic mixture of 13-*cis*-5,8-ERA as more able to prevent perturbations by the P23H mutation than 11-*cis*-retinal (Fig. 7A,B). Indeed, whereas for 11-*cis*-retinal the percentage of native hubs holding ≥ 50% of perturbed links is 41, for the 13-*cis*-5,8-ERA mixture such percentage lowers to 33 (Fig. 7A). Those divergences in favor of 13-*cis*-5,8-ERA increase by considering only the RBP-hubs (Fig. 7B), i.e. 45% for 11-*cis*-retinal and 23% for 13-*cis*-5,8-ERA, indicating a better ability of the new small chaperone to prevent structural perturbation caused by the P23H mutation especially in the RPB, which is predicted to participate in the stability core of rod opsin. Consistently, for the 13-*cis*-5,8-ERA mixture, the distributions of native hub-link perturbation versus evolutionary conservation highlights the main basins at low perturbation (i.e. between 0% and 20%) independently of sequence conservation, whereas for P23H rhodopsin the majority of low conserved hubs (e.g. those in the retinal binding site) hold relatively low perturbation (between 10% and 30%), whereas hubs corresponding to highly conserved amino acids tend to hold slightly higher link perturbations (i.e. between 30% and 40%) (Fig. 7C).

In summary, the computational model proved effective in predicting if a small compound is a more potent pharmacological chaperone than 11-*cis*-retinal, providing also insights into the structural determinants of small chaperone action at misfolded adRP *RHO* mutants.

4. Discussion

Missense mutations, which affect folding or trafficking of integral membrane proteins cause an ever increasing number of genetic diseases, including adRP linked to *RHO* mutations, an incurable disease that awaits light on the pathogenic mechanisms and, among those, its molecular structure determinants [1,5,51].

An understanding of the structural determinants of the disease implies deciphering the effects of misfolding mutations on the molecular forces deputed to the folding and stability of the protein. In this study, we faced this issue by comparing *in silico* mechanical unfolding of WT rod opsin with those of 39 adRP-linked mutants under the effects of the small chaperone 11-*cis*-retinal. Indeed, the folding of an unfolded polypeptide chain into a functional protein derives from the interplay of intramolecular interactions, and the same is true for protein destabilization/misfolding and unfolding. *In vitro* approaches to decipher the basic language of membrane protein folding relied since year 2000 on single-molecule force spectroscopy (SMFS) based on Atomic Force Microscopy (AFM) [64]. Since then, such strategy has been adopted to gain insights into the unfolding and folding mechanisms of a number of membrane proteins including bovine rhodopsin [64]. SMFS experiments on bovine rhodopsin probed molecular interactions within the native protein and discovered structural segments of well-defined mechanical stability [65,66]. An effective way to increase the resolution of *in vitro* SMFS data is to combine them with atomistic SMD simulations and analyses [16]. A strategy to infer those regions, which are essential for protein stability and oppose resistance to mechanical unfolding, is to represent protein structures as networks (PSNs) of interacting amino acids and to investigate perturbations in the native PSGs during the unfolding process and under the destabilizing effects of misfolding mutations [16]. Herein, extensive SMD simulations on WT and 33 adRP rhodopsin mutants with known degree of ER retention [18] in different rotameric and oligomeric states, coupled with PSN analysis, were instrumental in investigating the structural determinants of adRP linked to rod opsin mutations. We could define an index of structural perturbation by mutation (i.e. NP$_{SMD}$), relying on the weakening of native hubs and their links under the effect of point mutations. A correlative model based on NP$_{SMD}$ index and PCC index of ER retention was built and validated here in its ability to predict ER retention, under the effect of 11-*cis*-retinal, of six adRP rhodopsin mutants not included in the training set (i.e. A164E, P171E, Y178N, G182S, G188E, and P347S). Due to its higher stability, 9-*cis*-retinal substituted for 11-*cis*-retinal in our previous [18] and present *in vitro* determinations, while simulations were carried out on the natural chromophore. The *in vitro* response of cluster-2 and cluster-3 mutants may have been overestimated just because of the higher stability of the 9-*cis*-retinal analog, whose regeneration in rod and cone pigments differs from that of 11-*cis*-retinal [67,68]. Yet, other studies reported that the two retinoids behave similarly in improving trafficking of T17M, Q28H (both in cluster2) and P23H (in cluster3) [11,15,55–58]. In line with that, the good correlation between the structural NP$_{SMD}$ index computed on the rhodopsin forms and the PCC determined on the isorhodopsin forms further validate our choice to use the more stable 9-*cis*-analog in sub-cellular localization experiments.

