ACTA ABP BIOCHIMICA POLONICA

Vol. 70, No 3/2023
643–653
https://doi.org/10.18388/abp.2020_6789

Regular paper

CRISPR-Cas9 guided RNA based model for the treatment of Amyotrophic Lateral Sclerosis: A progressive neurodegenerative disorder

Muhammad Naveed[1]✉, Muhammad Aqib Shabbir[1], Tariq Aziz[2]✉, Hafiz Muhammad Hurraira[1], Sayyeda Fatima Zaidi[1], Ramsha Athar[1], Hassan Anwer Chattha[1], Metab Alharbi[3], Abdulrahman Alsahammari[3] aand Abdullah F. Alasmari[3]

[1]Department of Biotechnology, Faculty of Science and Technology, University of Central Punjab, Lahore Pakistan; [2]Department of Agriculture University of Ioannina Arta 47100 Greece; [3]Department of Pharmacology and Toxicology, College of Pharmacy, King Saud University, P.O. Box 2455, Riyadh 11451, Saudi Arabia

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disorder that leads to the degeneration of motor neurons and the weakening of muscles. Despite extensive research efforts, there is currently no cure for ALS and existing treatments only address its symptoms. To address this unmet medical need, genome editing technologies, such as CRISPR-Cas9, have emerged as a promising solution for the development of new treatments for ALS. Studies have shown that CRISPR-Cas9-guided RNAs have the potential to provide accurate and effective silencing in the genetic disease of ALS. Results have demonstrated a 67% on-target score and a 98% off-target score with GC content within the range of 40–60%. This is further validated by the correlation between the gRNA's structural accuracy and the minimum free energy. The use of CRISPR-Cas9 provides a unique opportunity to target this disease at the molecular level, offering hope for the development of a more effective treatment. In silico and computational therapeutic approaches for ALS suggest that the CRISPR-Cas9 protein holds promise as a future treatment candidate. The CRISPR mechanism and the specificity of gRNA provide a novel therapeutic approach for this genetic disease, offering new hope to those affected by ALS. This study highlights the potential of CRISPR-Cas9 as a promising solution for the development of new treatments for ALS. Further research is required to validate these findings in preclinical and clinical trials and to establish the safety and efficacy of this approach in the treatment of ALS.

Keywords: CRISPR-Cas9, RNA-based model, ALS, neurodegenerative disorder, progressive disease, pre-clinical trials

Received: 09 April, 2023; revised: 21 May, 2023; accepted: 28 May, 2023; available on-line: 06 September, 2023

✉e-mail: naveed.quaidian@gmail.com (MN) iwockd@gmail.com (TA)
Acknowledgments of Financial Support: The authors greatly acknowledge and express their gratitude to the Researchers Supporting Project number (RSP2023R335), King Saud University, Riyadh, Saudi Arabia.
Abbreviations: ALS, Amyotrophic lateral sclerosis; SOD1, Superoxide dismutase-1; EMA, European Medicine Agency

INTRODUCTION

Amyotrophic Lateral Sclerosis (ALS), commonly referred to as Lou Gehrig's disease, is a degenerative neurological condition that predominantly impacts the motor neurons that regulate voluntary muscle movement. The ailment was initially documented by the French neurologist Jean-Martin Charcot in 1869 and has subsequently been acknowledged as among the most catastrophic motor neuron disorders (Longinetti & Feng, 2019). ALS is a condition that impacts individuals of diverse ethnicities and age groups across the globe, with a marginal preponderance in males. The Etiology of ALS remains largely elusive, albeit a small proportion of cases, ranging from 5% to 10%, are deemed familial, stemming from heritable genetic mutations (Hardiman *et al*., 2017). The vast majority of ALS cases are sporadic, lacking a discernible familial background. The prevalence of ALS exhibits heterogeneity among diverse populations, with an annual occurrence rate of 1 to 3 cases per 100 000 persons. ALS has an estimated prevalence of approximately 4-8 cases per 100 000 individuals, although certain regions have reported higher rates (Longinetti & Fang, 2019). Symptoms of ALS may first appear in the muscles that govern speech and swallowing, as well as in the hands, arms, legs, and feet. ALS has two types, depending upon the gene due to which it is caused. *SOD1* and C9ORF72 are the genes that are responsible for this fatal disease. These genes play a crucial role in many processes including the control of motor neurons (Masrori & Vane Damme, 2020; Longinetti & Feng, 2019).

Regrettably, at present, there exists no remedy for ALS but the modern artificial intelligence-based technologies can be proved as enhanced treatment strategies with revolutionized drug design approaches (Naveed *et al*., 2023a; Naveed *et al*., 2023b). The mainstay of treatment approaches is centred on symptom management, provision of supportive care, and enhancement of the patient's quality of life. The implementation of a multidisciplinary approach entails the collaboration of a diverse group of healthcare professionals, such as neurologists, physical therapists, occupational therapists, speech-language pathologists, nutritionists, and palliative care specialists (Morimoto *et al*., 2019). The primary objectives of therapy entail the deceleration of the disease advancement, effective symptom management, and the consideration of the emotional and psychological dimensions of the ailment. ALS is a disease that leads to paralysis and causes premature death of patients. A drug named Riluzole is a glutamate antagonist that increases the survival of patients by a few months. Another drug named Endaravone was recently approved by FDA in 2017 (Saitoh & Takahashi, 2020; Yoshino, 2019). However, the European Medicine Agency (EMA)

has not approved this drug. There is no viable treatment for ALS patients despite decades of constant efforts to understand and to develop drugs for this illness. Clinical therapy, including early gastrostomy tube installation for nutritional intervention or non-invasive positive pressure ventilation for respiratory intervention, can help patients to feel better and live longer (Shoesmith *et al.*, 2020). Nevertheless, the condition cannot be cured by these treatments. The complexity of the condition, late diagnosis, and difficulties in developing a treatment that can effectively reach the central nervous system are all factors that can be connected to the failure of conventional pharmacological and interventional therapies (Varghese *et al.*, 2020; Norris *et al.*, 2020).

This research aims to design a proper gene therapy (Insilco) that aids the treatment of ALS. As we know that currently, there is no therapy for ALS other than the anti-glutamatergic compound riluzole, which is less effective in improving symptoms and only increases the survival period of a patient by a few months. The implementation of Insilco approaches for the assessment of current and future treatment strategies in ALS could shorten trial durations and reduce costs and burden on patients, thereby providing hope that effective therapies can be rapidly translated into the ALS clinic.

MATERIALS AND METHODS

Sequence Retrieval

The reference sequence of the Superoxide dismutase-1 (*SOD1*) gene was retrieved from NCBI (Nation Centre for Biotechnology Information) (https://www.ncbi.nlm.nih.gov) from the specifically allocated accession ID NM_000454. The sequence was retrieved in the FASTA format and saved in a text document for further utilization (Schoch *et al.*, 2020).

Local alignment of the sequence

To identify the mutations, present in the *SOD1* gene sequence, the sequence was analysed for local alignment by NCBI BLASTx (https://blast.ncbi.nlm.nih.gov/Blast.cgi). The BLASTx is a tool of NCBI that performs the local alignment of the query sequences with already available sequences in the database and gives results in the form of amino acids to identify the SNPs that might be responsible for the disease. The disease-causing SNPs were validated by the comparison with the NCBI db-SNP database (https://www.ncbi.nlm.nih.gov/snp).

Cas9 Endonuclease Protein

For genome engineering, the particular enzyme Cas9 Endonuclease protein was utilized. This enzyme should be used because it is simple and has all the basic capabilities to split DNA molecules into single strands. The Cas9 genome editing protein includes blunt ends and cuts sites, RNA-guided endonucleases, PAM site (NGG), crRNA, and tracrRNA. It can cut specific regions of the genomic DNA. Cas9 protein caused a double-strand break in the genomic DNA of *SOD1* (Koonin *et al.*, 2023).

