# Effects of Pathogenic Variations in the Human Rhodopsin Gene (*hRHO*) on the Predicted Accessibility for a Lead Candidate Ribozyme

Beau R. Froebel,[1–3] Alexandria J. Trujillo,[1,2,4] and Jack M. Sullivan[1–7]

[1]Research Service, VA Western New York Healthcare System, Buffalo, New York, United States
[2]Department of Ophthalmology, State University of New York, University at Buffalo, Buffalo, New York, United States
[3]The Ross Eye Institute of University at Buffalo, Buffalo, New York, United States
[4]Department of Pharmacology/Toxicology, State University of New York, University at Buffalo, Buffalo, New York, United States
[5]Department of Physiology/Biophysics, State University of New York, University at Buffalo, Buffalo, New York, United States
[6]Neuroscience Program, State University of New York, University at Buffalo, Buffalo, New York, United States
[7]The RNA Institute, University at Albany-State University of New York, Albany, New York, United States

Correspondence: Jack M. Sullivan, Pharmacology/Toxicology, and Physiology/Biophysics, University at Buffalo, The Ross Eye Institute, Veterans Administration Western New York Healthcare System, Medical Research, Building 20, Room 245, 3495 Bailey Avenue, Buffalo, NY 14215, USA;
js354@buffalo.edu.

**PURPOSE.** The mutation-independent strategy for hammerhead ribozyme (hhRz) or RNA interference (RNAi)-based gene therapeutics to treat autosomal dominant diseases is predicated on the hypothesis that a single therapeutic would equivalently suppress all/most of the diverse mutant mRNAs in patients with the disease phenotype. However, the hypothesis has not been formally tested. We address this through a comprehensive bioinformatics study of how mutations affect target mRNA structure accessibility for a single lead hhRz therapeutic (725GUC↓), designed against human rod rhodopsin mRNA (*hRHO*), for patients with *hRHO* mutations that cause autosomal dominant retinitis pigmentosa.

**METHODS.** A total of 199 in silico coding region mutations (missense, nonsense, insert, deletion, indel) were made in *hRHO* mRNA based on Human Gene Mutation Database and Database of Single Nucleotide Polymorphisms. Each mRNA was folded with MFold, SFold, and OligoWalk algorithms and subjected to a bioinformatics model called multiparameter prediction of RNA accessibility. Predicted accessibility of each mutant over both a broad local region and the explicit lead ribozyme annealing site were compared quantitatively to wild-type *hRHO* mRNA.

**RESULTS.** Accessibility of the 725GUC↓ site is sensitive to some mutations. For single nucleotide missense mutations, proximity of the mutation to the hhRz annealing site increases the impact on predicted accessibility, but some distant mutations also influence accessibility.

**CONCLUSIONS.** A mutation-independent strategy appears viable in this specific context but certain mutations could significantly influence ribozyme or RNAi efficacy through impact on accessibility at the target annealing site/region. This possibility must be considered in applications of this gene therapy strategy.

Keywords: gene therapy, ribozyme, post-transcriptional gene silencing agents, mutation-independent strategy, rhodopsin

Among the many known disease genes (current number: 22; RetNet, https://sph.uth.edu/retnet/sum-dis.htm) that cause autosomal dominant retinitis pigmentosa (adRP), the first gene identified and source of much ongoing investigation is disease-promoting mutations in the human rod rhodopsin mRNA (*hRHO*) gene (Online Mendelian Inheritance of Man: 180380).[1,2] Rhodopsin, the visual pigment of the rod photoreceptor, constitutes up to 90% of the total protein in the rod outer segment where phototransduction is initiated. A functional *RHO* protein depends on correct folding, post-translational modifications, cell trafficking, and interaction with 11-*cis*-retinal to form a ground state visual pigment that can be activated by light. Mutant *RHO* proteins frequently do not properly fold into functional conformations, blocking mature visual pigment formation.[3,4] This can activate cell stress machinery to promote apoptotic photoreceptor cell death.[5] This results in the severe, progressive vision impairment that characterizes adRP. The rationale for gene therapy to treat adRP and other autosomal dominant retinal degenerative diseases is to develop therapeutic agents capable of reducing the levels of toxic mutant proteins in the cell, allowing cells to maintain viability and (ideally) function to support vision longer into the life of the affected patient. As gene therapy studies to treat dominant retinal degenerations, and in particular *hRHO* mutations, are moving toward clinical trial, this study is both relevant and timely.

The retina has often been selected as a focus for the development of post-transcriptional gene silencing (PTGS) agents. Among these agents are ribozymes, small catalytic RNAs that on binding a target mRNA, undergo a conformational change, and cleave the target at NUH sites (N = any nucleotide,

U = U, H = C, U, A). This leads to increased degradation of targeted mRNA and decreased translation (expression) of the protein encoded by it. Although ribozymes are catalytic as RNA elements, short hairpin RNAs (shRNAs) anneal to target mRNAs and recruit host cell proteins to cleave the target mRNA. As gene therapy agents, hammerhead ribozyme (hhRzs) and shRNAs act in trans and interact with a target mRNA by a second-order biochemical reaction based on concentration-dependent biophysical collision frequency. The hhRz or shRNA must colocalize with target mRNA and generally be in concentrative excess to anneal and then cleave the target.

Challenges arise while developing PTGS agents against target mRNAs, reflected in the very small number of agents that successfully passed clinical trials and are in use.[6] A major barrier and bottleneck to the development of PTGS agents is the dense structure of any target mRNA, because the first limiting step in PTGS is the ability of any agent to anneal to the target mRNA, which is severely limited by structural inaccessibility.[7] Most potential hhRz or shRNA cleavage sites in a target mRNA are inaccessible due to dense and stable secondary folding, tertiary folding, protein coating, and dynamical fluctuations.[5-10] Successful annealing requires the target region to be largely single stranded under physiologic conditions. Such sites in target mRNAs are rare.

Our focus is on hhRzs as therapeutics because they are likely to have greater specificity than shRNAs, which have extensive off-target effects that are potentially toxic in the therapeutic context.[11-14] Because of this, *hRHO* has been targeted in multiple gene-based therapy development studies attempting to suppress its expression in the context of adRP with the use of ribozymes.[5,10,15-19]

Due to the increasingly vast number of mutations that may be found in a given autosomal dominant disease gene, strategies have been developed to target mutant genes or their mRNAs that are mutation independent.[6,15,16] Rather than attempt to make a new therapeutic for each individual mutation, which is both fiscally and biologically impractical (i.e., most single nucleotide [nt] mutations are buried in the mRNA structure and inaccessible), a single therapeutic is sought that can be used to target and suppress all or most known mutant mRNAs/proteins in a given autosomal dominant disease gene. For instance, if one identifies a highly accessible site in a target mRNA that can be attacked by a ribozyme or other PTGS agents (e.g., Rz725GUC↓), then that therapeutic might be useful to treat most/all known mutations as long as the mutations do not profoundly affect the accessibility of the target site to therapeutic RNA annealing. Because such a therapeutic will also suppress the wild-type (WT) mRNA, the knockdown agent must be combined with an allelic variant expression construct that encodes the WT protein through an mRNA that cannot be cleaved by the given therapeutic (e.g., silent mutation for valine of in-frame cleavable 725GUC↓ to noncleavable 725GUG). We call this strategy "combined knockdown and reconstitute" (CKDRT).[6]

*hRHO* mRNA has multiple pathogenic and tabulated variants thought to be pathogenic (The Human Gene Mutation Database [HGMD]; http://www.hgmd.cf.ac.uk).[20-22] No study to date has used the rich repertoire of known *hRHO* mutations, or in fact diverse mutations in any gene, to address the question of the viability of a CKDRT approach to treat adRP as a model autosomal dominant retinal degeneration. In this study, we use an approach to map mRNA target accessibility (developed in-house), which we call multiparameter prediction of mRNA accessibility (mppRNA). We used mppRNA to successfully predict potential ribozyme cleavage sites in *hRHO* in silico, which has identified lead hhRzs candidates with intracellular efficacy.[5,10] MFold is used to predict the most stable secondary structure of a target mRNA through a free



**Figure 1.** Lead candidate ribozyme. Hammerhead ribozyme developed by Yau et al. to target *hRHO* hhRz cleavage site 725GUC↓. The lead candidate is composed of 7-nt upstream and downstream antisense flanking regions and a 4-bp length helix II capped by a GAAA tetraloop. This ribozyme was selected among 32 tested ribozymes for cleavage in cellulo via secreted alkaline phosphatase assay as previously described.[10]

energy minimization approach. The single structure with the minimum folding energy (MFE) and an ensemble of structures within a designated degree of higher free energy (less stable) in the folding neighborhood are identified.[23] SFold uses a Boltzmann-weighted sampling of all folding space to estimate the true probability of single stranded-ness at each nucleotide position.[24] Another algorithm, OligoWalk (OW), available in RNAStructure, uses the output of MFold and calculates the ribozyme annealing and local structure breakage energies with the concept that a more accessible site is likely to exist as a single-stranded, unstructured form with less energy.[25] We take the vectorial outputs of MFold, SFold, and OW and combine them in a simple multiplicative model (set of intersection) to achieve an accessibility map along the target mRNA that has proven versatility.[5,10]

Our lead hhRz candidate against *hRHO*, Rz725GUC↓ (Fig. 1), targets the GUC↓ cleavage motif at nucleotide 725, which resides in a local region of high-weighted accessibility in *hRHO*. This lead candidate was identified by first using mppRNA to map accessibility followed by a high throughput screen (HTS) to test a multitude of hhRz expression constructs in cellulo.[10] mppRNA has been successfully used with *hRHO* to screen for available target sites for ribozyme cleavage.[5,10] Our studies illustrate there is a relationship between mppRNA predicted accessibility and the ability for a ribozyme to cleave in vitro and in cellulo.