Remarkably, the model was able to predict also the chaperone effect of a racemic mixture of the small compound 13-*cis*-5,8-ERA on the P23H mutant opsin, estimating such mixture as more effective than 11-*cis*-retinal in mitigating ER retention by the P23H mutant. Even if the small compound is not able to revert ER retention of the cluster-3 mutant to the WT levels, considering that, like the majority of adRP mutant opsins, P23H causes a slow retinal degeneration, the small chaperone may be able to decrease degeneration, thus giving the patient further years of vision. Based on these considerations, the computational protocol developed here can help identifying small molecules for precision therapeutic intervention for a rare highly-debilitating incurable disease.
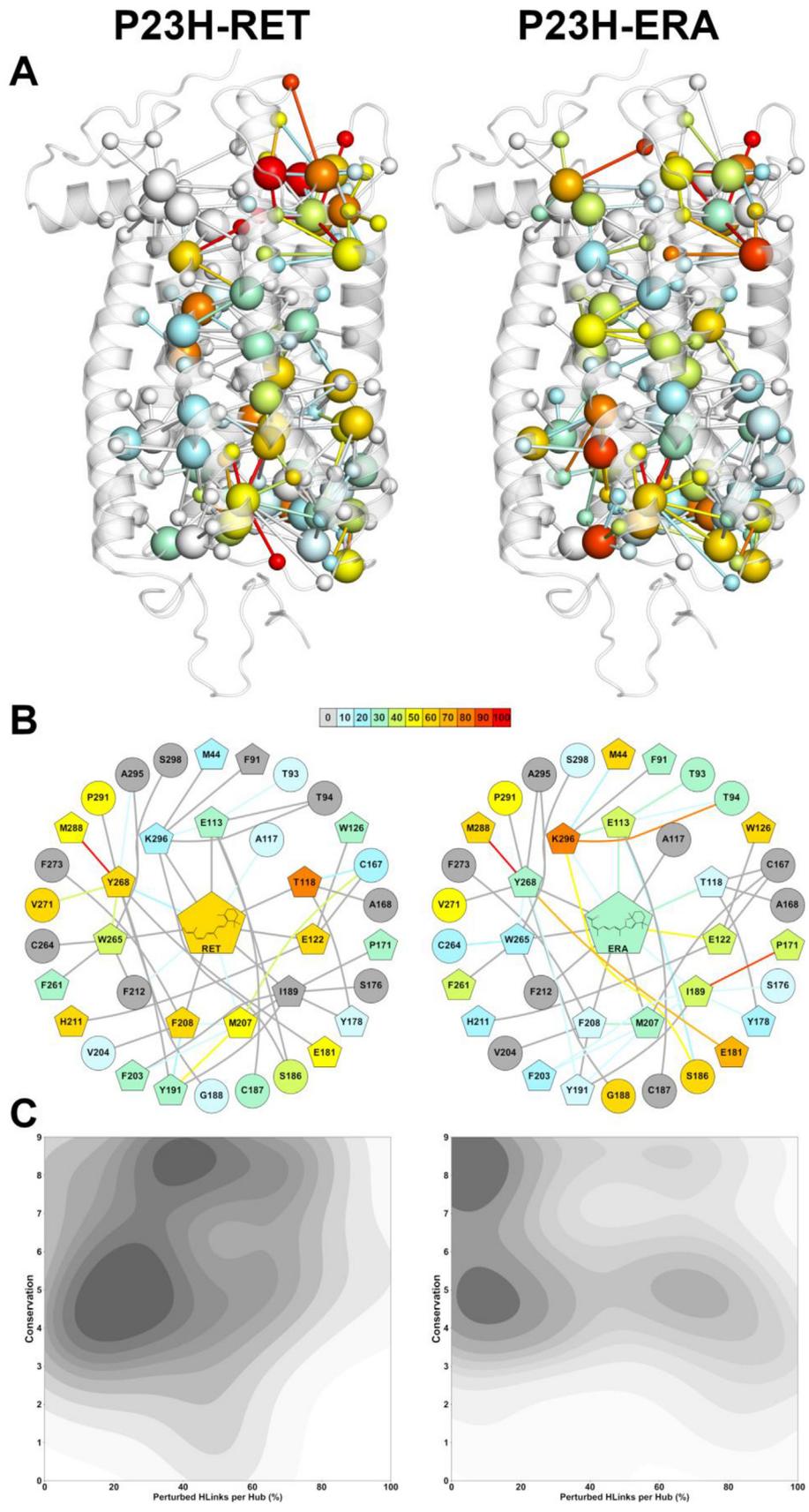
The structural determinants of mutant rod opsin misfolding would include weakening of the connections made by selected native hubs, which include non-conserved amino acids in the retinal binding site and highly conserved amino acids in the cytosolic side of the seven-helix bundle. The small chaperone action relies on the establishment of direct connections with the RBP hubs, thus preventing mutation-caused destabilization of both the RBP network and the distal hubs corresponding to highly conserved amino acids. The latter include the glutamate and the arginine of the E/DRY motif at the cytosolic end of H3, W161 in the middle of H4, Y223 in the cytosolic end of H5, F261 and W265 of the aromatic cluster in H6, and N302 and Y306 of the NPxxY motif in H7. The destabilization of such conserved residues may concur to cause not only adRP linked to *RHO* mutations but also other incurable conformational diseases hitting Family A GPCRs. Lack of retinal responsiveness by the mutants of cluster 4 likely relates to the deformations induced by mutation in the RBP, which prevents small chaperones from establishing the required interactions.