Formation of gRNA

The gRNA (guide RNA) is a small synthetic sequence of RNA that functions as a guide for DNA targeting enzymes. These enzymes are used for various purposes like deletion, insertion, or targeting of RNA (Collins *et al.*, 2021). The design of gRNA can be automated through the use of various tools available online. Among these tools we prefer CHOPCHOP (https://chopchop.cbu.uib.no), an online tool that offers a wide range of inputs and alignments to minimize search times. This tool is the most authentic and reliable for designing gRNA (Labun *et al.*, 2021).

Steps for generating gRNA

The CHOPCHOP tool provides a user-friendly interface for designing gRNA sequences for the CRISPR-Cas9 system. To start, the user simply enters the gene name on the home page and selects the desired genome version or species. Next, the user selects the specific version of the Cas9 protein to be used. The user then specifies the desired outcome of the experiment, whether it be gene knockout, inhibition, or repression. The general settings of CHOPCHOP can then be left unchanged. Upon clicking the "Find Target" button, the tool provides a list of all gRNA sequences present in the selected gene. The user can then select the desired gRNA sequence for their experiment (Labun *et al.*, 2021).

Verification of gRNA

To verify the effectiveness of gRNA, it is important to assess its on-target and off-target binding capacity. Several online tools are available for this purpose, in this study IDT (Integrated DNA Technology) (https://www.idtdna.com/site/order/designtool/index/) was chosen. IDT is a database that facilitates molecular-level analysis of DNA and RNA, where gRNA sequence was submitted for the checking off-target and on-target. The more suitable the gRNA was, the higher its on-target and off-target binding capacity became.

Indicating the off-Target and on-Target Scores

It was found that GC, which includes three hydrogen bonds and is more stable, had a stronger connection between guanine and cytosine. As a result, GC content of up to 50% destabilized off-target hybridization while stabilizing the DNA-RNA duplex. The choice of the number range was made because it fell within the *SOD1* coding zone. This numerical range represented the off-target and on-target ratings for gRNA, with numbers running from 0 to 100 and 100 being considered the best score in each situation. The genomic area of the targeted gene had to be carefully chosen. In the study, the *SOD1* gene was used to obtain accurate off-target scores by IDT (Integrated DNA Technology).

Assembling of gRNA Expression Vector

Assembling of gRNA was done after the selection of the desired gRNA. Since it is being genetically modified for this purpose, a specific expression vector that optimized lentiCRISPR v2 (52963) was selected. It has two expression cassettes, the chimeric guide RNA and hSpCas9. The vectors provided on this list were given by Feng Zhang's laboratory (https://zlab.bio) at MIT. All the enlisted plasmids were different, some expressed nuclease-deficient Cas9 that are used for gene activation and repression while others are optimized for genome editing by using CRISPR. For this validation, a new plasmid lentiCRISPR v2 was selected because it includes Cas9 within the same vector that was optimized for the Human Genome (*SOD1* gene).

Calculation Thermodynamic ensemble prediction

The RNAcofold program was a reliable tool for estimating hybridization energy and analyzing the thermodynamic interaction between the target gene and the expected sgRNA. Additionally, the Vienna RNA program was also used for this purpose. Both programs showed how RNA sequences base-paired. The calculated communication energies and equilibrium duplex structure concentrations were determined using the RNAcofold program.

Secondary Structure Prediction

The Mfold Web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi), which included the UNAfold program, was used in the study and considered one of the most reliable and frequently referenced tools for predicting secondary structures. The UNAfold program was used to predict the secondary structure of the sgRNA strand by minimizing free energy, and it was able to predict sgRNA architectures containing pseudoknots.

RESULTS

Sequence Retrieval and Analysis

In this study, the *SOD1* (Superoxide dismutase 1) gene sequence was retrieved from the NCBI (National Center for Biotechnology Information) database (https://www.ncbi.nlm.nih.gov) using the accession number NM_000454.4. The selection of the sequence was based on a comprehensive literature review. The chosen sequence was the start codon of the *SOD1* receptor gene located on chromosome X at position 31659693-31668931.

Open reading frame (ORF) screening

The screening of the coding region of the *SOD1* gene was performed using the open reading frame finder (https://www.ncbi.nlm.nih.gov/orffinder). The ORF finder is a widely used tool for predicting and analysing open reading frames. In the case of the *SOD1* gene, 56 ORFs were identified and were calculated from the genomic coordinates 31659693-31668931. Figure 1 below represents the ORFs of *SOD1* gene.

Analysis of the designed guide RNA

The gRNA design was based on the objective of identifying or cutting the target gene. The CHOPCHOP tool, a web-based platform, was utilized to design the gRNA sequence. The *SOD1* gene was inserted into the software as the targeted gene, with the exon region represented by a blue box and the introns indicated by red lines. The structural representation of *SOD1* gene is given below in the Fig. 2 and the Table 1 below depicts the target sequences with their genomic locations inside the gene.

Table 1. Target sequences and genomic locations of the target sites in *SOD1* gene

Rank	Target sequence	Genomic location	Strand	GC content (%)	Self-complementarity	MM0	MM1	MM2	MM3	Efficiency
1	CTAGCGAGTTATGGCGACGAAGG	chr21:31659760	+	55	2	0	0	0	0	67.55
2	TAACTCGCTAGGCCACGCCGAGG	chr21:31659748	-	65	4	0	0	0	1	69.6
3	CTTCGTCGCCATAACTCGCTAGG	chr21:31659759	-	55	2	0	0	0	1	55.63

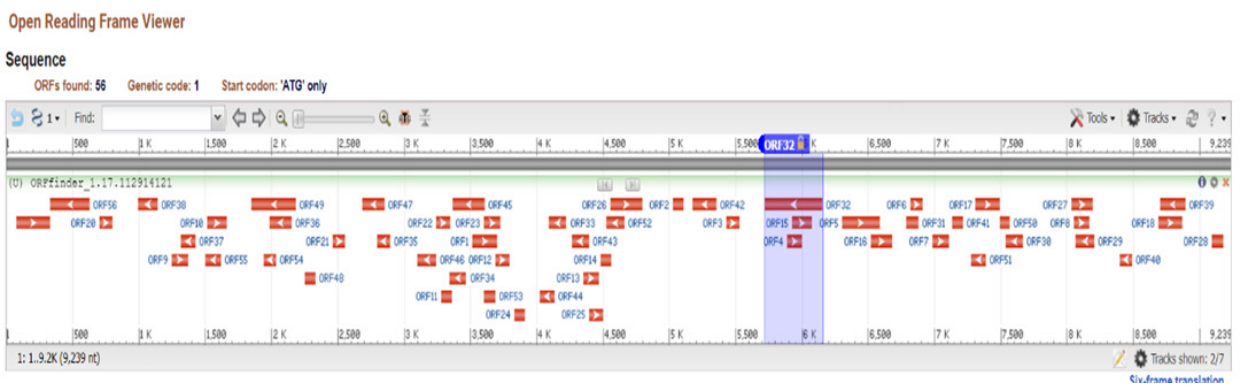

Figure 1. Graphical representation of the open reading frame (ORF) of the *SOD1* gene