The human *hRHO* gene is subject to hundreds of pathogenic changes, including single nucleotide missense/nonsense; small insertions, deletions, or indels and gross insertions or deletions. Most of these mutations result in adRP.[2,26] It is established that a mutation within the target binding site of a ribozyme can decrease its efficacy by either altering its rate of binding or release of products from the antisense flanking regions or directly altering its enzymatic ability to cleave at the target if the mutation affects the NUH site directly.[27-30] Here, we took advantage of our predictive mppRNA model to explore the possibility that pathogenic missense mutations, nonsense mutations, insertions, or deletions can lead to a sufficient structural change in a target mRNA that could perturb accessibility for our lead PTGS agent (Rz725GUC↓). Because prospective patients who would be receiving a therapy of this sort will have varied mutations in

**FIGURE 2.** Distribution of known *hRHO* mutations. An illustration of the types and locations of 199 mutations surveyed for predicted cleavage accessibility via mppRNA analysis. All known mutations are located in the coding region of the mRNA. The types of mutations are color coded. Sites with multiple surveyed mutations are represented as having an increased frequency. For mutants affecting multiple nucleotides, the first nucleotide affected was depicted. Mutations were selected from HGMD and NCBI's dbSNP. Mutations including single nucleotide point mutations, small and large insertions, deletions, or indels were sampled. The common P23H single nucleotide missense mutation responsible for most cases of adRP in the United States, as well as six single nucleotide missense mutations causing mutations at amino acid P347, a severe form of adRP, were surveyed.

the gene being suppressed, understanding the effect of mutations on the ability of the single agent to knock down levels of mutant mRNA is highly important for the eventual clinical application of PTGS agents for adRP or other autosomal dominant diseases.

## MATERIALS AND METHODS

### Mutants and Ribozyme

The in silico *hRHO* mRNA transcript used was a 1532-nt mRNA originally cloned and sequenced by Nathans and Hogness (GenBank Entry NM_000539.3).[26] The transcript has a 95-nt 5′untranslated region (UTR), full coding region (1044 nt), stop codon (UAA), and a 394-nt 3′UTR that contains the first dominant poly-A signal (nt 1506) and continues 26 nt past the polyA signal to nt 1532 at the approximate site where 3′ cleavage and polyA addition would naturally occur. PolyA tails (~200 to 250 nt) have no structure and therefore are not included in predictive folding experiments. This 1532-nt sequence is the dominant form of polyadenylated *hRHO* mRNA made in vivo.[26] The dominant mature *hRHO* mRNA is approximately 1.7 to 1.8 kb due to unstructured polyadenylation at the 3′ end. Nondominant transcripts exist but were not used for in silico analysis in this study. All *hRHO* mutations were made in silico in the 1532-nt dominant form of polyadenylated *hRHO* mRNA.

The lead ribozyme candidate for an mutation-independent (MI) CDKRT lead gene therapy strategy targets the GUC↓ site at position 725 in the coding region of the *hRHO* transcript and is composed of 7-nt antisense flanking regions on each side of the cleavage (nonhybridizing) nt (C of GUC↓), and a 4-bp-length Stem II helix capped by a GAAA tetraloop (Fig. 1).[10] This agent was the best-performing hhRz among 32 tested ribozymes for cleavage in cellulo in a HTS in which the hhRzs were embedded in an engineered high-flexibility region of the adenoviral viral-associated RNA I (VAI) scaffold RNA, demonstrated as a 60% mRNA knockdown measured by secreted alkaline phosphatase reporter assay.[10] The NUH↓ sites selected for screening in the prior studies were chosen based on their predicted accessibility (rank ordered) in the *hRHO* mppRNA map. The 725 region had the highest level of predicted accessibility, which correlated with the level of experimental target suppression.

A total of 199 *hRHO* mutations in the coding region of the mRNA (nts 96 to 1142) were surveyed in this study and were found in the HGMD, National Center for Biotechnology Information's (NCBI) Database of Single Nucleotide Polymorphisms (dbSNP), with one novel mutation characterized from an adRP patient seen in Buffalo, NY (Sullivan et al., unpublished observations, 2016).[20,22] All known single nucleotide missense and nonsense mutations found on HGMD that were pathogenic

for adRP were used in this study (Fig. 2). Notably, we evaluated six missense mutations at the known hotspot amino acid residue P347, which all cause severe early-onset adRP and the P23H mutation responsible for the most common form of *hRHO* adRP in the United States.[2,31–33] A smaller number (15) of nonpathogenic single nucleotide polymorphisms (SNPs) within 200 nts of the target nucleotide 725 were also sampled from NCBI's dbSNP that did not appear on the HGMD list to evaluate the impact of local or regional variations on PTGS target accessibility. Pathogenic small insertions, small deletions, small indels, and large insertions or deletions were also explored, totaling 24. Insertions, deletions, and indels were generated from those listed at HGMD.[34,35]

### mppRNA Approach to Target mRNA Accessibility Mapping

A mutant *hRHO* sequence was created for each of the mutations from the 1532-nt WT *hRHO* mRNA using the information in the databases and the source papers and folded using three secondary structure predictive algorithms. All foldings were done under standard conditions commonly used for these algorithms (37°C, 1 M salt). Briefly, the first algorithm, MFold (http://mfold.rna.albany.edu/), which examines the frequency of being single stranded in all structures in the acquired ensemble, uses energy minimization to predict the MFE structure and an ensemble of structures of higher (less stable) free energies within a set criterion range of energy difference from the MFE structure (parameters used: 10% window of free energy from the MFE structure, difference window size = 3, up to 100 structures in the ensemble).[23] The MFE secondary structural fold generated from MFold for WT *hRHO* shows the dense secondary structure of a typical target mRNA for PTGS therapeutic development (Fig. 3). The output of MFold used for mppRNA is the normalized ssCount frequency vector (Fig. 4A), which represents the percentage of predicted structures where each nucleotide is single stranded. This is a biased probability estimate of accessibility only in the structural neighborhood of the MFE. The second algorithm, SFold (http://sfold.wadsworth.org/), uses a Boltzmann-weighted sampling of the total structural state ensemble and outputs a direct true estimate of the probability of being single stranded along the target string over the entire folding space.[24] The output of SFold used for mppRNA is the sStrand vector (Fig. 4B). Another output of SFold shows how the sample of structures (1000/sample) clusters into discrete distributions with repeatable probability densities. A method of displaying such distributions is through multidimensional scaling (MDS), which in effect collapses the three-dimensional (3D) map of location in the space (base pairing differences) versus free energy into a plane with unitless dimensionless axes that displace the distribution of structures.[36,37] MDS is a

## *hRHO* mRNA (1532 nt)



FIGURE 3. Example of 2D fold of WT *hRHO* mRNA generated using MFold. (A) Full 1532-nt hRHO mRNA. The MFE structure is shown. (B) Expanse of the 725 region. *Blue arrow* indicates the 725GUC↓ cleavage site.

**FIGURE 4.** Independent accessibility maps. (**A**) MFold, (**B**) SFold, and (**C**) OligoWalk outputs. An mppRNA map is generated by multiplying probabilities of each output for each nucleotide position. Large span representing the local regional integration (*blue bar*) (640 to 764 nt) and hhRz anneal site representing the ribozyme site integration (*red bar*) (718 to 732 nt) are indicated.