The computational structural model developed here relies on two different oligomeric models of rhodopsin, a nonamer and a hexamer, both consistent with experimental information [39,40]. Merging the results of simulations on two different oligomeric arrangements provided more robustness to the computational structural model. In those supramolecular arrangements only one monomer was subjected to pulling and mutation, whereas, the dynamics of the non-pulled monomers were limited to the side chains in direct contact with the pulled monomer, leading to an implicit solvent/protein cage model. Whether rod opsin oligomerization is an artifact or a physiologic condition has long been debated [69–73]. Along the same line, whether rod opsin is monomeric, dimeric or oligomeric into the ER is not known. Although the role of rod opsin oligomerization awaits clarification, functional characterization of rhodopsin monomers and dimers in detergents demonstrated that monomeric rhodopsin is able to activate transducin, though the oligomeric form is more active [74,75]. *In vitro* experiments in solution and in nanodiscs containing one and two rhodopsin molecules strengthened the evidence that the receptor monomer holds the structural determinants for transducin activation and, in this regard, it is the functional unit [70,71,76,77]. Oligomerization may be essential for the ontogeny and/or desensitization of rod opsin and, hence, for the control of light signaling. Aggregation between WT rod opsin and some partially misfolded adRP *RHO* mutants may have implications in the etiology of the disease and in responsiveness to retinoid chaperones [78,79].

In the context of this study, the employment of oligomeric models of rhodopsin merely served to prevent artifacts such as anomalous rotations of the protein during pulling in implicit membrane/water, which violate experimental observation, as already discussed [16,41]. Therefore, investigating aggregation of partially misfolded rod opsin mutants with WT or impairment in rod opsin oligomerization by mutation as possible determinants of the disease goes beyond the scope of this study. In summary, we merged

the results of simulations inferred from two significantly different oligomeric rod opsin models to further strengthen the already demonstrated independence of the model from the input structural arrangement [41]. On these bases, we may reasonably neglect possible changes in oligomeric order and architecture of WT and mutant rod opsin on changing cellular compartment, stage of opsin's life-cycle, and functional state.

In conclusion, a structure-based automated approach was developed, which is able to predict ER retention of novel adRP mutants and the ability of small chaperones to mitigate such ER retention with therapeutic implications. In this framework, the distributions of stable native hubs in the four clusters of mutants as a function of hub-link weakening and amino acid conservation provided valuable profiles of rhodopsin misfolding by mutation, which could distinguish mutants that do not respond to small chaperones from mutants either non misfolded or responsive to small chaperones. This is one of the remarkable novelties of this study with respect to studies published by ourselves and by others. Another novelty of the present model is that non-misfolded or chaperone-responsive mutants can be distinguished from misfolded mutants non-responsive to small chaperones with no need to simulate the opsin form.