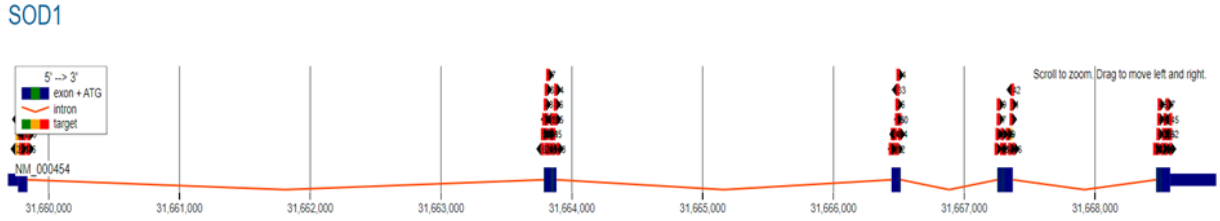

Figure 2. Structural representation of *SOD1* gene

Table 2. The primer specifications and their off-targets for the designed primers of gRNA for rank 1

Pair	Left primers co-ordinates	Left primer	Left primer Tm	Right primer coordinates	Right primer	Right primer Tm	Pair off--target	Product size
1	Chr21:31659647-31659669	CGGAGGTCTG-GCCTATAAAGTA	59.6	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	195
2	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCGGAGA	59.9	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	185
3	Chr21:31659660-31659682	TATAAAG-TAGTCGCGGA-GACGG	60.6	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	182
4	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCGGAGA	59.9	Chr21:31659821-31659843	CCTTCTGCTC-GAAATTGAT-GAT	60.6	0	186
5	Chr21:31659647-31459647	CGGAGGTCTG-GCCTATAAAAG-TA	59.6	Chr21:31659821-31659843	CCTTCTGCTC-GAAATT-GATGTGAT	60.6	0	196

off-Targets		
Location	No. of Mismatches	Sequence
No of Targets		
Shen *et al*., 2018 Predictions of Repair Profile – Statistics of GRNA1		
Reference Sequence	GGGGTTTCCGTTGCAGTCCTCGGAACCAGGACCTCGGCGTGGCCTAGCGAGTTATGGCGA < > CGAAGGCCGTGTGCGTGCTGAAGGGCGACGGCCCAGTGCAGGGCATCATCAATTTCGAGC	
Frameshift Frequency	34.94	
Precision Score	0.59	
Frame+0 Frequency	65.06	
Frame+1 Frequency	65.06	
Frame+2 Frequency	18.52	
1-Bp Ins Frequency	7.46	
Highest Del Frequency	44.27	
Highest Ins. Frequency	4.16	
Highest Outcome Frequency	44.27	
Microhomology Deletion Frequency	82.80	
Microhomology-Less Deletion Frequency	9.74	

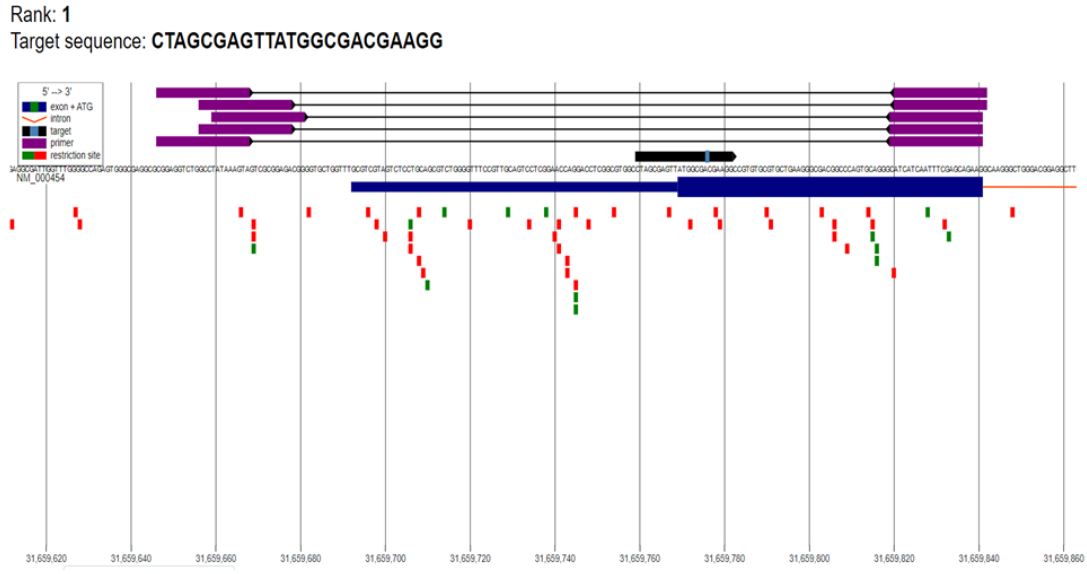

Figure 3. Graphical representation of gRNA along with its primers and off-Targets for rank 1

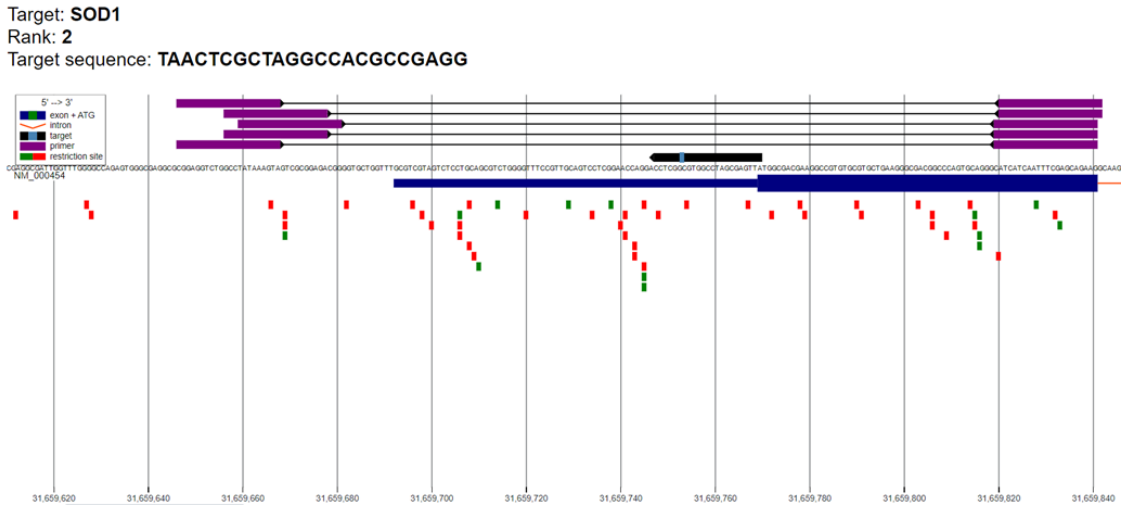

Figure 4. Graphical representation of gRNA along with its primers and off-Targets for rank 2

Table 3. The primer specifications and their off-targets for the designed primers of gRNA for rank 2