**FIGURE 5.** *hRHO* mppRNA full map. (**A**) Full-length mppRNA map of WT *hRHO*, using vector outputs from MFold, SFold, and OligoWalk. (**B**) Expanse of the local 725 region. The 725GUC↓ cleavage site is located in a region of high predicted accessibility, identified as a mesa or elevated plateau in the mppRNA map. *Blue bar* (near abscissa) represents the local regional integration from nts 640 to 764. *Red bar* represents the ribozyme binding site integration from nts 718 to 732.

way of looking at how single nucleotide polymorphisms can affect RNA structure.[38–42] The third algorithm requires the RNAStructure package (http://rna.urmc.rochester.edu/RNAstructure.html). RNAStructure has a version of MFold that generates an ensemble of structures. The OW algorithm uses the MFold output and calculates the local target unfolding energy (target break energy) for overlapping windows (of arbitrary size) of RNA over the entire length of mRNA.[25] The 15-nt window was used in OW to represent the span of annealing for a standard hhRz (including the Rz725GUC↓) with 7-nt antisense flanks on either side of the cleavage nucleotide (C↓; Fig. 1), which does not base pair with the hhRz. The OW output (.rep file) displays the free energy (kcal/mol) of the unfolding energy for each overlapping 15-nt segment (1-nt

steps) along the entire length of the mRNA (Fig. 4C); it is already a filtered version because of the stipulated window size for a standard symmetrical hhRz. We conduct a linear transform to convert this energy vector into a unitless positive probabilistic scale by adding a scalar constant to all centered points and then normalizing the output by the same constant. The constant chosen was 41.6 kcal/mol or the absolute value of the total free energy (−41.6 kcal/mol) of the RNA sequence 5′-GGGGGGGGGGGGGGGGGAAAAACCCCCCCCCCCCCCC-3′ (calculated via RNAStructure). This sequence folds into only one predicted low energy structure, a stem loop of 15 high energy base pairs (G≡C), connected by a small single stranded loop of five A residues. The total energy of this structure was used for normalization of the OW output, as it represents the maximum folding free energy of a 15-nt stable segment of folded mRNA (such a sequence does not occur in *hRHO*, but this is used because it is a normalization parameter that can extend beyond this particular target in future comparisons). In doing this normalization, we convert the free energy scale into a unitless scale between 0 and 1, which is an additional probability estimator of accessibility (local high free energy regions are likely to be single stranded).

The raw MFold and SFold outputs are then smoothed by a 15-point averaging window to substantially decrease probability of a 0 accessibility score at one or a few nucleotides in one vector output from being projected into the final mppRNA multiplicative product vector output. The 15-point smoothing also allows the MFold and SFold outputs to represent, nucleotide for nucleotide, the accessibility of the same 15-nt segment as the OW output along the entire length of the mRNA. The 15-nt smoothed vector outputs of each of these

three algorithms ($0 \leq X \leq 1$) are then multiplied together, with current assumption of equivalent weighting among the output vectors, and an mppRNA map is generated that represents the intersection access probability at each nt residue (Fig. 5). mppRNA seeks regions that are accessible as assessed by all three algorithms (MFold, SFold, and OW) (intersection set of the three probability estimators). Note the broad local region of accessibility (640 to 764 nts) that includes the targeting site for Rz725GUC↓ (718 to 732 nts).

By integrating the area under the curve of the 15-nt span of the mppRNA map centered at the site of cleavage, one can generate a quantitative scalar value that represents the overall predicted accessibility for the explicit target region, referred to here as the ribozyme site integration (718 to 732 nt). mppRNA was performed on WT *hRHO* and each mutant, and the site integration was determined for each mutant. For mutants that cause the numbering to shift, the area around the residue that corresponds to nt 725 in WT *hRHO* was integrated. A larger integration of the entire peak representing the broad area of increased accessibility, referred to here as the local integration (640 to 764 nt), was performed to assess the impact of the mutations on a broader area of the mRNA secondary structure. Because it has been shown that efficacy of ribozymes against *hRHO* correlates to the areas of peaks in the mppRNA map, it is possible to make judgements about altered efficacy of our lead hhRz among various mutant genotypes.[5,10]

## Quantitative Analysis

Statistical and graphical analysis was conducted in Origin (Versions 6.1, 8.1, and 2016; MicroCal, Northampton, MA, USA) using ANOVA and post hoc *t*-tests for between-condition comparisons with a previously chosen significance level of 0.05. Tests for uniformity of variance and normal distribution of data were conducted (uniformity of variance: Levene's, Shapiro's, Wilk's; Gaussian distribution: Anderson-Darling, Shapiro-Wilk, Kolmogovov-Smirnov). Post hoc parametric Student *t*-test and Welch's *t*-test were conducted. Welch's *t*-test does not depend on equivalence of variance and is robust to non-Gaussian distributions. Degrees of freedom (df) are shown where pertinent.

## RESULTS

### Single Nucleotide Missense/Nonsense Mutations

The two integrations in both a global (total mRNA) and a local mppRNA map of the WT *hRHO* mRNA are displayed in Figure 5. The global mppRNA map (Fig. 5A) shows that accessibility varies considerably throughout *hRHO* mRNA. The most accessible region (peak with the greatest integration weight) of the mppRNA map is broad and covers the span of nucleotides 640 to 764, which constitutes the local integration region that is illustrated by the blue bar underneath the expanded region of the global map (Fig. 5B). Within the local area is the actual 15-nt target annealing site (718 to 732 nts) in *hRHO* mRNA for our lead hhRz (Rz725GUC↓; Fig. 5B, red bar). Both the local and annealing site integral values from the mppRNA map were analyzed in the context of mutational effects on accessibility for a single lead hhRz candidate therapeutic (Rz725GUC↓).

A total of 176 in silico transcripts of single nucleotide point *hRHO* mutations were analyzed for impact on accessibility within the local and 725 hhRz annealing site spans in the mppRNA maps. The data for the MFold, SFold, OW, and mppRNA maps for the WT and each *hRHO* mutant are provided (Supplementary Table S1). Qualitatively, the region



**FIGURE 6.** Distribution of accessibility relative to WT *hRHO*. (**A**) Local integration and (**B**) site integration. Relative accessibilities were found by normalizing the area under the curve (integral) for each mutant by the area under the curve for the WT mppRNA over the appropriate span of sequence (mutant accessibility/WT accessibility). Local integration shows minimal influence of mutation on structure. A *line* (*red*) could be fit through the data with essentially zero slope and a y-intercept of 1 (slope: 2.9776E-6 ± 1.10249E-5; intercept: 0.98886 ± 0.00773). Site integration indicates a substantially greater impact of mutation. An insulator sequence is observed from approximately nts 500 to 600, flanked by regions of high influence on accessibility at the 725 region. Another region of high accessibility is found at the end of the coding region (1114 to 1140; coding region is 96 to 1142).

of accessibility around the 725 cleavage site was maintained in most single nucleotide mutations (Supplementary Table S2). Most single nucleotide mutations exerted little effect on either local or site accessibility relative to WT *hRHO* mRNA (17 of 176 [9.7%] showed a statistically significant difference in site accessibility by *t*-test; one showed a statistically significant difference in the local accessibility by *t*-test [U755G, relative integration 0.84803, $t(248) = -2.28418$; $P = 0.02321$; Supplementary Table S2). The distribution of predicted accessibility for each of the 176 single nucleotide missense/ nonsense mutations is shown (Fig. 6). Accessibility within or around the 725 hhRz target site could theoretically increase or decrease due to a proximate or remote mutation in the *hRHO* mRNA. The local accessibility of the mutants remained within ±15% of WT and a line with nonsignificant slope and y-intercept approximating 1 could be readily fit through the data (Fig. 6A). Sequence regions where mutations show the largest impact on 725 accessibility are within the actual hhRz binding site and those flanking an insulator sequence from nucleotides 500 to 600. Mutations within the insulator region show low impact on accessibility in the 725 region. An area at the tail end

of the coding region (nts 1116 to 1129) also has a high number of mutations with some impact on 725 accessibility; however, none of these single nucleotide mutations exerted a statistically significant shift in accessibility in the local integration analysis (Supplementary Table S2). In contrast, the hhRz annealing site accessibility around 725 was substantially impacted by mutational variation, with a range from +28.2% to −67.9% (Fig. 6B; + values indicate enhanced accessibility and − values indicate decreased accessibility). Not surprisingly, the largest areas of change are mutations immediately within and bracketing the annealing sequence of the Rz725GUC↓ cleavage site. There are no known mutations at the 725 cleavage nucleotide reported to date. Several mutations (G720A, G723U, and A727U) within the 725 annealing site region exert a statistically significant difference on the integrated site accessibility values. This effect may be nucleotide-specific because an alternative mutation, A727G, does not exert a significant change in site accessibility. There were significant changes in annealing site accessibility from single nucleotide mutations surrounding the hhRz annealing span (U715A, U715G, C738A, A741U, and U742G) and from single nucleotide mutations bracketing the insulator region (U469G, C488G, C498U, U639A, U654G, G655A, G658A, and C673U). There is a consistent relationship between the proximity of the *hRHO* mutation and its impact on accessibility at the site of hhRz cleavage (Table 1; Fig. 6). Other *hRHO* mutants that promoted statistically significant change in accessibility at the 725 annealing site include, surprisingly, the single nucleotide mutation U355A, which is remote from the 725 region, and the large insertion mutation, 1032ins150del8, which may be expected to largely alter overall secondary structure of the *hRHO* mRNA fold.