Collectively, the correlation between the indices of structural perturbation and ER retention got herein and in the previous study [18] validate our initial hypothesis that the adRP rod opsin mutants biochemically classified as misfolded (reviewed in [10–12]), represent quasi-native states of the protein. Indeed, the mutants T58R, G89D, G106R, L125R, A164V, and C167R, ascribed to Class-2 misfolded and retinal non-responsive mutants by Athanasiou and co-workers [12], belong to our cluster-1 mutants, poorly structurally perturbed, poorly retained into ER, and able to reach plasma membrane independently of the retinal presence. Moreover, our cluster-2 and cluster-3 mutants, previously biochemically classified as misfolded (e.g. falling in classes IIa/II or IIb/III or Class-2) [10–12], according to our experiments are variedly relieved in their ER retention by the presence of 9-*cis*/11-*cis*-retinal. In summary, all cluster 1–3 mutants, i.e. 28 out of the 39 mutants considered in this study, either independently or not of the presence of 9-*cis*/11-*cis*-retinal, are somehow able to be transported to the plasma membrane. Remarkably, some cluster-1 mutants while not mislocalized may not respond to the retinal, due to deformations in the retinal binding site. This is, for example, the case of L125R, whose replacing amino acid causes perturbations at the interface between H3 and H5 in proximity to the β-ionone ring of the chromophore. Incidentally, we expect the same deformations also in the E122L/L125R double mutant, in line with the observations that it was still unable to bind 11-*cis*-retinal in spite of its improved expression, glycosylation and trafficking compared to L125R [44]. For some of our cluster-1 mutants, e.g. G89D, L125R, and A164V, early mass spectrometry determinations indicated partial or complete misfolding of rod opsin due to the formation of a non-native disulphide bridge between C185 and C187 in EL2 [80]. Our *in vitro* data suggest that, if present, such a non-native

▶

Fig. 7. Structural effects of small chaperones bound to the P23H mutant opsin. (A) The native hub-links and their connected nodes are shown on the crystal structure of dark rhodopsin (PDB: 1U19). PSN analysis was carried out on the SMD trajectories of P23H mutant opsin bound to 11-*cis*-retinal (left) and 13-*cis*-5,8-ERA (right). As for the latter, the perturbations averaged over the PSNs of the two enantiomers in the two oligomeric models are shown here. The big spheres (centered on the Cα-atoms) are the 51 native stable hubs, whereas the small spheres are the nodes linked to such hubs. Each of the two structures shows the perturbation, expressed by coloring (see the color-legend bar), of the native hub-links. In detail, the color of each hub represents the percentage of perturbed hub-links (i.e. with reduced frequency with respect to the WT). Each link is colored according to its perturbation compared to the WT. Each non-hub node is colored according to the average perturbation of its links. (B) The 2D-network concerning the RBP native hubs and nodes linked to such hubs is shown for the simulated complexes between P23H opsin and 11-*cis*-retinal (left) or 13-*cis*-5,8-ERA (right). As for the latter, the perturbations averaged over the PSNs of the two enantiomers in the two oligomeric models are shown here. Native hubs are represented as pentagons. See panel A legend for coloring explanation. (C) The distribution of native stable hubs concerning the simulated complexes between P23H opsin and 11-*cis*-retinal (left) or 13-*cis*-5,8-ERA (right) is shown according to two coordinates: percentage of perturbed native links per hub and conservation grade of the amino acid residue behaving as a hub. For 13-*cis*-5,8-ERA, all native hubs of the two distinct enantiomers contributed to the surface.

disulphide bridge would be compatible with a protein state targetable to the plasma membrane even in the absence of 9-*cis*-retinal. Thus for cluster-1 mutants, which reach plasma membrane and show low ER retention either in the opsin or rhodopsin states, the chaperone effect of the chromophore, if any, is masked. The structural instability of the remaining 11 mutants considered in this study, almost all previously classified as misfolded [10–12] and ascribed to our cluster 4 of retinal non-responsive and ER-retained mutants, may be due to some deformation in the retinal binding site that prevents them from binding the chromophore. Yet, our hypothesis is that they represent quasi-native states. In summary, in the context of the adRP rod opsin mutants, the adjective misfolded should be interpreted as a folded but less stable state than the native one, due to structural deformations inside or outside the retinal binding site.

Future applications of the computational protocol include the structure-based optimization of the 13-*cis*-5,8-ERA chaperone and the discovery/design of non-retinoid small chaperones with translational potential as therapeutic agents for the cure of adRP linked to rod opsin mutations. The discovery/design of non-retinoid chaperones will take into account the information on two completely unrelated chemical scaffolds published almost simultaneously to our chaperone [35,36]. Furthermore, the approach is exportable to other conformational diseases linked to protein missense mutations.