Pair	Left primers coordinates	Left primer	Left primer Tm	Right primer coordinates	Right primer	Right primer Tm	Pair off--target	Product size
1	Chr21:31659647-31659669	CGGAGGTCTGGC-CTATAAAGTA	59.6	Chr21:31659820-31659842	CTTCTGCTC-GAAATTGAT-GATG	59.8	0	195
2	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCGGAGA	59.9	Chr21:31659820-31659842	CTTCTGCTC-GAAATTGAT-GATG	59.8	0	185
3	Chr21:31659660-31659682	TATAAAG-TAGTCGCGGA-GACGG	60.6	Chr21:31659820-31659842	CTTCTGCTC-GAAATTGAT-GATG	59.8	0	182
4	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCGGAGA	59.9	Chr21:31659821-31659843	CCTTCTGCTC-GAAATTGAT-GAT	60.6	0	186
5	Chr21:31659647-31459647	CGGAGGTCTGGC-CTATAAAAGTA	59.6	Chr21:31659821-31659843	CCTTCTGCTC-GAAATT-GATGTGAT	60.6	0	196
Highest Outcome Frequency					24.03			
Microhomology Deletion Frequency					77.97			
Microhomology-Less Deletion Frequency					17.49			

off-Targets		
Location	No. of Mismatches	Sequence
chr1:157046129	3	CCTCGGCGcGGCCTAGCGgGTTc
Shen *et al.*, 2018 Predictions of Repair Profile – Statistics of GRNA1		
Reference Sequence	ACTGGGCCGTCGCCCTTCAGCACGCACACGGCCTTCGTCGCCATAACTCGCTAGGCCACG<>CCGAGGT CCTGGTTCCGAGGACTGCAACGGAAACCCCAGACGCTGCAGGAGACTACGACG	
Frameshift Frequency	60.99	
Precision Score	0.49	
Frame+0 Frequency	39.01	
Frame+1 Frequency	44.70	
Frame+2 Frequency	16.29	
1-Bp Ins Frequency	4.54	
Highest Del Frequency	24.03	
Highest Ins. Frequency	2.37	

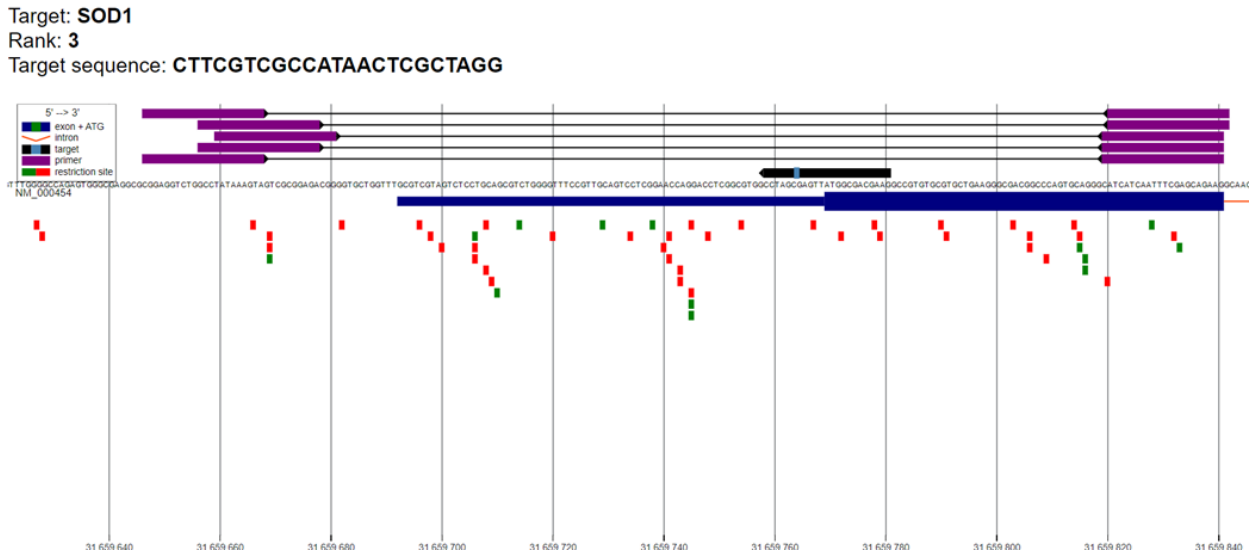

Figure 5. Graphical representation of gRNA along with its primers and off-Targets for rank 3

Table 4. The primer specifications and their off-targets for the designed primers of gRNA for rank 3

Pair	Left primers coordinates	Left primer	Left primer Tm	Right primer coordinates	Right primer	Right primer Tm	Pair off-target	Product size
1	Chr21:31659647-31659669	CGGAGGTCTG-GCCTATAAAG-TA	59.6	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	195
2	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCG-GAGA	59.9	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	185
3	Chr21:31659660-31659682	TATAAAG-TAGTCGCGGA-GACGG	60.6	Chr21:31659820-31659842	CTTCTGCTCGA-AATTGATGATG	59.8	0	182
4	Chr21:31659657-31659679	GCCTATAAAG-TAGTCGCG-GAGA	59.9	Chr21:31659821-31659843	CCTTCTGCTCGA-AATTGATGAT	60.6	0	186
5	Chr21:31659647-31459647	CGGAGGTCTG-GCCTATA-AAAGTA	59.6	Chr21:31659821-31659843	CCTTCTGCTCGA-AATTGATGTGAT	60.6	0	196

off-targets		
Location	No. of Mismatches	Sequence
chr16:49703781	3	CTagGTCGCCcTAACTCGCTGGG
Shen *et al*., 2018 Predictions of Repair Profile – Statistics of GRNA1		
Reference Sequence	TGATGCCCTGCACTGGGCCGTCGCCCTTCAGCACGCACACGGCCTTCGTCGCCATAACTC<>GC TAGGCCACGCCGAGGTCCTGGTTCCGAGGACTGCAACGGAAACCCCAGACGCTGCAGG	
Frameshift Frequency	75.30	
Precision Score	0.46	
Frame+0 Frequency	24.70	
Frame+1 Frequency	35.83	
Frame+2 Frequency	39.47	
1-Bp Ins Frequency	10.15	
Highest Del Frequency	16.98	
Highest Ins. Frequency	7.20	
Highest Outcome Frequency	16.98	
Microhomology Deletion Frequency	68.14	
Microhomology-Less Deletion Frequency	21.70	

Table 5. The target sequences, chromosomal position, and target sources of the *SOD1* target

sgRNAs	Target sequence (5′-3′)	Position of *SOD1* gene	off-target sources	on-target sources	GC% Content
sg(*SOD1*)1	CTAGCGAGTTATGGCGACGA	Ch21:31659760	98	67	55
sg(*SOD1*)2	CTTCGTCGCCATAACTCGCT	Ch21:31659759	96	57.6	55
sg(*SOD1*)3	GTCGCCCTTCAGCACGCACA	Ch21:31659783	82	53	65

Table 6. Validation of the designed primers

Chr21:31659647-31659669 Assembly 1 FWD 5′ CGGAGGTCTGGCCTATAAAGTA3′ Position +/ T_m 60°C Length 22 GC Content 50% Melting Temp. T_m 59.6°C	Chr21: 31659820-31659842 Assembly 1 REV 5′ CTTCTGCTCGAAATTGATGATG 3′ Position –/ T_m 60°C Length 22 GC Content 40.9% Melting Temp. T_m 59.8°C
Chr21:31659657-31659679 Assembly 2 FWD 5′ GCCTATAAAGTAGTCGCGGAGA3′ Position +/ T_m 60°C Length 22 GC Content 50% Melting Temp. T_m 59.9°C	Chr21:31659821-31659843 Assembly 2 REV 5′ CCTTCTGCTCGAAATTGATGAT3′ Position –/ T_m 60°C Length 22 GC Content 40.9% Melting Temp. T_m 60.6°C
Chr21:31659657-3659679 Assembly 2 FWD 5 GCCTATAAAGTAGTCGCGGAGA 3′ Position +/ T_m 60°C Length 22 GC Content 50% Melting Temp. T_m 59.9°C	Chr21:31659902-31659920 Assembly 2 REV 5′CCCGCTCCTAGCAAAGGT3′ Position –/ T_m 60°C Length 18 GC Content 61.1% Melting Temp. T_m 60.4°C

The CHOPCHOP tool was utilized to generate multiple gRNA sequences for the target gene [3]. CHOPCHOP displayed the sequence with the minimum number of mismatches at the top, as it was crucial for the number of mismatches to be zero. If the number of mismatches were greater than one, both the targeted and non-targeted regions of the gene would have been cut. The top three target sequences were carefully selected from the CHOPCHOP results, taking into consideration both efficiency and 100% matching with the target. Figure 3 depicts the graphical representation of gRNA along with its primers and off-targets below. The specifications of the designed primers and their off-targets are given below in the Table 2.

