BLA STN 125785

Casgevy
Exagamglogene autotemcel (exa-cel)

Reviewer
Komudi Singh, Bioinformatics Reviewer OTP/OCTHT

1. **BLA#:** STN 125785

2. **APPPLICANT NAME AND LICENSE NUMBER**
   Vertex Pharmaceuticals Incorporated; License # 2279

3. **PRODUCT NAME/PRODUCT TYPE**
   Non-Proprietary/Proper/USAN: Exagamglogene autotemcel (exa-cel)
   Proprietary Name: Casgevy
   Company codename: CTX001

4. **GENERAL DESCRIPTION OF THE FINAL PRODUCT**
   a. Pharmacological category: Autologous Genome Edited Hematopoietic
                                   Stem Cell-Based Gene Therapy
   b. Dosage form: Suspension for infusion
   c. Strength/Potency: >3 x 10^6 cells/mL
   d. Route of administration: Intravenous infusion
   e. Indication: Treatment of transfusion-dependent
      thalassemia (TDT)

5. **MAJOR MILESTONES**
   Initial IND Submission (BB-IND 18143)   April 27, 2018
   IND allowed to proceed   October 10, 2018
   Orphan Drug Designation granted   May 11, 2020
   Regenerative Medicine Advanced Therapy Designation
   granted   May 05, 2020
   Pre-BLA Meeting   August 9, 2022
   BLA Submission   April 3, 2023
   First Committee Meeting   April 21, 2023
   Filing Meeting   May 11, 2023
   BLA Filed   May 30, 2023
   Mid-Cycle Meeting   September 14, 2023
   External Late-Cycle Meeting   December 18, 2023
   PDUFA action due date   March 30, 2024

6. **BIOINFORMATICS REVIEW TEAM**

Reviewer/Affiliation	Section/Subject Matter
Komudi Singh, Bioinformatics Reviewer CBER/OTP/OCTHT/DCT1	Off-target analysis using bioinformatics, TIDE analysis, (b) (4)

7. **INTER-CENTER CONSULTS REQUESTED**
No inter-center consults were requested.

8. SUBMISSION(S) REVIEWED

Date Received	Submission	Comments
02/24/23	125785.001	Bioinformatics Module 4
04/28/23	125785.011	Response to Bioinformatics IR#1
05/10/23	125785.013	Response to Bioinformatics IR#2
5/19/23	125785.019	Response to Bioinformatics IR#3
05/22/23	125785.021	Response to Bioinformatics IR#4
08/15/23	125785.036	Response to Bioinformatics IR#5
09/08/23	125785.039	Response to Bioinformatics IR#6
09/18/23	125785.042	Response to Bioinformatics IR#7
09/26/23	125785.047	Response to Bioinformatics IR#8
11/03/23	125785.057	Response to Bioinformatics IR#9
01/03/24	125785.087	Bioinformatics PMR2 IR

9. Referenced REGULATORY SUBMISSIONS (e.g., IND BLA, 510K, Master File, etc.)

No referenced regulatory submissions for bioinformatics information.

10. REVIEWER SUMMARY AND RECOMMENDATION

A. EXECUTIVE SUMMARY

This review memo summarizes the Agency's assessment of bioinformatics information submitted in Modules 3 and 4 of the BLA. The Applicant used *in silico* and cellular assays for off-target assessment of exa-cel. *In silico* analysis methods relied on the use of the guide RNA (gRNA) sequence information and user-provided mismatch criterion while scanning the human reference genome to identify potential off-target editing sites based on sequence homology. Since this analysis scans only the reference genome sequence, potential off-target editing sites in the target sickle cell disease (SCD) population could arise due to nucleotide variations present in individual genomes. To account for genome heterogeneity in the assessment of off-target editing in the target population, the Applicant performed variant-aware homology search and identified 50 additional off-target loci. The Applicant used the 1,000 genomes project database that has whole genome sequencing (WGS) data for 2,504 individuals. Of this, sequencing data were from 661 individuals representative of the target population with only 61 WGS data that were collected from individuals in the United States. The small sample size in the database may not adequately capture variants in the patients across United Stated. Additionally, the Applicant used hematopoietic stem and progenitor cells (HSPCs) from a limited number of samples from healthy subjects (n=3), transfusion dependent thalassemia (TDT) subjects (n=3), and SCD subjects (n=3) to perform cellular off-target analysis. The healthy and TDT samples would add to the total samples tested using cell-based off-target analysis method. The Applicant subsequently performed hybrid capture sequencing as confirmatory testing and reported that no off-target editing was observed at the sites tested. However, due to sample limitations, only a fraction of the variants harboring potential off-target loci were empirically tested. To more comprehensively account for genetic heterogeneity in SCD target populations in the United States and to assess the impact of these variants on the potential for off-target

editing, a bioinformatics off-target study that would allow for both including a greater number of variants in their in silico analysis and experimentally testing the editing potential at variant contributed off-target loci was implemented as a post-marking requirement (PMR).