Acknowledgments

Funding

Author contributions

A.F. implemented and performed the PSN analyses. D.S. and A.C. performed the *in vitro* experiments. V.M. designed the strategy for *in vitro* analyses. F.F. conceived the study, ran molecular simulations and performed structural analyses. A.F., V.M. and F.F. contributed to writing and editing the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.csbj.2021.10.040.

References

[1] Sanders CR, Myers JK. Disease-related misassembly of membrane proteins. Annu Rev Biophys Biomol Struct. 2004;33(1):25–51.
[2] Hartong DT, Berson EL, Dryja TP. Retinitis pigmentosa. Lancet 2006;368 (9549):1795–809.
[3] Ferrari S, Di Iorio E, Barbaro V, Ponzin D, Sorrentino FS, Parmeggiani F. Retinitis Pigmentosa: Genes and Disease Mechanisms. Curr Genomics 2011;12 (4):238–49.
[4] Kennan A, Aherne A, Humphries P. Light in retinitis pigmentosa. Trends Genet. 2005;21(2):103–10.
[5] Mendes HF, van der Spuy J, Chapple JP, Cheetham ME. Mechanisms of cell death in rhodopsin retinitis pigmentosa: implications for therapy. Trends Mol Med. 2005;11(4):177–85.
[6] Palczewski K. G protein-coupled receptor rhodopsin. Annu Rev Biochem. 2006;75(1):743–67.
[7] Fanelli F, De Benedetti PG. Update 1 of: Computational Modeling Approaches to Structure-Function Analysis of G Protein-Coupled Receptors. Chem Rev. 2011;111(12):PR438–535.
[8] Okada T, Ernst OP, Palczewski K, Hofmann KP. Activation of rhodopsin: new insights from structural and biochemical studies. Trends Biochem Sci. 2001;26 (5):318–24.
[9] Katritch V, Cherezov V, Stevens RC. Structure-function of the G protein-coupled receptor superfamily. Annu Rev Pharmacol Toxicol. 2013;53 (1):531–56.
[10] Briscoe AD, Gaur C, Kumar S. The spectrum of human rhodopsin disease mutations through the lens of interspecific variation. Gene 2004;332: 107–18.
[11] Krebs MP, Holden DC, Joshi P, Clark CL, Lee AH, Kaushal S. Molecular mechanisms of rhodopsin retinitis pigmentosa and the efficacy of pharmacological rescue. J Mol Biol. 2010;395(5):1063–78.
[12] Athanasiou D, Aguila M, Bellingham J, Li W, McCulley C, Reeves PJ, et al. The molecular and cellular basis of rhodopsin retinitis pigmentosa reveals potential strategies for therapy. Progress in Retinal and Eye Research. 2018;62:1–23.
[13] Lin JC, Liu HL. Protein conformational diseases: from mechanisms to drug designs. Curr Drug Discov Technol. 2006;3(2):145–53.
[14] Conn PM, Ulloa-Aguirre A. Trafficking of G-protein-coupled receptors to the plasma membrane: insights for pharmacoperone drugs. Trends Endocrinol Metab. 2010;21(3):190–7.
[15] Mendes HF, Cheetham ME. Pharmacological manipulation of gain-of-function and dominant-negative mechanisms in rhodopsin retinitis pigmentosa. Hum Mol Genet. 2008;17(19):3043–54.
[16] Fanelli F, Seeber M. Structural insights into retinitis pigmentosa from unfolding simulations of rhodopsin mutants. FASEB J. 2010;24(9):3196–209.
[17] Felline A, Seeber M, Rao F, Fanelli F. Computational Screening of Rhodopsin Mutations Associated with Retinitis Pigmentosa. J Chem Theory Comput. 2009;5(9):2472–85.
[18] Behnen P, Felline A, Comitato A, Di Salvo MT, Raimondi F, Gulati S, et al. A small chaperone improves folding and routing of rhodopsin mutants linked to inherited blindness. iScience. 2018;4:1–19.
[19] Vishveshwara S, Ghosh A, Hansia P. Intra and inter-molecular communications through protein structure network. Curr Protein Pept Sci. 2009;10(2):146–60.