The first rank was selected based on the target sequence CTAGCGAGTTATGGCGACGA and its genomic location on chromosome 21 at position 31659760. The target had a GC content of 55% and 2 instances of self-complementarity. The first pair of primers for the left coordinates on chromosome 21, located between positions 31659647 and 31659669, had a sequence of CGGAGGTCTGGCCTATAAAGTA with a melting temperature (TM) of 59.6°C. The right primer coordinates, located between positions 31659820 and 31659842, had a primer sequence of CTTCTGCTCGAAATTGATGATG and had zero off-targets with a product size of 195. The location of the first gRNA was in the first exon, but in the case of general, mutations near the C terminus of a protein are less likely to result in loss of function for that protein [4]. The graphical representation of the gRNA and primers for rank 2 are given below in the Fig. 4 and the primer specifications and off-targets are given below in the Table 3.

As represented in Fig. 4, the selection was made for rank 3 with a target sequence of CTTCGTCGCCATAACTCGCT and a genomic location located at chr21:31659759, exhibiting a GC content of 59.6% and 2 instances of self-complementarity. The primer pair for the left primers was situated at chr21:31659647-31659669, with a primer sequence of CGGAGGTCTGGCCTATAAAGTA and a melting temperature (TM) of 59.6°C. The right primer was positioned at chr21:31659820-31659842, possessing a primer sequence of CTTCTGCTCGAAATTGATGATG, with zero off-target pairs and a product size of 195. It was generally observed that mutations close to the C-terminus of a protein had a reduced probability of causing a loss-of-function mutation in the protein [5]. The graphical representation of the gRNA and primers for rank 3 are given below in the Fig. 5 and the primer specifications and off-targets are given below in the Table 4.

Above figure represents the selection of rank 6, with a target sequence of GTCGCCCTTCAGCACGCACA located on chr21:31659783 and a GC content of 65%, with zero self-complementarity. The first primer

Table 7. The approximate prediction of the secondary structure free energy model

Target No	Target RNA sequence	Free Energy of Thermodynamic	Frequency of the MFE	Ensemble Diversity
01	CTAGCGAGTTATGGCGACGAAGG	**–1.87** kcal/mol	**75.45** %	2.14
02	CTTCGTCGCCATAACTCGCTAGG	**–0.89** kcal/mol	**45.27** %	2.58
03	GTCGCCCTTCAGCACGCACACGG	**–1.73** kcal/mol	**42.28** %	3.55

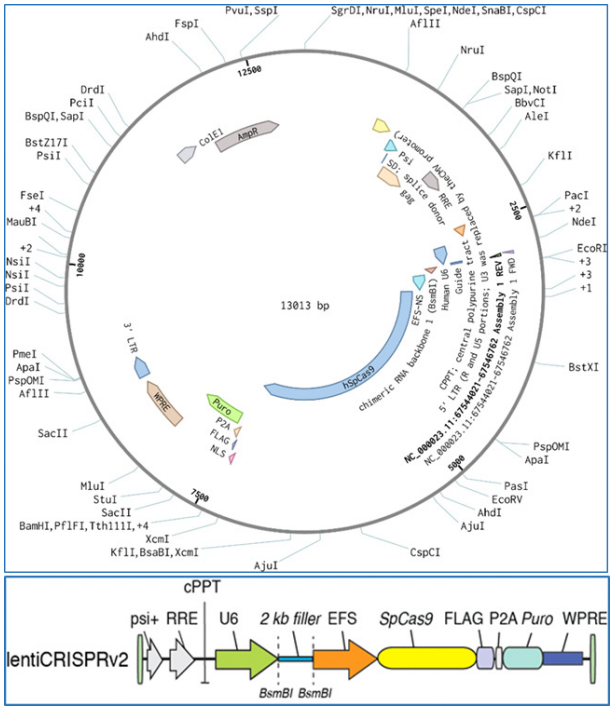

Figure 6. Designed plasmid for the expression of edited target in *SOD1*

pair had a left primer at coordinate's chr21:31659647-31659669, with a primer sequence of CGGAGGTCTGGCCTATAAAGTA and a melting temperature of 59.6°C. The right primer was located at coordinate's chr21:31659828-31659848, with a primer sequence of CCTTGCCTTCTGCTCGAAAT, and its paired off-targets were zero, with a product size of 201. There were three miss-matches in this gRNA, which were caused by the off-target, and it was generally noted that mutations close to the C terminus of a protein are less likely to result in a loss of function mutation of that protein.

Three gRNAs were designed, sgRNA1, sgRNA2, and sgRNA3. SgRNA1 was positioned at chrX:31659760 in the *SOD1* gene, within the target sequence CTAGCGAGTTATGGCGACGAAGG. SgRNA2 was positioned at chrX:31659759 in the *SOD1* gene sequence, within the target sequence CTTCGTCGCCATAACTCGCTAGG, and sgRNA3 was positioned at chrX:31659783 in the *SOD1* gene sequence, within the target sequence GTCGCCCTTCAGCACGCACACGG. The results are explained in Table 5.

Primer validation

For the validation of the primer CHOPCHOP tool was used and the results obtained from it are shown in Table 6.

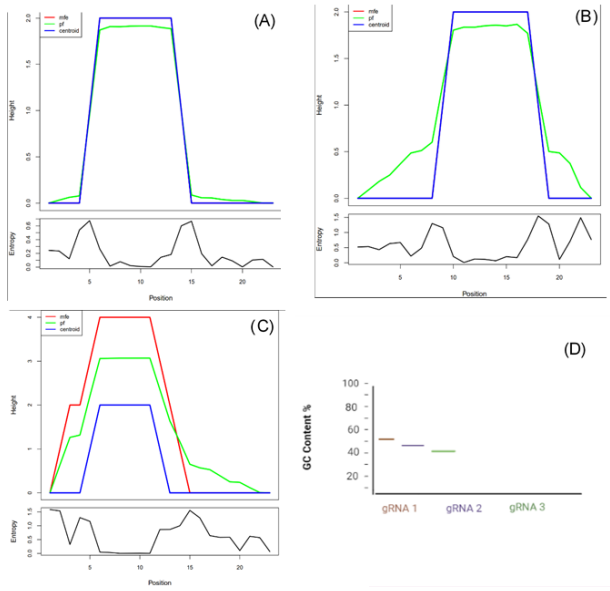

Figure 7. (A) The minimum free energy (MFE) structure, a thermodynamic ensemble of RNA structures, and the centroid structure were represented. The optimal secondary structure, CTAGCGAGTTATGGCGACGAAGG was shown to have a minimum free energy of –1.87 kcal/mol and was depicted in the graph. (B) The MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure were represented. The graph depicted the optimal secondary structure CTTCGTCGCCATAACTCGCTAGG with a frequency of minimum free energy –0.89 kcal/mol. (C) the MFE structure, thermodynamic ensemble of RNA structures, and the centroid structure were represented. The graph displayed the optimal secondary structure, GTCGCCCTTCAGCACGCACACGG, with a minimum free energy of –1.73 kcal/mol. (D) The percentage GC Content.

Design of lentiCRISPR v2 for genome editing

The design of the detailed vector plasmid was obtained through the use of a benching tool, and the different regions were depicted in Fig. 6. The plasmid consisted of two expression cassettes, including the chimeric guide RNA and hSpCas9. The vector was digested using BsmBI, and a pair of annealed oligos were cloned into the single gRNA scaffold. The oligos were designed based on the target sequence site (20bp) and were flanked at the 3' end by a 3bpPAM sequence (AGG) upstream of the protospacer.