Table 1 contains noteworthy examples of single nucleotide mutants that have different effects on the predicted accessibility of the mRNA, chosen based on the magnitude of changes in accessibility, shape of the local/site mppRNA map, statistical significance, and clinical relevance. Some mutants had considerable differences of predicted accessibility in either one or both regions of integration seemingly dependent on the mutation's proximity to the 725 residue. Maps of the single nucleotide mutations with the largest relative changes from WT *hRHO* are presented (Fig. 7). Mutation A664G (Asp190Gly) shows the second highest local accessibility (112.4% relative to WT, $t[df = 248] = 1.4189$; $P = 0.15718$). The site accessibility also increases but is also not statistically significant (116.9% relative to WT, $t[28] = 1.42455$; $P = 0.16534$). A664G is located in the large region of susceptibility just after the insulating region. Likewise, mutation U755G (Phe220Leu) shows the largest significant decrease (84.8% of WT, $t[248] = −2.28418$; $P = 0.02321$) in accessibility of the local integration. The site accessibility change shows a nonsignificant increase (114.9% of WT, $t[28] = 1.64737$; $P = 0.11066$). The distant mutation, G1116A (Glu341Lys), has the greatest but statistically nonsignificant increase on the local integration (113.3% of WT, $t[248] = 1.64317$; $P = 0.10162$). The site integration for G1116A shows a nonsignificant increase in accessibility (114.5% of WT, $t[28] = 1.05567$; $P = 0.30014$). The site accessibilities of the mutants were subject to larger changes than the local accessibilities relative to WT. The largest statistically significant variations occurred when the mutation was in close proximity to the site of ribozyme annealing and cleavage. Two mutations illustrate this: G720A (Val209Met), which has the highest site accessibility of all mutants (128.2% of WT, $t[28] = 2.39529$; $P = 0.02354$), and G723U (Val210Phe), which has the lowest of all (32.1% of WT, $t[28] = −7.89601$; $P = 1.34E-8$). Although these mutations are within the antisense binding region of our ribozyme, we acknowledge the predicted accessibility measured here does not address the potential impact of mutation

on the thermodynamics and kinetic behavior of 725 hhRz due to a mismatch of base pairing in helices I and III.[43] These examples likely are representative of potential unreported mutations.

The 725 target site in *hRHO* resides approximately at 50% of the length of the mRNA. We explored adRP mutations near the (remote) 5′ and 3′ ends of the coding region. The mutation C163A causes Pro23His, the most common mutation causing adRP in the United States. The predicted accessibilities, both local and site, for this mutation did not differ significantly from that of WT *hRHO* (Table 1; Fig. 8A) (local accessibility, 102.7% of WT, $t[248] = 0.35738$; $P = 0.72111$; site accessibility, 107.2% of WT, $t[28] = 0.57014$; $P = 0.57313$). Multiple mutations at codon 347, including Pro347Ser, are responsible for a severe early-onset and rapidly progressive form of adRP. These P347X mutations (P347A, P347L, P347Q, P347R, P347S, and P347T) occur through changes at nts 1134 and 1135 as shown in Table 1. The mutation Pro347Ser (C1134U) showed the largest decrease in relative site accessibility of the six mutations (84.3% of WT, $t[28] = −1.80676$; $P = 0.08156$; Fig. 8B). The mutant C1134A responsible for Pro347Thr also showed a nonsignificant decrease (88% of WT, $t[28] = −1.32941$; $P = 0.19444$). Other P347X mutants did not show changes in relative site accessibility substantially different from *hRHO*. The effects of P347X mutations on local accessibility quantification show values within 3% of WT that were not significant. At distances of more than 400 nts, different single nucleotide mutations can promote some changes in the accessible architecture of a targeted region, but none have proven statistically significant.

To further evaluate quantitative accessibility differences between WT and mutant *hRHO* mRNAs, the integrated vector weights from the site and local regions surrounding the 725GUC↓ cleavage site were compared using a two-way *t*-test (Supplementary Table S2). Of the 199 mutations, 18 (9.0%) lead to statistically significant changes in the site 725 accessibility (17 missense mutations and one large insertion [1032ins150del8]). For the local integration around 725, there are only two mutations (U755G and 1032ins150del8, 1.0%), which promote a statistically significant change in accessibility.

Because only a relatively small subset of mutations could affect structural accessibility in *hRHO* targeting by our lead 725 hhRz PTGS agent, this is evidence in support of the MI-CKDRT therapeutic strategy for this construct.

To explore the extent to which the site accessibility is related to the local accessibility, a linear regression was performed on the accessibility predicted for each mutant by each method and plotted as site versus local integration. A significant positive correlation was observed ($R = 0.51531$, $P < 0.0001$), suggesting that the site and local accessibility around nt 725 follow similar trends when affected by a given mutation (Fig. 9). This is consistent with the idea that RNA folds locally and that the annealing site of the 725 hhRz is a subset (nested within) of the larger local regional accessible region.

We further investigated the class of mutants that exert statistically significant effects and the class of mutants that do not exert statistically significant effects for their impact on the 725 region (Fig. 10). The mppRNA map of the 725 region of the WT *hRHO* mRNA is shown for comparison (Fig. 10A). The average mppRNA map in the 725 region of all of the *hRHO* mutants that exert statistically significant effects demonstrates substantial differences relative to the WT (Fig. 10B). The average mppRNA map in the 725 region of all of the *hRHO* mutants that do not exert statistical significant effects shows similarity to the WT *hRHO* mRNA map (Fig. 10C). The average site integrals for both classes were compared with WT and the mean of only the statistically significant class was different from WT (Fig. 10D).