In addition, this memo summarizes the Agency's assessment of the computational approach to measuring on-target editing frequency that the Applicant has proposed for lot release testing. Finally, this memo provides the bioinformatics review of (b) (4)

B. RECOMMENDATION

I. APPROVAL

The Applicant has provided a sufficiently detailed assessment of off-target editing risks in exa-cel using orthogonal methods. However, several variants harboring potential off-target loci were not experimentally tested to measure editing potential at these off-target loci. Hence additional analysis for a more comprehensive assessment of heterogeneity will be needed as a PMR (see below). They have also provided a detailed description of analytical procedures in place to validate computational software and scripts used to measure on-target editing frequency. They have provided an adequate description of

(b) (4)

Their report and their approach for determining on-target editing frequency and (b) (4) testing are acceptable.

Bioinformatics PMR (PMR#2):
Conduct studies to comprehensively assess and screen for the impact of sequence heterogeneity on the risk of off-target editing in the patient population for exa-cel. Specifically:
i. Perform a new in silico off-target analysis using publicly available databases/datasets to allow for inclusion of more variants. Specifically, perform analysis using all variants with at least 0.5% allele frequency in at least one of the 5 continental groups (Africa, Europe, East Asia, South Asia, and the Americas).
ii. Perform confirmatory testing, as appropriate and feasible, of all the off-target loci nominated from the new in silico analysis from (i) as well as those that were not accounted for in the previous study using appropriate samples harboring variants. Specifically,
    a. Screen for the presence of all previously identified variants (e.g., CPS1) as well as any variants identified in study (i) and (ii) in the patients treated in Studies 121, 111, 141, 151, 161, and 171.

b. For patients with a confirmed variant(s), assess for indels and chromosomal changes at each respective locus in appropriate samples.

II.   COMPLETE RESPONSE (CR)

Not applicable

III.   SIGNATURE BLOCK

Reviewer/Title/Affiliation	Concurrence	Signature and Date
Komudi Singh, PhD Bioinformatics Reviewer, OCTHT	Concurred	
Cinque Soto, PhD Bioinformatics Team Lead, OCTHT	Concurred	
Steven Oh, PhD Deputy Director, OCTHT	Concurred	

Table of Contents

List of Tables

List of Figures

Glossary of Terms

BCL11A	B-cell lymphoma/leukemia 11A
bp	base pair
CD34+	cluster of differentiation 34+
CRISPR/Cas9	clustered, regularly interspaced, short palindromic repeats-associated 9 nucleases
dsODN	double-stranded oligodeoxynucleotides
DSB	double-strand DNA break
exa-cel	exagamglogene autotemcel
FDA	Food and Drug Administration
GUIDE-seq	Genome-wide Unbiased Identification of DSBs Enabled by Sequencing
gRNA	guide RNA
HbF	fetal hemoglobin
HD	healthy donor
HSPC	hematopoietic stem and progenitor cell
IR	information request
LOD	limit of detection
NGS	next-generation sequencing
PAM	protospacer adjacent motif
SCD	sickle cell disease
SD	standard deviation
sgRNA	single guide RNA
SNP	single nucleotide polymorphism
TDT	transfusion-dependent β-thalassemia
TIDE	Tracking of Indels by Decomposition
WGS	whole genome sequencing

Module 2.2.4.1

Section I. – Genotoxicity and Off-Target Analysis

The Applicant outlined a list of bioinformatics tools they used to assess the safety of the gRNA SPY101 in their preclinical study shown in **Figure 1**. The tools are listed in the "On and off target analysis" box. As a part of the lead gRNA characterization, the Applicant used high-depth hybrid capture and GUIDE-Seq on the gene edited cells from healthy donors and from patient cells. The Applicant described the steps used to perform GUIDE-Seq and the *in-silico* tools used in assessing the off-targets loci in Module 2 (section 2.2.4.1) of the BLA.