The use of LentiCRISPR v2 plasmids was studied for the delivery of CRISPR materials and targeted genome editing in the *SOD1* gene, work on high-efficiency multiplex genome editing of the *SOD1* gene using an engineered mechanism of CRISPR-Cas9. In this study, the LentiCRISPR v2 plasmid was designed by inserting a gRNA expression cassette, from which the Cas9 gRNA was co-expressed with a gRNA that recognized the target sequence or gene. As a result, it was hypothesized that the gRNA expression cas-

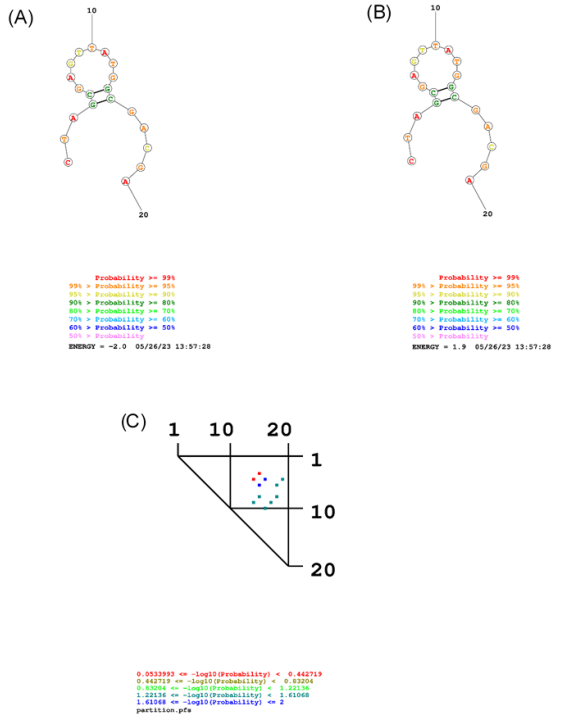

Figure 8. (A) Secondary structure of target gRNA1 fold (B) Secondary structure of target gRNA1 MaxExpect (C) Secondary structure of target gRNA1 Partition.

sette within the LentiCRISPR viral vector would be targeted and destroyed along with the target sequence of the gene, leading to sustained expression of Cas9. The LentiCRISPR v2 carried gRNAs targeting both the Cas9 and *SOD1* genes.

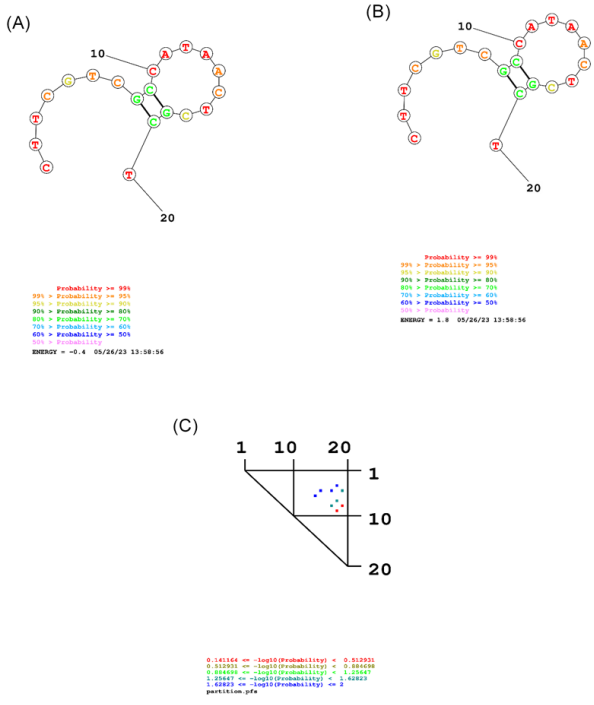

Figure 9. (A) Secondary structure of target gRNA1 fold (B) Secondary structure of target gRNA1 MaxExpect (C) Secondary structure of target gRNA1 Partition.

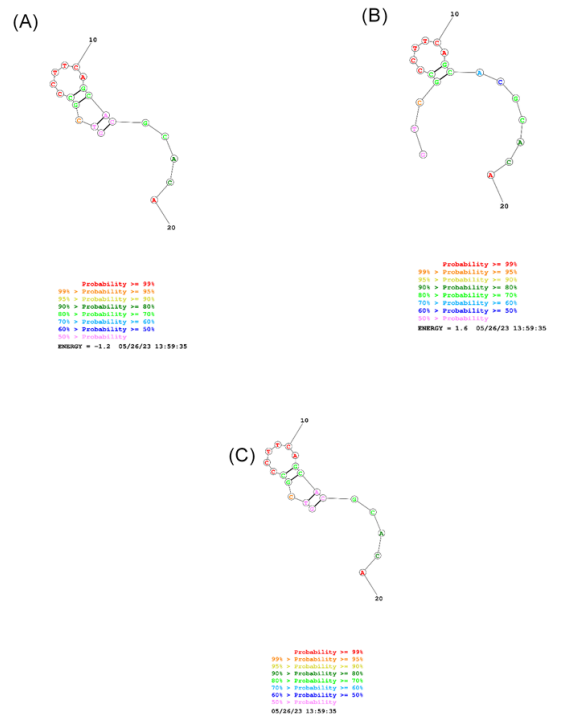

Figure 10. (A) Secondary structure of target gRNA1 fold (B) Secondary structure of target gRNA1 MaxExpect (C) Secondary structure of target gRNA1 Partition.

Thermodynamic ensemble prediction

The approximate prediction of the secondary structure free energy model was represented in Table 7. The frequency of the minimum free energy and ensemble diversity were analyzed against the target gRNA sequences. The RNAfold web server (http://rna.tbi.univie.ac.at/cgi-bin/RNAWebSuite/RNAfold.cgi) was used to predict the thermodynamic ensemble prediction, as shown in Figs 9, 10, and 11 (Graphical Representation of Thermodynamic Ensemble Energy).

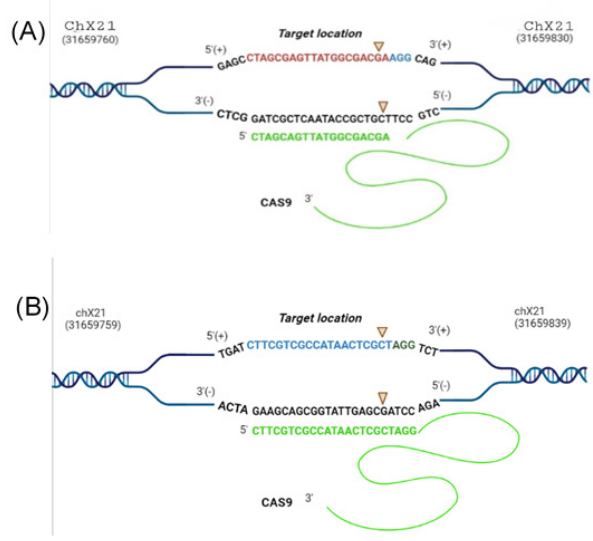

Figure 11. The specificity of the complex was encoded in the first 20 nucleotides of the gRNA

Calculation of the GC Content

In the study, the "RNA /DNA GC Content Calculator" (http://www.endmemo.com/bio/gc.php) was used to predict the GC content. The gene to be knocked down needed to have a standard percentage of GC content as a requirement for use as a sgRNA CRISPR-Cas9. The Mfold server was used to predict the free energy of folding and secondary structure for the sgRNA. By using this server, it was found that the gRNA had 52%, 48%, and 43% GC content, respectively. The percentage GC content is given above in Fig. 8 (D).