Investigative Ophthalmology & Visual Science

TABLE 1. Accessibility of Selected Missense Mutations in bRHO mRNA

| Mutation | AA Change | Relative Local Accessibility | Local T | Local Probability | df Students Local | Relative Site Accessibility | Site T | Site Probability | DF Students Site | Significance | Welch's Local Probability | DF Welch's Local | Welch's Site Probability | df Welch's Site | Significance |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Statistically significant mutations** | | | | | | | | | | | | | | | |
| U355A | Val87Asp | 0.90677 | 1.2809 | 0.20143 | 248 | 0.78749 | | **0.02846** | 28 | Yes | 0.20143 | 247.11626 | **0.02939** | 25.05657 | Yes |
| U469G | Leu125Arg | 0.91411 | 1.16052 | 0.24695 | 248 | 0.711 | 3.22343 | **0.00321** | 28 | Yes | 0.24695 | 247.37373 | **0.00332** | 26.74611 | Yes |
| C498G | Arg135Gly | 0.86431 | 1.72835 | 0.08517 | 248 | 0.4141 | 7.58253 | **2.93E-08** | 28 | Yes | 0.08518 | 246.1811 | **4.49863E-07** | 18.41736 | Yes |
| C498U | Arg135Trp | 0.8747 | 1.64221 | 0.10181 | 248 | 0.52879 | 6.19254 | **1.09E-06** | 28 | Yes | 0.10181 | 247.71518 | **9.52704E-06** | 17.12176 | Yes |
| U639A | Gly182Ser | 0.89917 | 1.32067 | 0.18783 | 248 | 0.76258 | 2.47618 | **0.01959** | 28 | Yes | 0.18783 | 247.77943 | **0.01963** | 27.83217 | Yes |
| U654G | Cys187Gly | 0.9426 | 0.72191 | 0.47103 | 248 | 0.3646 | 8.01767 | **9.89E-09** | 28 | Yes | 0.47104 | 244.90456 | **7.23E-08** | 21.27865 | Yes |
| G655A | Cys187Tyr | 0.94704 | 0.73536 | 0.46281 | 248 | 0.74415 | 2.87015 | **0.00773** | 28 | Yes | 0.46282 | 246.31445 | **0.00796** | 26.52961 | Yes |
| G658A | Gly188Glu | 1.0268 | −0.36523 | 0.71525 | 248 | 0.70768 | 3.1712 | **0.00366** | 28 | Yes | 0.71525 | 247.56162 | **0.00372** | 27.3687 | Yes |
| C673U | Thr193Met | 0.91983 | 1.01484 | 0.31117 | 248 | 0.75762 | 2.11463 | **0.04349** | 28 | Yes | 0.31118 | 245.51325 | **0.0453** | 23.44325 | Yes |
| U715A | Met207Lys | 0.91535 | 0.979 | 0.32853 | 248 | 0.43892 | 5.23288 | **1.47E-05** | 28 | Yes | 0.32859 | 233.54038 | **1.66E-05** | 26.8037 | Yes |
| U715G | Met207Arg | 1.0389 | −0.45103 | 0.65236 | 248 | 0.59593 | 4.12343 | **3.02E-04** | 28 | Yes | 0.65239 | 233.87996 | **3.02E-04** | 27.99794 | Yes |
| G720A | Val209Met | 1.00175 | −0.02324 | 0.98147 | 248 | 1.28201 | −2.39529 | **0.02354** | 28 | Yes | 0.98147 | 247.94813 | **0.02423** | 25.59975 | Yes |
| G723U* | Val210Phe | 0.91998 | 0.97833 | 0.32886 | 248 | 0.32131 | 7.89601 | **1.34E-08** | 28 | Yes | 0.32889 | 241.74562 | **2.46E-08** | 25.6809 | Yes |
| A727U | His211Leu | 0.934 | 0.87215 | 0.38397 | 248 | 0.53614 | 5.53627 | **6.42E-06** | 28 | Yes | 0.38397 | 247.90269 | **1.07E-05** | 24.02735 | Yes |
| C738A | Pro215Thr | 0.89583 | 1.44397 | 0.15001 | 248 | 0.77245 | 2.51358 | **0.01798** | 28 | Yes | 0.15002 | 246.47001 | **0.01827** | 26.80654 | Yes |
| A741U | Met216Leu | 0.98431 | 0.19514 | 0.84544 | 248 | 1.20162 | −2.56638 | **0.01591** | 28 | Yes | 0.84545 | 243.92743 | **0.02046** | 25.22943 | Yes |
| U742G | Met216Arg | 0.87092 | 1.88717 | 0.06031 | 248 | 0.7883 | 2.4819 | **0.01934** | 28 | Yes | 0.06035 | 238.26391 | **3.52E-06** | 24.01659 | Yes |
| U755G† | Phe220Leu | 0.84803 | −1.64737 | **0.02321** | 248 | 1.14861 | 2.28418 | 0.11066 | 28 | Yes | **0.02327** | 231.34936 | 0.11081 | 27.63342 | Yes |
| **Clinically important mutations** | | | | | | | | | | | | | | | |
| C163A | Pro23His | 1.02715 | −0.35738 | 0.72111 | 248 | 1.07158 | −0.57014 | 0.57313 | 28 | No | 0.72111 | 247.82468 | 0.57338 | 26.51889 | No |
| C1134A | Pro347Thr | 0.97119 | 0.3878 | 0.6985 | 248 | 0.87964 | 1.32941 | 0.19444 | 28 | No | 0.6985 | 247.92512 | 0.1962 | 24.03038 | No |
| C1134G | Pro347Ala | 0.98177 | 0.24655 | 0.80546 | 248 | 0.93464 | 0.66532 | 0.51129 | 28 | No | 0.80546 | 247.81182 | 0.51169 | 26.04784 | No |
| C1134U | Pro347Ser | 0.98742 | 0.16636 | 0.86801 | 248 | 0.84335 | 1.80676 | 0.08156 | 28 | No | 0.86801 | 247.93754 | 0.08418 | 22.5303 | No |
| C1135A | Pro347Gln | 1.01357 | −0.18092 | 0.85658 | 248 | 1.01593 | −0.1202 | 0.90518 | 28 | No | 0.85658 | 247.99645 | 0.90519 | 27.6065 | No |
| C1135G | Pro347Arg | 0.99525 | 0.0634 | 0.9495 | 248 | 0.95979 | 0.41254 | 0.68308 | 28 | No | 0.9495 | 247.99456 | 0.68309 | 27.92877 | No |
| C1135U | Pro347Leu | 1.02113 | −0.28125 | 0.77875 | 248 | 1.04932 | −0.44675 | 0.65849 | 28 | No | 0.77875 | 247.99223 | 0.65853 | 27.66197 | No |
| **Other notable mutations** | | | | | | | | | | | | | | | |
| A664G | Asp190Gly | 1.12406 | −1.4189 | 0.15718 | 248 | 1.1686 | −1.42455 | 0.16534 | 28 | No | 0.15727 | 231.65985 | 0.16622 | 25.91649 | No |
| G1116A | Glu341Lys | 1.13277 | −1.64317 | 0.10162 | 248 | 1.1452 | −1.05567 | 0.30014 | 28 | No | 0.10164 | 243.19716 | 0.30147 | 24.37739 | No |

Normalized accessibility of select mutants calculated by integration (area under the curve) of the mppRNA map over the ribozyme site integration (nt positions 718–732) and over the local regional integration (nt positions 640 to 764). Single nucleotide missense mutations are capable of both substantially increasing or decreasing mppRNA predicted accessibility for a ribozyme targeting bRHO mRNA even when outside of the 15-nt target annealing region. Changes can have both local impact, represented by the small integral, and more regional impact, represented by the large integral. Local changes are not always translated into regional changes, nor vice versa. Changes in accessibility are dependent on the specific nucleotide change. Bolded entries indicate statistical significance (P < 0.05).

* Smallest accessibility seen for the site integral.
† Smallest accessibility seen for the local change. Bolded entries indicate statistical significance (P < 0.05).

**FIGURE 7.** Select mutants showing different results. All 199 surveyed mutations had mppRNA maps generated, and area under the curve integrations were performed over the local regional (*blue bar*) and ribozyme site (*red bar*) regions. In these select mutants, qualitative and quantitative shifts in the regional/local structure of the mppRNA map are observed in the vicinity of the 725 hhRz targeting site. This is an indication for changes in stability over the target region. Different mutants were compared with the same region of the WT mRNA mppRNA map. (**A**) WT *hRHO*. (**B**) A664G. (**C**) G720A. (**D**) G723U. (**E**) U755G. (**F**) G1116A.

Similarly, the average local integrals for the two classes of mutants were compared with WT, but the means were not significantly different (Fig. 10E). It is clear that the individual effects of certain mutants on the annealing site accessibility are represented as an averaged class effect.

To evaluate the potential impact of a remote mutation that had a significant impact of the site accessibility we further investigated the U355A mutation. This mutation occurs in a region of substantial stable secondary structure and within the 356 GUC↓ cleavage site that does not suppress with hhRzs.[5] We first explored the MFE structure generated by MFold for the U355A mutant versus the WT mRNA (Fig. 11A). An impact of

U355A is seen on the local structure relative to the WT mRNA but does not extend into the 725 region of the mutant mRNA, which shows identical local structure to the WT mRNA. Evaluating the ensemble of possible structures the U355A mutation has an effect on the region around the 725 cleavage site and the mppRNA site integral shows significantly different accessibility relative to WT mRNA but the local integration did not (Fig. 11B). Reasoning that the effect of the mutation on remote accessibility could be reflected in the ensemble of all potential structures, we used SFold to investigate for propagated differences in structure due to the mutation (Fig. 11C). Using the multidimensional scaling (MDS) analysis of SFold, one can

**Figure 8.** mppRNA maps of select clinically significant *hRHO* mutants. (**A**) Local regional map of the WT *hRHO* mRNA. (**B**) Single nucleotide missense mutation C163A, representing the common adRP mutant P23H. An insignificant increase in site accessibility of 7.2% is noted ($t[28] = 0.57014$; $P = 0.57313$), and the local map architecture is maintained compared with WT. (**C**) A map illustrating the mutation C1134U, responsible for the P347S mutation. A decrease of 15.7% is noted for the site integration ($t[28] = -1.80676$; $P = 0.08156$). Local regional (*blue bar*) and ribozyme site (*red bar*) regions are indicated by the bars.

see that the WT *hRHO* mRNA demonstrates two clustered distributions of structures and the single MFE structure resides in only one of those clusters. The U355A mutation promotes a change in the probabilistic distribution of structures, which is consistent with a larger scale effect of the mutation.

## Other Forms of Mutations

Of the 23 samples that included small insertions, deletions, or indels, only one showed a statistically significant effect on site or local accessibility. This mutation (1032in150del8), a large



**Figure 9.** Comparison of 15- vs. 125-nt integrations. A linear correlation ($R = 0.51531$, $P < 0.0001$) is noticed between changes in the local integration versus the ribozyme site integration when each is normalized to the integral for WT *hRHO* for that region. This suggests that, although related (the ribozyme site sequence is nested within the local regional sequence), there is some independence between changes in the accessibility on a ribozyme binding site versus a regional local scale.

insertion of 150 bp, reduced the site accessibility by 51.3% ($t[28] = -6.25557$; $P < 0.05$).[35] 1032ins150del8 is also the only insertion, deletion, or indel to reduce the local integration significantly, by 26.7% ($t[248] = -3.99756$; $P < 0.05$). Other than this very large insertion, eight others had a change site in accessibility of 10% or more from WT. Of these, seven are insertions or deletions of 3 or more nts, whereas only one was a single nucleotide insertion. Table 2 lists the relative accessibilities of each insertion, indel, or deletion. These results contradict the thought that mutations involving more than 1 nt would have larger influence on accessibility surveyed in the area proximal to the 725GUC↓ cleavage site (four total in the area of nt 725 ± 200).