[20] Seeber M, Felline A, Raimondi F, Mariani S, Fanelli F. WebPSN: a web server for high-throughput investigation of structural communication in biomacromolecules. Bioinformatics 2015;31(5):779–81.
[21] Seeber M, Felline A, Raimondi F, Muff S, Friedman R, Rao F, et al. Wordom: A user-friendly program for the analysis of molecular structures, trajectories, and free energy surfaces. J Comput Chem. 2011;32(6):1183–94.
[22] Felline A, Seeber M, Fanelli F. webPSN v2.0: a webserver to infer fingerprints of structural communication in biomacromolecules. Nucleic Acids Res. 2020;48 (W1):W94–W103.
[23] Amitai G, Shemesh A, Sitbon E, Shklar M, Netanely D, Venger I, et al. Network analysis of protein structures identifies functional residues. J Mol Biol. 2004;344(4):1135–46.
[24] Brinda KV, Vishveshwara S. A network representation of protein structures: implications for protein stability. Biophys J. 2005;89(6):4159–70.
[25] del Sol A, Fujihashi H, Amoros D, Nussinov R. Residues crucial for maintaining short paths in network communication mediate signaling in proteins. Mol Syst Biol. 2006;2006(2):0019.
[26] Bode C, Kovacs IA, Szalay MS, Palotai R, Korcsmaros T, Csermely P. Network analysis of protein dynamics. FEBS Lett. 2007;581(15):2776–82.
[27] Chennubhotla C, Bahar I, Levitt M. Signal propagation in proteins and relation to equilibrium fluctuations. PLoS Comput Biol. 2007;3(9):e172.
[28] Nussinov R, Tsai C-J, Csermely P. Allo-network drugs: harnessing allostery in cellular networks. Trends Pharmacol Sci. 2011;32(12):686–93.
[29] Pandini A, Fornili A, Fraternali F, Kleinjung J. Detection of allosteric signal transmission by information-theoretic analysis of protein dynamics. FASEB J. 2012;26(2):868–81.
[30] Papaleo E, Lindorff-Larsen K, De Gioia L. Paths of long-range communication in the E2 enzymes of family 3: a molecular dynamics investigation. Phys Chem Chem Phys. 2012;14(36):12515–25.
[31] Csermely P, Sandhu KS, Hazai E, Hoksza Z, Kiss HJ, Miozzo F, et al. Disordered proteins and network disorder in network descriptions of protein structure, dynamics and function: hypotheses and a comprehensive review. Curr Protein Pept Sci. 2012;13(1):19–33.
[32] Venkatakrishnan AJ, Deupi X, Lebon G, Tate CG, Schertler GF, Babu MM. Molecular signatures of G-protein-coupled receptors. Nature 2013;494 (7436):185–94.
[33] Felline A, Belmonte L, Raimondi F, Bellucci L, Fanelli F. Interconnecting Flexibility, Structural Communication, and Function in RhoGEF Oncoproteins. J Chem Inf Model. 2019;59(10):4300–13.
[34] Rader AJ, Anderson G, Isin B, Khorana HG, Bahar I, Klein-Seetharaman J. Identification of core amino acids stabilizing rhodopsin. Proc Natl Acad Sci U S A. 2004;101(19):7246–51.
[35] Chen Y, Chen Yu, Jastrzebska B, Golczak M, Gulati S, Tang H, et al. A novel small molecule chaperone of rod opsin and its potential therapy for retinal degeneration. Nature. Communications. 2018;9(1). https://doi.org/10.1038/s41467-018-04261-1.
[36] Mattle D, Kuhn B, Aebi J, Bedoucha M, Kekilli D, Grozinger N, et al. Ligand channel in pharmacologically stabilized rhodopsin. PNAS 2018;115 (14):3640–5.
[37] Liang Y, Fotiadis D, Filipek S, Saperstein DA, Palczewski K, Engel A. Organization of the G protein-coupled receptors rhodopsin and opsin in native membranes. J Biol Chem. 2003;278(24):21655–62.