Prediction of the secondary structure

The secondary structure of gRNA was predicted using an online tool from Mathew's lab (https://rna.urmc.rochester.edu/RNAstructure.html). The RNA sequence of interest was first input into the software, which then performed calculations including energy minimization and partition function analysis to predict the most likely secondary structure of the RNA molecule. The tool predicted the structure in a graphical format, such as a dot plot or a secondary structure diagram and outputted the structure in various formats for further analysis.

The Mathew Lab RNA Structure tool was widely used in the RNA research community, and it was capable of predicting the secondary structure of RNAs with high accuracy. However, it is important to note that these predictions were based on thermodynamics, which may not always reflect the true structure of an RNA molecule, particularly in the presence of specific proteins or other molecules that could impact its structure.

Designed gRNA evaluation

In this step, two sgRNAs (sgRNA1 and sgRNA2) were designed to target a specific sequence in the *SOD1* gene. The specificity of the complex was encoded in the first 20 nucleotides of the gRNA, which were represented in Fig. 11 (A) and (B) respectively, represented in green. The alteration of these 20 nucleotides resulted in changes to the DNA sequence.

The analysis of the gRNAs was performed. The Cas-9RNP was joined with the (-) sense strand of the gene and produced the double-strand break at the locations 31659693 and 31659843.

It was crucial to ensure that the gRNA was attached to the complementary sequence. The results confirmed the strong incorporation of the guided strand. The ability of the sgRNA to bind with the complementary sequence of the target region demonstrated its capability as a gRNA.

DISCUSSION

Amyotrophic lateral sclerosis (ALS) is a progressive neurodegenerative disease that affects nerve cells in the brain and spinal cord. It is also known as Lou Gehrig's disease, after the famous baseball player who was diagnosed with the disease in the 1930s. The disease causes the degeneration and death of motor neurons, which are the nerve cells that control muscle movement. As the motor neurons die, the muscles they control weaken and eventually stop working, leading to symptoms such as difficulty in speaking, swallowing, and breathing. The progression of the disease varies widely, with some people experiencing a rapid decline in their health, while others may live for many years with the disease (Krishnan *et al.*, 2020).

The current study demonstrates the proof-of-concept utilizing the cellular disorder model whereby genetic defect ALS could be corrected using CRISPR-Cas9 with results indicating that they are more similar to 99%. It demonstrates the on-target and off-target scores that hold GC content within 40-60%, observed by RNA / DNA GC Content Calculator (Meijboom et al., 2022). This prediction is considered significant for the implementation of sgRNAs action. The CRISPR-Cas9 mechanism on the mutant *SOD1* gene is shown in Figure 6. Three target sites of the *SOD1* gene were selected for designing target sgRNAs based on their location in the exonic region. Notably, the designed sgRNA1 and sgRNA2 targeted a sequence in the *SOD1* gene. The specificity of this complex is encoded in the first 20nt of the gRNA as shown in Figure 5 (A) and (B). The binding ability of sgRNA with the complementary sequence of the target region proves their aptitude for working as gRNA (Raikwar et al., 2019).

The target region in the mutant human *SOD1* gene is near the start codon. Scissors indicate the double-strand break location; black color highlights the CAG repeats in the *SOD1* gene, and orange or yellow indicates sgRNA cassettes. Further validations are the minimum free energy that is considered a benchmark of sgRNAs structural accuracy. It measures the stability of the guide strand. The Mfold web server was used to calculate the minimum free energy (Tyagi et al., 2020).. RNA structure webserver was used to predict the secondary structure of an oligonucleotide by folding minimum free energy. This server predicts the most stable structures of an oligonucleotide with max-expected accuracy as shown in the graphical representation in Graphs 1, 2, and 3. These graphs represent the MFE structure, the thermodynamic ensemble of RNA structures, and the centroid structure. The folded structures of oligonucleotides were predicted at a specific temperature of 39°C. These results indicate that CRISPR-Cas9 can provide one of the best therapeutic approaches to ALS.

Imbued by the CRISPR mechanism and suitability of gRNA, pharmaceuticals and researchers are working to utilize this strategy as a therapeutic approach for the next generation. By using CRISPR technology researchers can easily modify the gene function and alter the DNA sequence. The emergence of CRISPR technology opens a new avenue to correct genetic disorders (Duan et al., 2020; Tyagi et al., 2020). Recently, it has been used as an efficient tool for site-specific genome editing in single cells and entire organisms in a specific manner. This study presents a specific possible future candidate in the treatment of SBMA that holds a tremendous potential therapeutic approach for a genetic disorder (Muhammad *et al.*, 2023a; Muhammad *et al.*, 2023b, Muhammad *et al.*, 2023c; Naveed *et al.*, 2022)

Although these methodologies offer significant perspectives, they are unable to comprehensively depict the intricate biological mechanism within a living organism. Additional investigation utilizing animal models and clinical trials involving humans is imperative to authenticate the effectiveness and safety of CRISPR-Cas9 as a treatment for ALS (Kim *et al.*, 2020).

CONCLUSION

Using bioinformatics techniques, we predicted three sgRNAs against the *SOD1* gene on the X chromosome. The fast development of genome editing technologies and applications has been greatly aided by computational

and experimental work. Advanced alignment techniques are being used simultaneously to enhance gRNA design and forecast off-target regions, which might hasten the development of better and more precise editing tools. It is extremely advantageous to create multiple sgRNAs with the aid of CRISPR technology for the silencing of important genes in different biological systems. CRISPR technology may be an emerging key to a novel therapy against *SOD1*. Compared to other gene-editing methods including TALENs and ZFNs, the CRISPR-Cas9 method is being widely used for gene editing and gene correction treatment strategy for several genetic disorders. All this is because of its cost-effectiveness and ability to edit multiple genes at the same time. This research may open a milestone corridor in the therapy of ALS. Research revealed that CRISPR technology is an effective gene silencing method to cure genetic disorders in the future with much higher efficiency.

RECOMMENDATIONS

Since ALS is a disease caused by the production of toxin proteins, in this research we attempted to block the disease-causing expression of the gene by using CRISPR. The designed gRNA will bind to the target site allowing the enzyme to perform cleavage and this will result in the removal of those extra CAG repeats. This will cause the restoration of regular protein production similar to the use of CRISPR in other diseases, which are caused by abnormal numbers of CAG repeats in the gene, e.g., Huntington's disease. The bioinformatics-based approach has shown accurate results, which need to be backed by in vitro and in vivo results for the clinical manifestation of this method. We encourage further research on this particular and seemingly possible way of curing this disease as CRISPR-based gene editing has a tremendous potential for doing miracles in medical science.

Declarations

Funding: No External funding was received.

Conflicts of Interest: The authors declare no conflict of interest.