## Discussion

PTGS therapy has great potential in the treatment of a wide variety of diseases, including adRP. Part of understanding the complexity of developing these treatments includes understanding how the molecular genetic basis of the disease might affect therapeutic efficacy. This bioinformatics study shows that using a predictive model of target RNA structure and ribozyme target accessibility provides expedited insight on how mutational variations that cause adRP could affect ribozyme annealing and subsequent target mRNA cleavage. Although not tested in this study, shRNA- or miRNA-type therapies also target local regions, and one would expect similar outcomes. shRNAs typically use a span of 19 nt for antisense binding, 4 bp larger than sampled for the standard 7 nt/7 nt hhRz tested here in the context of our lead agent with therapeutic potential. We found that most of the comprehensive *hRHO* mutations surveyed had no statistical effect on predicted accessibility at or around the Rz725GUC↓ cleavage motif. However, a small subset of mutations was shown to either increase or decrease accessibility around the 725GUC↓ site, and for most mutations the variability is not extreme. Overall, these outcomes are highly favorable for a mutation-independent PTGS strategy.

As expected, the effect on accessibility at a single target annealing site increases with proximity of the mutation to

FIGURE 10. Comparison of local structural mRNA impact by mutant classes. We compare the effect on structure of the class of mutants that exert statistically significant impact on the 725 region versus the larger class of mutants that do not exert statistically significant impact, in comparison to the WT mRNA. (A) Local mppRNA map of the WT *hRHO*. (B) Local average mppRNA map of all *hRHO* mutants that exerted statistically significant effects on local/regional mppRNA maps around the 725 target site. (C) Local average mppRNA map of all *hRHO* mutants that did not have statistical effect on 725 region accessibility. (D) Quantitative statistical comparison of accessibility in the immediate hhRz binding site region (nt 718 to 732) for WT versus class of statistically significant mutants (SS) and the class of nonstatistically significant mutants (Non-SS) (mean ± SEM, *P* values; WT mean = 0.19342 ± 0.01346; SS mean = 0.13653 ± 0.00865; Non-SS mean = 0.19413 ± 0.01404) (SS: *t*[28] = 3.55571; *P* = 0.00136, significant; Non-SS: *t*[28] = −0.03625; *P* = 0.97134). Asterisk indicates statistical significance. (E) Statistical comparison of accessibility in the local region of hhRz binding (nt 640 to 764) for WT versus mutants with and without significant impact (WT mean = 0.1982 ± 0.0105; SS mean = 0.18192 ± 0.00968; Non-SS mean = 0.19745 ± 0.01026) (SS: *t*[248] = 1.1402; *P* = 0.2553; Non-SS: *t*[248] = 0.05082; *P* = 0.95951).

the region of annealing. However, it was also shown that single nucleotide mutations in relatively distant positions can have an appreciable effect, as shown with the P347 mutation, the two mutations at nt 498 that reduced accessibility, and the global impact of mutation U355A. The effect of single nucleotide mutations on the efficacy of ribozymes has only been explored in the context of a mutation being located within the antisense binding region, where it directly affects initial target annealing and product leaving rates, or directly affects the NUH site, which has general inhibitory effects on the rate of ribozyme cleavage.[28] Our study surveyed a total of 199 mutants, 176 of which were single nucleotide variations. The results of this study strengthen the potential utility of the mutation-independent PTGS strategy but also raise caution that some mutations may affect accessibility in a target mRNA and alter the expected impact of cleavage-active lead therapeutic hhRzs (or other PTGS agents). Prior to using our Rz725GUC↓ therapeutic agent on mutations that induce significant predicted accessibility differences, it would be prudent to experimentally evaluate accessibility to annealing and cleavage on mutant *hRHO* mRNAs both in vitro and in cellulo. Only five of the known *hRHO* mutations directly affect nucleotides within the 15-nt Rz725GUC↓ annealing region in the *hRHO* transcript, and only one directly modifies the hhRz NUH↓ triplet recognition motif (G723U). Knowledge that only a small subset of known *hRHO* mutations could affect accessibility to ribozyme

annealing, a factor critical to therapeutic efficacy, shows that a mutation-independent hhRz therapy targeting the 725GUC↓ site has utility for gene therapy in a large pool (>90%) of individuals with *hRHO* mutations. For those mutations that have a significant predicted effect on accessibility at the Rz725GUC↓ site/region, an experimental approach to evaluate accessibility and efficacy of the given therapy against model mRNAs in vitro and in cellulo is a reasonable approach.

We also found that the ability for a *hRHO* mutation to affect target site accessibility is dependent on the domain the mutation occupies within the transcript. This is evidenced by areas of insulation where mutations had little to no effect on accessibility at the 725 site. The dominant area of insulation (nts 500 to 600) is an area of strong negative folding energy. This likely affects the impact of local single nucleotide mutations on the nearby 725 region as they are unable to propagate structural destabilization out of the insulating region. There were also areas in the target (nts 400 to 500 and 600 to 800) where accessibility around 725 was altered by mutation, albeit not always in a statistically significant manner. Although the impact of any mutation is most likely to have local structural impact on at least secondary RNA folding, as we showed in this study, there is also precedence for longer-range impacts of mutations. Several studies have shown that single nucleotide polymorphisms are most likely to have local effect on RNA structure but can also have longer range effects.[38–42] Having comprehensively surveyed the

**FIGURE 11.** Impact of a single mutation (U355A, Val87Asp) on global mRNA structure. **(A)** Local regions of MFE structures of the *hRHO* WT mRNA and U355A mRNA. The MFE structures had similar free energies and global structures (WT: −530.89 kcal/mol; U355A: −535.48 kcal/mol; Supplementary Fig. S1). The mutation (*red arrows*) has a local impact on structure in U355A versus WT mRNA in the same region but does not impact the local structure around the 725 hhRz target cleavage site (*blue arrows* in U355A and WT mRNAs). **(B)** mppRNA comparison of the WT *hRHO* mRNA map and the U355A map in the region around the 725 target site. The site integration was significantly different for the U355A mutation relative to the WT *hRHO* mRNA ($t[28] = -2.31025$; $P = 0.02846$), whereas the local/regional integration was not ($t[248] = -1.2809$; $P = 0.20143$). **(C)** Multidimensional scaling plots generated by the SFold RNA statistical module for the WT *hRHO* and mutant U355A mRNAs. WT *hRHO* mRNA shows two clusters of sampled structures, and the MFE structure is located in the cluster with the highest probability (WT: two clusters; probabilities: 0.543* [*red dots*], 0.457 [*green dots*]; *cluster with MFE structure; second trial: probabilities 0.537, 0.463 [data not shown]). U355A *hRHO* mRNA also shows two clusters like the WT *hRHO* mRNA, but the distributions are substantially different (U355A: two clusters; first trial: probabilities: 0.847 [*red dots*], 0.153* [*green dots*]; second trial: probabilities: 0.861 and 0.139 [data not shown]). Note that the low probability cluster (*green dots*) now containing the MFE is substantially diffuse compared with the lower probability cluster for the WT mRNA. Each MDS analysis was repeated twice to demonstrate the repeatability of clusters in the sampling of 1000 structures for both the WT and U355A mRNAs. These outcomes suggest that the U355A single nucleotide mutation has a more global effect on the entire set of possible structures engaged by the *hRHO* mRNA.

known *hRHO* mutations, it is reasonable to conclude that our results are representative of the impact of single nucleotide polymorphisms on mRNAs as found in other studies. We expect that newly identified mutations in *hRHO* will follow the same paradigm identified for the current set of variations as to how they could impact target mRNA accessibility.

This study is limited to work with a single predictive model (mppRNA) to identify accessible regions in *hRHO* that are amenable to attack by hhRzs, both in vitro and in cellulo, and focuses on one known accessible region in *hRHO* at

which successful PTGS agents have already been realized. mppRNA is a proven useful method of screening for target accessibility to design, screen, and develop ribozymes for cleavage of *hRHO* mRNA.[5,10] In its current state, the mppRNA model identified our lead candidate hhRz for *hRHO*: Rz725GUC↓. Although the model is proven useful in its current state, it is possible that predictive output could be improved by future bioinformatics studies. At this time, we know that mppRNA has predicted accessibility in *hRHO* that has resulted in meaningful therapeutic mRNA cleavage and

**B**



WT hRHO mppRNA Map

U355A hRHO mppRNA Map

**C**



WT hRHO

U355A hRHO

**FIGURE 11.** Continued.

knockdown by the lead 725 hhRz.[10] The 725 region is therefore a valid reference site to evaluate the impact of *hRHO* mutations on lead agent target accessibility. There are other regions in *hRHO* to which we and other investigators have made successful PTGS agents. A broader bioinformatics study to evaluate the now established database of accessibility in the set of known *hRHO* mutants at these PTGS target sites is feasible (Froebel and Sullivan, unpublished results, 2017). A similar approach could be applied to autosomal dominant disease genes (e.g., *PRPH2* and *BEST1*) for which there are many mutations and are likely to be candidates for PTGS therapeutics in the future.