[38] Okada T, Sugihara M, Bondar AN, Elstner M, Entel P, Buss V. The retinal conformation and its environment in rhodopsin in light of a new 2.2 A crystal structure. J Mol Biol. 2004;342(2):571–83.

[39] Fotiadis D, Liang Y, Filipek S, Saperstein DA, Engel A, Palczewski K. Atomic-force microscopy: Rhodopsin dimers in native disc membranes. Nature 2003;421(6919):127–8.

[40] Zhao DY, Pöge M, Morizumi T, Gulati S, Van Eps N, Zhang J, et al. Cryo-EM structure of the native rhodopsin dimer in nanodiscs. J Biol Chem. 2019;294 (39):14215–30.

[41] Seeber M, Fanelli F, Paci E, Caflisch A. Sequential unfolding of individual helices of bacterioopsin observed in molecular dynamics simulations of extraction from the purple membrane. Biophys J. 2006;91(9):3276–84.

[42] Davidson FF, Loewen PC, Khorana HG. Structure and function in rhodopsin: replacement by alanine of cysteine residues 110 and 187, components of a conserved disulfide bond in rhodopsin, affects the light-activated metarhodopsin II state. Proc Natl Acad Sci U S A. 1994;91(9):4029–33.

[43] Fahmy K, Jager F, Beck M, Zvyaga TA, Sakmar TP, Siebert F. Protonation states of membrane-embedded carboxylic acid groups in rhodopsin and metarhodopsin II: a Fourier-transform infrared spectroscopy study of site-directed mutants. Proc Natl Acad Sci U S A. 1993;90(21):10206–10.

[44] Stojanovic A, Hwang I, Khorana HG, Hwa J. Retinitis pigmentosa rhodopsin mutations L125R and A164V perturb critical interhelical interactions: new insights through compensatory mutations and crystal structure analysis. J Biol Chem. 2003;278(40):39020–8.

[45] Dunbrack RL, Karplus M. Backbone-dependent rotamer library for proteins. Application to side-chain prediction. J Mol Biol. 1993;230(2):543–74.

[46] Ponder JW, Richards FM. Tertiary templates for proteins. Use of packing criteria in the enumeration of allowed sequences for different structural classes. J Mol Biol. 1987;193(4):775–91.

[47] Sutcliffe MJ, Hayes FRF, Blundell TL. Knowledge based modelling of homologous proteins, Part II: Rules for the conformations of substituted sidechains. Protein Eng. 1987;1(5):385–92.

[48] Brooks BR, Brooks CL, Mackerell AD, Nilsson L, Petrella RJ, Roux B, et al. CHARMM: the biomolecular simulation program. J Comput Chem. 2009;30 (10):1545–614.

[49] Im W, Feig M, Brooks CL. An implicit membrane generalized born theory for the study of structure, stability, and interactions of membrane proteins. Biophys J. 2003;85(5):2900–18.

[50] Vishveshwara S, Brinda KV, Kannan N. Protein structure: insights from graph theory. J Theor Comput Chem. 2002;01(01):187–211.

[51] Kannan N, Vishveshwara S. Identification of side-chain clusters in protein structures by a graph spectral method. J Mol Biol. 1999;292(2):441–64.

[52] Ashkenazy H, Abadi S, Martz E, Chay O, Mayrose I, Pupko T, et al. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Res 2016;44(W1):W344–50.

[53] Angeletti B, Löster J, Auricchio A, Gekeler F, Shinoda K, Ballabio A, et al. An in vivo doxycycline-controlled expression system for functional studies of the retina. Invest Ophthalmol Vis Sci. 2003;44(2):755. https://doi.org/10.1167/iovs.02-0340.

[54] Dunn KW, Kamocka MM, McDonald JH. A practical guide to evaluating colocalization in biological microscopy. Am J Physiol Cell Physiol. 2011;300(4): C723–42.

[55] Li T, Sandberg MA, Pawlyk BS, Rosner B, Hayes KC, Dryja TP, et al. Effect of vitamin A supplementation on rhodopsin mutants threonine-17 –> methionine and proline-347 –> serine in transgenic mice and in cell cultures. Proc Natl Acad Sci U S A. 1998;95(20):11933–8.

[56] Noorwez SM, Malhotra R, McDowell JH, Smith KA, Krebs MP, Kaushal S. Retinoids assist the cellular folding of the autosomal dominant retinitis pigmentosa opsin mutant P23H. J Biol Chem. 2004;279(16):16278–84.

[57] Noorwez SM, Sama RRK, Kaushal S. Calnexin improves the folding efficiency of mutant rhodopsin in the presence of pharmacological chaperone 11-cis-retinal. J Biol Chem. 2009;284(48):33333–42.

[58] Kiser PD, Golczak M, Palczewski K. Chemistry of the retinoid (visual) cycle. Chem Rev. 2014;114(1):194–232.

[59] Palczewski K. Retinoids for treatment of retinal diseases. Trends Pharmacol Sci. 2010;31(6):284–95.

[60] Comitato A, Di Salvo MT, Turchiano G, Montanari M, Sakami S, Palczewski K, et al. Dominant and recessive mutations in rhodopsin activate different cell death pathways. Hum Mol Genet. 2016;25:2801–12.

[61] Sakami S, Kolesnikov AV, Kefalov VJ, Palczewski K. P23H opsin knock-in mice reveal a novel step in retinal rod disc morphogenesis. Hum Mol Genet. 2014;23(7):1723–41.