REFERENCES

Collins SP, Rostain W, Liao C, Beisel CL (2021) Sequence-independent RNA sensing and DNA targeting by a split domain CRISPR–Cas12a gRNA switch. *Nucleic Acids Res* **49**: 2985–2999. https://doi.org/10.1093/nar/gkab100

Duan W, Guo M, Yi L, Liu Y, Li Z, Ma Y, Zhang G, Liu Y, Bu H, Song X, Li C (2020) The deletion of mutant SOD1 *via* CRISPR/Cas9/sgRNA prolongs survival in an amyotrophic lateral sclerosis mouse model. *Gene Ther* **27**: 157–169. https://doi.org/10.1038/s41434-019-0116-1

Hardiman O, Al-Chalabi A, Chio A, Corr EM, Logroscino G, Robberecht W, Shaw PJ, Simmons Z, van den Berg LH (2017) Erratum: Amyotrophic lateral sclerosis (Nature reviews. Disease primers (2017) 3 (17071)). *Nat Rev Dis Primers* **3**: 17085. https://doi.org/10.1038/NRDP.2017.85

Kim BW, Jeong YE, Wong M, Martin LJ (2020) DNA damage accumulates and responses are engaged in human ALS brain and spinal motor neurons and DNA repair is activatable in iPSC-derived motor neurons with SOD1 mutations. *Acta Neuropathol Commun* **8**: 7. https://doi.org/10.1186/s40478-019-0874-4

Koonin EV, Gootenberg JS, Abudayyeh OO (2023) Discovery of diverse CRISPR-Cas systems and expansion of the genome engineering toolbox. *Biochemistry* https://doi.org/10.1021/acs.biochem.3c00159

Krishnan G, Zhang Y, Gu Y, Kankel MW, Gao F-B, Almeida S (2020) CRISPR deletion of the C9ORF72 promoter in ALS/FTD patient motor neurons abolishes production of dipeptide repeat proteins and rescues neurodegeneration. *Acta Neuropathol* **140**: 81–84. https://doi.org/10.1007/s00401-020-02154-6

Labun K, Krause M, Torres Cleuren Y, Valen E (2021) CRISPR Genome editing made easy through the CHOPCHOP website. *Curr Protoc* **1**: https://doi.org/10.1002/cpz1.46

Longinetti E, Fang F (2019). Epidemiology of amyotrophic lateral sclerosis: an update of recent literature. https://doi.org/10.1097/WCO.0000000000000730

Masrori P, Van Damme P (2020) Amyotrophic lateral sclerosis: a clinical review. *Eur J Neurol* **27**: 1918–1929. https://doi.org/10.1111/ENE.14393

Meijboom KE, Abdallah A, Fordham NP, Nagase H, Rodriguez T, Kraus C, Gendron TF, Krishnan G, Esanov R, Andrade NS, Rybin MJ, Ramic M, Stephens ZD, Edraki A, Blackwood MT, Kahriman A, Henninger N, Kocher J-PA, Benatar M, Brodsky MH et al. (2022) CRISPR/Cas9-mediated excision of ALS/FTD-causing hexanucleotide repeat expansion in C9ORF72 rescues major disease mechanisms *in vivo* and *in vitro*. *Nat Commun* **13**: 6286. https://doi.org/10.1038/s41467-022-33332-7

Morimoto S, Takahashi S, Fukushima K, Saya H, Suzuki N, Aoki M, Okano H, Nakahara J (2019) Ropinirole hydrochloride remedy for amyotrophic lateral sclerosis – Protocol for a randomized, double-blind, placebo-controlled, single-center, and open-label continuation phase I/IIa clinical trial (ROPALS trial). *Regen Ther* **11**: 143–166. https://doi.org/10.1016/j.reth.2019.07.002

Muhammad N, Waseem M, Aziz T, Hassan JU, Makhdoom SI, Ali U, Alharbi M, Alsahammari A (2023) Identification of bacterial strains and development of anmrna-based vaccine to combat antibiotic resistance in *Staphylococcus aureus via in vitro* and *in silico* approaches. *Biomedicines* **11**: 1039. https://doi.org/10.3390/biomedicines11041039

Muhammad N, Ain Nu, Aziz T, Javed K, Shabbir MA, Alharbi M, Alsahammari A, Alasmari AF (2023) Artificial intelligence assisted pharmacophore design for Philadelphia chromosome-positive leukemia with gamma-tocotrienol: a toxicity comparison approach with Asciminib. *Biomedicines* **11**: 1041. https://doi.org/10.3390/biomedicines11041041

Muhammad N, Ain Nu, Aziz T, Shabbir MA, Ayesha S, Zafar A, Ghulam N, Alharbi M, Alsahammari A, Alasmari AF (2023). Side chain inset of neurogenerative amino acids to metalloproteins: a therapeutic signature for huntingtin protein in Huntington's disease. *Eur Rev Med Pharmacol Sci* **27**: 6831—6842. https://doi.org/10.26355/eurrev_202307_33154

Naveed M, Noor-ul-Ain, Shabbir MA (2023a) Computational drug shifting towards drug-drug conjugates and monoclonal antibody conjugates in the contradictory excursion of asthma. *Lett Drug Des Discov* **20**: 1219–1229. https://doi.org/10.2174/1570180819666220422114450

Naveed M, Ain N, Aziz T, Javed K, Shabbir M, Alharbi M, Alsahammari A, Alasmari A (2023b) Artificial intelligence assisted pharmacophore design for Philadelphia chromosome-positive leukemia with gamma-tocotrienol: a toxicity comparison approach with Asciminib. *Biomedicines* **11**: 1041. https://doi.org/10.3390/biomedicines11041041

Naveed M, Sheraz M, Amin A, Waseem M, Aziz T, Khan AA, Ghani M, Shahzad M, Alruways MW, Dablool AS, Elazzazy AM, Almalki AA, Alamri AS, Alhomrani M (2022) Designing a novel peptide-based multi-epitope vaccine to evoke a robust immune response against pathogenic multidrug-resistant *Providencia heimbachae*. *Vaccines* **10**: 1300. https://doi.org/10.3390/vaccines10081300

Norris SP, Likanje M-FN, Andrews JA (2020) Amyotrophic lateral sclerosis: update on clinical management. *Curr Opin Neurol* **33**: 641–648. https://doi.org/10.1097/WCO.0000000000000864

Raikwar SP, Kikkeri NS, Sakuru R, Saeed D, Zahoor H, Premkumar K, Mentor S, Thangavel R, Dubova I, Ahmed ME, Selvakumar GP, Kempuraj D, Zaheer S, Iyer SS, Zaheer A (2019) Next generation precision medicine: CRISPR-mediated genome editing for the treatment of neurodegenerative disorders. *J Neuroimmune Pharmacol* **14**: 608–641. https://doi.org/10.1007/s11481-019-09849-y

Saitoh Y, Takahashi Y (2020) Riluzole for the treatment of amyotrophic lateral sclerosis. *Neurodegener Dis Manag* **10**: 343–355. https://doi.org/10.2217/nmt-2020-0033

Schoch CL, Ciufo S, Domrachev M, Hotton CL, Kannan S, Khovanskaya R, Leipe D, Mcveigh R, O'Neill K, Robbertse B, Sharma S, Soussov V, Sullivan JP, Sun L, Turner S, Karsch-Mizrachi I (2020) NCBI Taxonomy: a comprehensive update on curation, resources and tools. *Database* **2020**: https://doi.org/10.1093/database/baaa062

Shoesmith C, Abrahao A, Benstead T, Chum M, Dupre N, Izenberg A, Johnston W, Kalra S, Leddin D, O'Connell C, Schellenberg K, Tandon A, Zinman L (2020) Canadian best practice recommendations for the management of amyotrophic lateral sclerosis. *Can Med Assoc J* **192**: E1453–E1468. https://doi.org/10.1503/cmaj.191721

Tyagi S, Kumar R, Das A, Won SY, Shukla P (2020) CRISPR-Cas9 system: A genome-editing tool with endless possibilities. *J Biotechnol* **319**: 36–53. https://doi.org/10.1016/j.jbiotec.2020.05.008

Varghese GM, John R, Manesh A, Karthik R, Abraham OC (2020) Clinical management of COVID-19. *Indian J Med Res* **151**: 401

Yoshino H (2019) Edaravone for the treatment of amyotrophic lateral sclerosis. *Expert Rev Neurother* **19**: 185–193. https://doi.org/10.1080/14737175.2019.1581610