This study is a first-of-kind evaluation of the potential impact mutations have on the accessibility of a single mutation-independent PTGS therapeutic for autosomal dominant retinal degeneration. Although given mutations can increase or decrease accessibility at the target site, overwhelmingly this study has shown that most mutations do not significantly affect accessibility at a single PTGS cleavage site. For some mutations, there is a significant change in predicted accessibility, and for these mutations, experiments should be performed with mutant transcripts to compare in vitro cleavage and in cellulo target knockdown with other mutants and the WT mRNA. To test all known and predicted mutations would be time consuming and cost prohibitive, so a bioinformatics approach could be used to test effects of diverse mutations on lead target site accessibility. However, it may be valuable to test a subset of

mutations to obtain experimental data to compare with predicted accessibility.

## CONCLUSIONS

A total of 199 mutations/variations of the *hRHO* gene known or suspected to cause adRP were surveyed for changes in predicted accessibility of the target mRNA, 176 of which were single nucleotide missense mutations. The effect of mutation on mRNA accessibility at the hhRz cleavage site is related to its position within the mRNA sequence and the proximity to the target cleavage site for the lead hhRz. Although many single nucleotide mutations exert little effect on predicted accessibility, some mutations can significantly increase or decrease accessibility when both proximate to and at large distances from the cleavage site. Mutations encompassing more than one nucleotide failed to show a greater chance of having a significant effect on accessibility, perhaps due to an insufficient number of sampled mutations of this sort in proximity to the 725GUC↓ cleavage site. An *hRHO* adRP mutation is more likely to have an impact on utility of the mutation-independent strategy of hhRz gene therapy if it occurs proximate to the target cleavage site or exerts gross mRNA folding perturbation. The outcomes in this study are likely to follow for other hhRz cleavage sites in *hRHO* and other targets. Although we found that some mutations can alter the predicted accessibility at a ribozyme cleaving region, overwhelmingly the mutation-

Investigative Ophthalmology & Visual Science

TABLE 2. Relative Accessibilities of Each Insertion, Indel, or Deletion

| Mutation | AA Change | Relative Local Accessibility | Local T | Local Probability | df Students Local | Relative Site Accessibility | Site T | Site Probability | df Students Site | Significance | Welch's Local Probability | df Welch's Local | Welch's Site Probability | df Welch's Site | Significance |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Insertions** | | | | | | | | | | | | | | | |
| 1037insG | Frameshift | 0.99768 | −0.03095 | 0.97533 | 248 | 0.99528 | −0.04514 | 0.96432 | 28 | No | 0.97533 | 247.99892 | 0.96432 | 27.99255 | No |
| 1094ins4 | Frameshift | 0.96682 | −0.44892 | 0.65388 | 248 | 0.96682 | −0.55903 | 0.65388 | 28 | No | 0.65388 | 247.80844 | 0.58059 | 27.98786 | No |
| 1116insG | Frameshift | 0.97151 | −0.38338 | 0.70176 | 248 | 0.84878 | −1.6995 | 0.10031 | 28 | No | 0.70176 | 247.92747 | 0.10173 | 24.80603 | No |
| **Deletions** | | | | | | | | | | | | | | | |
| 131delC | Frameshift | 1.0212 | 0.28052 | 0.77931 | 248 | 1.05883 | 0.52734 | 0.60211 | 28 | No | 0.77931 | 247.92959 | 0.60218 | 27.54981 | No |
| 297_308del12 | 68_71delLeuArgThrPro | 1.00606 | 0.07954 | 0.93667 | 248 | 1.14761 | 0.99938 | 0.32617 | 28 | No | 0.93667 | 247.74222 | 0.32794 | 23.14898 | No |
| 312_314del3 | 73delIAsn | 0.95148 | −0.75663 | 0.44999 | 248 | 0.89089 | −1.19553 | 0.24191 | 28 | No | 0.45005 | 248.55574 | 0.24359 | 23.93338 | No |
| 504_521del18 | 137_142del ValValVal CysLysPro | 0.98593 | −0.17849 | 0.85848 | 248 | 0.85147 | −1.41336 | 0.16857 | 28 | No | 0.85849 | 246.59342 | 0.16949 | 25.83083 | No |
| 711_719del9 | 206_208del TyrMetPhe | | −1.04853 | 0.29545 | 248 | 0.87861 | −1.03624 | 0.30896 | 28 | No | 0.29417 | 238.99641 | 0.31111 | 22.48861 | No |
| 863_865del3 | IleIle255_256Ile | 0.95254 | −0.64549 | 0.51921 | 248 | 1.01163 | 0.08908 | 0.92965 | 28 | No | 0.51921 | 247.62758 | 0.92966 | 27.64708 | No |
| 887_889del3 | CysTrp264_265Trp | 0.94922 | −0.69569 | 0.48728 | 248 | 1.07337 | 0.64859 | 0.52189 | 28 | No | 0.48728 | 247.28603 | 0.52206 | 27.09957 | No |
| 1032_1041del10 | Frameshift | 1.01289 | 0.17033 | 0.86489 | 248 | 0.94723 | −0.54275 | 0.5916 | 28 | No | 0.86489 | 247.88546 | 0.95611 | 27.9281 | No |
| 1032_1170del139 | Nonsense | 0.95864 | −0.55978 | 0.57613 | 248 | 0.8723 | −1.39594 | 0.1737 | 28 | No | 0.57613 | 247.79504 | 0.17737 | 20.91875 | No |
| 1048_1050del3 | LeuThr318_319Pro | 0.98028 | −0.26535 | 0.79096 | 248 | 0.94725 | −0.54112 | 0.5927 | 28 | No | 0.79096 | 247.92656 | 0.5927 | 27.99578 | No |
| 1074_1077del4 | Frameshift | 0.96665 | −0.45015 | 0.653 | 248 | 0.89308 | −1.16911 | 0.25222 | 28 | No | 0.653 | 247.86673 | 0.2542 | 23.25321 | No |
| 1075delC | Frameshift | 1.0461 | 0.60056 | 0.54868 | 248 | 1.0518 | 0.48323 | 0.63269 | 28 | No | 0.54869 | 247.43452 | 0.63272 | 27.80486 | No |
| 1090_1106del17 | Frameshift | 0.93749 | −0.86559 | 0.38755 | 248 | 0.96886 | −0.30355 | 0.76371 | 28 | No | 0.38756 | 246.5567 | 0.76371 | 27.96884 | No |
| 1098delG | Frameshift | 0.98988 | −0.13559 | 0.89225 | 248 | 0.98988 | −0.01177 | 0.89225 | 28 | No | 0.89225 | 247.9864 | 0.9907 | 27.18881 | No |
| 1112del42 | Nonsense | 0.98239 | −0.2362 | 0.81348 | 248 | 0.86318 | −1.50961 | 0.14235 | 28 | No | 0.81348 | 247.97336 | 0.14343 | 25.51882 | No |
| 1114delC | Frameshift | 0.97193 | −0.37972 | 0.70448 | 248 | 0.91567 | −0.88793 | 0.38214 | 28 | No | 0.70448 | 247.82002 | 0.38243 | 26.97505 | No |
| 1116_1123del8 | Frameshift | 1.0119 | 0.1601 | 0.87293 | 248 | 1.01089 | 0.11011 | 0.91311 | 28 | No | 0.87293 | 247.95375 | 0.91311 | 27.99332 | No |
| **Indels** | | | | | | | | | | | | | | | |
| 499_500delGGinsUU | Arg136Leu | 0.99816 | −0.02415 | 0.98075 | 248 | 0.99543 | −0.04 | 0.96838 | 28 | No | 0.98075 | 247.998 | 0.96838 | 27.98761 | No |
| 606_607delCCinsGA | Pro171Glu | 0.9939 | −0.08164 | 0.935 | 248 | 1.0026 | 0.02994 | 0.97633 | 28 | No | 0.935 | 247.9819 | 0.97633 | 27.99992 | No |
| 1032ins150del8 | Insertion | 0.73277 | −3.99756 | **8.44617E-05** | 248 | 0.48744 | −6.25557 | **9.22635E-07** | 28 | Yes | **8.59E-05** | 232.71225 | **3.52E-06** | 20.73333 | Yes |

Bold type represents the accessibilities for 1032ins150del8, the single mutation in this group that significantly affected both local and site accessibility.

independent strategy appears viable to treat diverse mutations in a given disease gene.