[62] Sakami S, Maeda T, Bereta G, Okano K, Golczak M, Sumaroka A, et al. Probing mechanisms of photoreceptor degeneration in a new mouse model of the common form of autosomal dominant retinitis pigmentosa due to P23H opsin mutations. J Biol Chem. 2011;286(12):10551–67.

[63] Chiang W-C, Kroeger H, Sakami S, Messah C, Yasumura D, Matthes MT, et al. Robust Endoplasmic Reticulum-Associated Degradation of Rhodopsin Precedes Retinal Degeneration. Mol Neurobiol. 2015;52(1):679–95.

[64] Kedrov A, Janovjak H, Sapra KT, Müller DJ. Deciphering molecular interactions of native membrane proteins by single-molecule force spectroscopy. Annu Rev Biophys Biomol Struct. 2007;36(1):233–60.

[65] Tanuj Sapra K, Park P-H, Filipek S, Engel A, Müller DJ, Palczewski K. Detecting molecular interactions that stabilize native bovine rhodopsin. J Mol Biol. 2006;358(1):255–69.

[66] Sapra KT, Park PS, Palczewski K, Muller DJ. Mechanical properties of bovine rhodopsin and bacteriorhodopsin: possible roles in folding and function. Langmuir 2008;24(4):1330–7.

[67] Herrera-Hernandez MG, Ramon E, Lupala CS, Tena-Campos M, Perez JJ, Garriga P. Flavonoid allosteric modulation of mutated visual rhodopsin associated with retinitis pigmentosa. Sci Rep. 2017;7(1):11167.

[68] Srinivasan S, Fernández-Sampedro MA, Morillo M, Ramon E, Jiménez-Rosés M, Cordomí A, et al. Human Blue Cone Opsin Regeneration Involves Secondary Retinal Binding with Analog Specificity. Biophys J. 2018;114(6):1285–94.

[69] Park PS, Filipek S, Wells JW, Palczewski K. Oligomerization of G protein-coupled receptors: past, present, and future. Biochemistry 2004;43 (50):15643–56.

[70] Chabre M, Deterre P, Antonny B. The apparent cooperativity of some GPCRs does not necessarily imply dimerization. Trends Pharmacol Sci. 2009;30 (4):182–7.

[71] Chabre M, le Maire M. Monomeric G-Protein-Coupled Receptor as a Functional Unit. Biochemistry 2005;44(27):9395–403.

[72] Dell'Orco D, Koch K-W. Transient complexes between dark rhodopsin and transducin: circumstantial evidence or physiological necessity? Biophys J. 2015;108(3):775–7.

[73] Chabre M, Cone R, Saibil H. Biophysics: is rhodopsin dimeric in native retinal rods? Nature 2003;426(6962):30–1. discussion 1.

[74] Jastrzebska B, Maeda T, Zhu L, Fotiadis D, Filipek S, Engel A, et al. Functional characterization of rhodopsin monomers and dimers in detergents. J Biol Chem. 2004;279(52):54663–75.

[75] Jastrzebska B, Fotiadis D, Jang G-F, Stenkamp RE, Engel A, Palczewski K. Functional and structural characterization of rhodopsin oligomers. J Biol Chem. 2006;281(17):11917–22.

[76] Bayburt TH, Leitz AJ, Xie G, Oprian DD, Sligar SG. Transducin activation by nanoscale lipid bilayers containing one and two rhodopsins. J Biol Chem. 2007;282(20):14875–81.

[77] Ernst OP, Gramse V, Kolbe M, Hofmann KP, Heck M. Monomeric G protein-coupled receptor rhodopsin in solution activates its G protein transducin at the diffusion limit. PNAS 2007;104(26):10859–64.

[78] Gragg M, Kim TG, Howell S, Park P- S-H. Wild-type opsin does not aggregate with a misfolded opsin mutant. Biochim Biophys Acta. 2016;1858(8):1850–9.

[79] Gragg M, Park P-H. Misfolded rhodopsin mutants display variable aggregation properties. Biochim Biophys Acta Mol Basis Dis. 2018;1864(9):2938–48.

[80] Hwa J, Klein-Seetharaman J, Khorana HG. Structure and function in rhodopsin: Mass spectrometric identification of the abnormal intradiscal disulfide bond in misfolded retinitis pigmentosa mutants. Proc Natl Acad Sci U S A. 2001;98 (9):4872–6.