Figure 1. Flow Chart of Studies for Exa-Cel
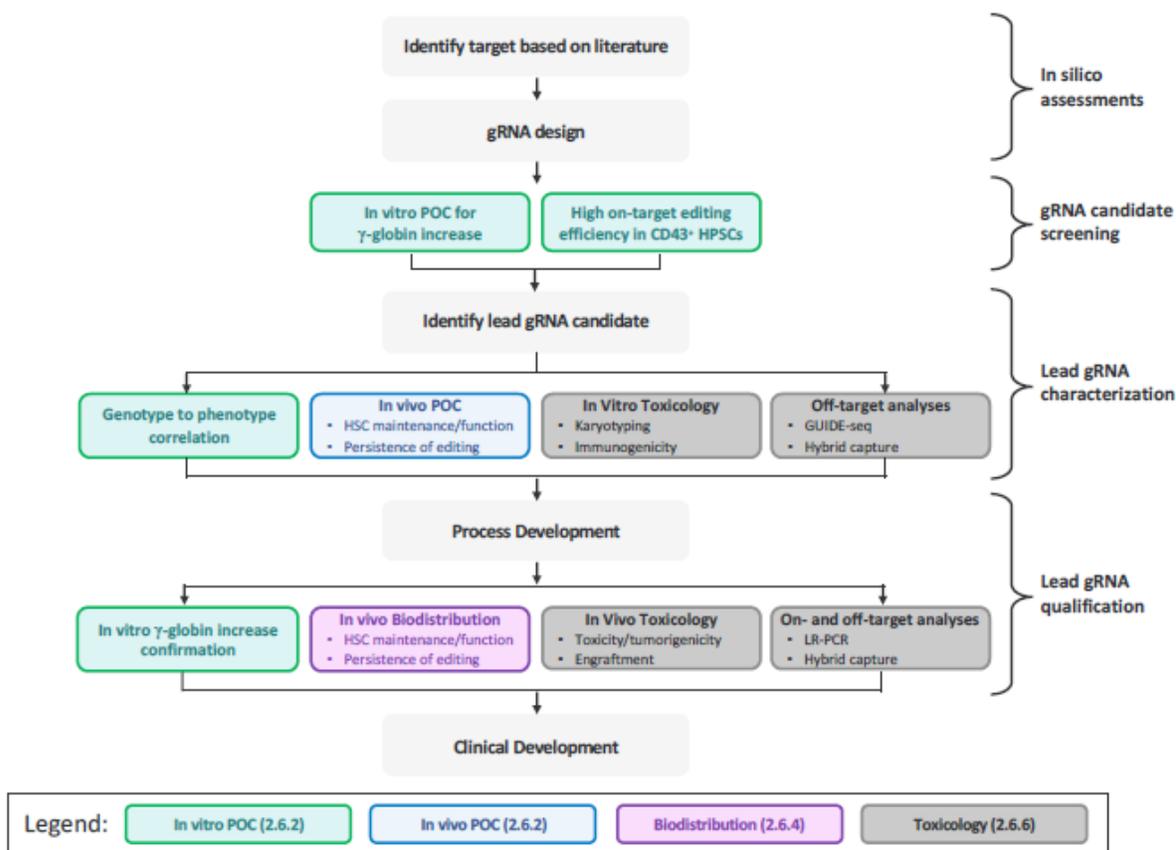
Source: Figure 1: Flow Chart of Studies for Exa-Cel (Module 2, nonclinical-overview.pdf, Amendment 001, 2022)

Reviewer Comment: The Applicant's off-target analysis study strategy is acceptable.

Module 4.3.3.1

Section I. – SPY101 Off-Target Assessment using *in silico* methods

The Applicant used three *in silico* algorithms to nominate potential off-target sites for the sgRNA *SPY101* based on its homology to the reference sequence. The three algorithms used were (b) (4) Each of these algorithms uses a homology-based off-target search that is unique to the tool. (b) (4) uses the (b) (4) sequence alignment and ranks the off-target sites by considering the position of the mismatch.

Figure 2: A cartoon depiction of the target genomic DNA and the components of the CRISPR-Cas9 sgRNA genome editing system
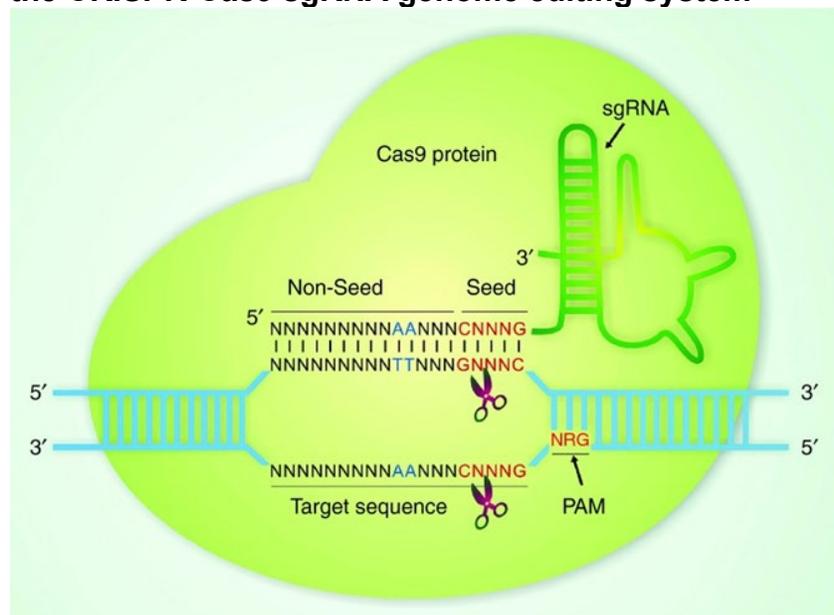
The seed sequence of the sgRNA that is adjacent to the PAM sequence is highlighted in red. Watson and Crick pairing between the sgRNA and target DNA in the presence of PAM results in cleavage of the target DNA resulting in a double stranded break. (Figure from PMID: 26575098)