## References

1. Dryja TP, McGee TL, Reichel E, et al. A point mutation of the rhodopsin gene in one form of retinitis pigmentosa. *Nature*. 1990;343:364–366.

2. Gal A, Apfelstedt-Sylla E, Janecke AR, Zrenner E. Rhodopsin mutations in inherited retinal dystrophies and dysfunctions. *Prog Retinal Eye Res*. 1997;16:51–79.

3. Wilson JH, Wensel TG. The nature of dominant mutations in rhodopsin and implications for gene therapy. *Mol Neurobiol*. 2003;28:149–158.

4. 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;11:177–185.

5. Abdelmaksoud HE, Yau EH, Zuker M, Sullivan JM. Development of lead hammerhead ribozyme candidates against human rod opsin for retinal degeneration therapy. *Exp Eye Res*. 2009;88:859–879.

6. Sullivan JM, Yau EH, Kolniak, TA, et al. Variables and strategies in development of therapeutic post-transcriptional gene silencing agents. *J Ophthalmol*. 2011;2011:531380.

7. Sullivan JM, Yau EH, Taggart RT, Butler MC, Kolniak TA. Bottlenecks in development of therapeutic post-transcriptional gene silencing agents. *Vision Res*. 2008;48:453–469.

8. Sullivan JM, Yau EH, Taggart RT, Butler MC, Kolniak TA. Relieving bottlenecks in RNA drug discovery for retinal diseases. *Adv Exp Med Biol*. 2012;723:145–153.

9. Scherr M, Rossi JJ, Sczakiel G, Patzel V. RNA accessibility prediction: a theoretical approach is consistent with experimental studies in cell extracts. *Nucleic Acids Res*. 2000;28:2455–2461.

10. Yau EH, Butler, MC, Sullivan, JM. A cellular high-throughput screening approach for therapeutic trans-cleaving ribozymes and RNAi against arbitrary mRNA disease targets. *Exp Eye Res*. 2016;151:236–255.

11. Grimm D, Streetz KL, Jopling CL, et al. Fatality in mice due to oversaturation of cellular microRNA/short hairpin RNA pathways. *Nature*. 2006;441:537–541.

12. Snøve O Jr, Rossi JJ. Toxicity in mice expressing short hairpin RNAs gives new insight into RNAi. *Genome Biol*. 2006;7:231.

13. Rao DD, Senzer N, Cleary MA, Nemunaitis J. Comparative assessment of siRNA and shRNA off target effects: what is slowing clinical development. *Cancer Gene Ther*. 2009;16:807–809.

14. Kleinman ME, Kaneko H, Cho WG, et al. Short-interfering RNAs induce retinal degeneration via TLR3 and IRF3. *Mol Ther*. 2012;20:101–108.

15. Millington-Ward S, O'Neill B, Tuohy G, et al. Strategems *in vitro* for gene therapies directed to dominant mutations. *Hum Mol Genet*. 1997;6:1415–1426.

16. Sullivan JM, Pietras KM, Shin BJ, Misasi JN. Hammerhead ribozymes designed to cleave all human rod opsin mRNAs which cause autosomal dominant retinitis pigmentosa. *Mol Vis*. 2002;8:102–113.

17. Gorbatyuk M, Justlien V, Liu J, Hauswirth WW, Lewin AS. Preservation of photoreceptor morphology and function in P23H rats using an allele independent ribozyme. *Exp Eye Res*. 2007;84:44–52.

18. O'Neill B, Millington-Ward S, O'Reilly M, et al. Ribozyme-based therapeutic approaches for autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 2000;41:2863–2869

19. Dresner KA, Timmers AM, Hauswirth WW, Lewin AS. Ribozyme-targeted destruction of RNA associated with autosomal dominant retinitis pigmentosa. *Invest Ophthalmol Vis Sci*. 1998;39:681–689.

20. Stenson PD, Ball EV, Mort M, et al. The Human Gene Mutation Database (HGMD®): 2003 update. *Hum Mutat*. 2003;21:577–581.

21. Database of Single Nucleotide Polymorphisms (dbSNP). Bethesda (MD): National Center for Biotechnology Information, National Library of Medicine. (dbSNP Build ID: 148). (Gene ID: 6010). Available at: http://www.ncbi.nlm.nih.gov/SNP/.

22. Sherry ST, Ward MH, Kholodov M, et al. dbSNP: the NCBI database of genetic variation. *Nucleic Acids Res*. 2001;29:308–311.

23. Zuker M. Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res*. 2003;31:3406–3415.

24. Ding Y, Chan CY, Lawrence CE. Sfold web server for statistical folding and rational design of nucleic acids. *Nucleic Acids Res*. 2004;32(suppl):W135–W141.

25. Mathews DH, Burkard ME, Freier SM et al. Predicting oligonucleotide affinity to nucleic acid targets. *RNA*. 1999;5:1458–1469.

26. Nathans J, Hogness DS. Isolation and nucleotide sequence of the gene encoding human rhodopsin. *Proc Natl Acad Sci U S A*. 1984;81:4851–4855.

27. Zoumadakis M, Neubert WJ, Tabler M. The influence of imperfectly paired helices I and III on the catalytic activity of hammerhead ribozymes. *Nucleic Acids Res*. 1994;22:5271–5278.

28. Werner M, Uhlenbeck OC. The effect of base mismatches in the substrate recognition helices of hammerhead ribozymes on binding and catalysis. *Nucleic Acids Res*. 1995;23:2092–2096.

29. Grassi G, Forlino A, Marini JC. Cleavage of collagen RNA transcripts by hammerhead ribozymes in vitro is mutation-specific and shows competitive binding effects. *Nucleic Acids Res*. 1997;25:3451–3458.

30. Kasai Y, Shizuku H, Takagi Y, Warashina M, Taira K. Measurements of weak interactions between truncated substrates and a hammerhead ribozyme by competitive kinetic analyses: implications for the design of new and efficient ribozymes with high sequence specificity. *Nucleic Acids Res*. 2002;30:2383–2389.

31. Oh KT, Longmuir R, Oh DM, et al. Comparison of the clinical expression of retinitis pigmentosa associated with rhodopsin mutations at codon 347 and codon 23. *Am J Ophthalmol*. 2003;136:306–313.

32. Gal A, Artlich A, Ludwig M, et al. Pro-347-Arg mutation of the rhodopsin gene in autosomal dominant retinitis pigmentosa. *Genomics*. 1991;11:468–470.

33. Macke JP, Hennessey JC, Nathans J. Rhodopsin mutation proline[347]-to-alanine in a family with autosomal dominant retinitis pigmentosa indicates an important role for proline at position 347. *Hum Mol Genet*. 1995;4:775–776.

34. Restagno G, Maghtheh M, Mhattacharya S, et al. A large deletion at the 3′ end of the rhodopsin gene in an Italian family with a diffuse form of autosomal dominant retinitis pigmentosa. *Hum Mol Genet*. 1993;2:207–208.

35. Al-Maghtheh M, Kim RY, Hardcastle A, et al. A 150 bp insertion in the rhodopsin gene of an autosomal dominant retinitis pigmentosa family. *Hum Mol Genet*. 1994;3:205–206.

36. Chan CY, Lawrence CE, Ding Y. Structure clustering features of the Sfold Web server. *Bioinformatics*. 2005;21:3926–3928.

37. Ding Y, Chan CY, Lawrence CE. Clustering of RNA secondary structures with applications to messenger RNAs. *J Mol Biol*. 2006;359:554–571.

38. Halvorsen M, Martin JS, Broadway S, Laederach A. Disease-associated mutations that alter the RNA structural ensemble. *PLoS Genet*. 2010;6:e1001074.

39. Ritz J, Martin JS, Laederach A. Evaluating our ability to predict the structural disruption of RNA by SNPs. *BMC Genomics*. 2012;13(suppl 4):S6.

40. Martin JS, Halvorsen M, Davis-Neulander L, et al. Structural effects of linkage disequilibrium on the transcriptome. *RNA*. 2012;18:77–87.

41. Sabarinathan R, Tafer H, Seemann SE, Hofacker IL, Stadler PF, Gorodkin J. The RNAsnp web server: predicting SNP effects on local RNA secondary structure. *Nucleic Acids Res*. 2013; 41:W475–W479.

42. Wan Y, Qu K, Zhang QC, et al. Landscape and variation of RNA secondary structure across the human transcriptome. *Nature*. 2014;505:706–712.

43. Stage-Zimmermann TK, Uhlenbeck OC. Hammerhead ribozyme kinetics. *RNA*. 1998;4:875–889.