The (b) (4) method considers mismatches closer to the protospacer adjacent motif (PAM) sequence (shown as the seed sequence in **Figure 2**). Mismatches in this region

(b) (4)

have been shown to decrease the likelihood of an off-target edit. In summary, (b) (4) can be inclusive of up to 5 mismatches but would not identify off-targets with more than (b)(4) mismatches in the seed region. (b) (4) uses the (b) (4) and searches for off-targets with either (b)(4) mismatches and (b)(4) indel or with just (b)(4) mismatches. (b) (4) uses the (b) (4) alignment algorithm and provides users with additional off-target identification options and scores such as cutting frequency determination.

For off-target nomination, the Applicant first used a mismatch criterion of 5bp when implementing (b) (4) and (b) (4) . They used default mismatch criteria when implementing (b) (4) . They also used suboptimal PAMs such as NGA, NAA, NCG, NGC, NTG, and NGT with 4 mismatches to nominate off-targets. From this analysis, a total of 5,007 loci were identified as potential homology-based off-target sites. These findings were reported in ctxsr-015.pdf.

Subsequently, the Applicant performed another *in silico* analysis in which they excluded loci with >3 mismatches. This filtering resulted in the nomination of 171 off-target loci that were subjected to confirmatory testing with hybrid capture sequencing using the genomic DNA samples from *SPY101* edited CD34+ HSPCs from 4 healthy donors (See **Section II: Table 2**).

Section II. –Overview of Confirmatory Testing by Hybrid Capture Sequencing

The Applicant used a Hybrid Capture Library to validate the off-targets that were identified from the GUIDE-Seq and *in silico* homology-based analysis.

Figure 3: Schematic representation of hybrid capture baits (120-mers, green) tiled across a potential off-target site (20-mer, salmon)
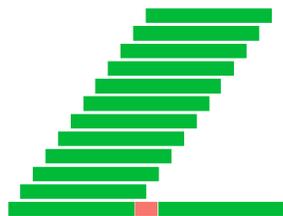
Source: Figure 1 (ctxsr-036.pdf, BLA 125785 Amendment 001)

The Applicant used the strategy shown in **Figure 3** to (b) (4)

For each healthy donor whose HSPCs were assessed for off-target editing, 223 non-overlapping candidate off-target regions were identified from the combined cellular and *in silico* analysis (see **Section V**). The 223 loci identified from the healthy donor samples were included in the confirmatory testing of the patient samples (N=6). Additional loci identified in each patient sample were also included in the confirmatory testing. The total number of off-targets in patient samples ranged from 237-249 (see **Section VII**).

CD34+ HSPCs from the mobilized peripheral blood (mPB) of 4 healthy subjects were used for validation of the off-targets identified from the *in silico* analysis. Briefly, (b) (4)

Sequencing data were aligned with the (b) (4) algorithm using default parameters to the human reference genome (hg38). De-duplication of the aligned reads was done with (b) (4). Sites with a >=0.2% difference in indel frequency between treated and untreated in any one donor were subject to statistical testing across all four donors.

Section III. – Confirmatory Testing of *In silico* Nominated Loci Using Hybrid Capture Sequencing

Confirmatory testing was performed using a hybrid capture library that targeted these 5007 sites. The test used genomic DNA from CRISPR-Cas9/SPY101 and control edited HSPCs from 4 healthy donor (HD) replicates. The Applicant did not provide data on the demographic information for the source of the samples used in the confirmatory testing. Statistical significance for confirmatory testing was set at ≥1% indel frequency between edited and control samples. A summary of the hybrid-capture assessment of off-target sites nominated from the *in silico* method is presented in **Table 1**.

Table 1. Hybrid Capture Characteristics and Results for Sites Identified Through Sequence Homology

Number of Mismatches: PAM Types	Number of Off-Target Sites Identified	Number of Sites With Sufficient Quality	Total Sites Confirmed by Hybrid Capture
5: NGG, NAG 4: NGA, NAA, NCG, NGC, NTG, NGT	5007	4340	0

Source: Table 6: Hybrid Capture Characteristics and Results for Sites Identified through Sequence Homology (ctxsr-015.pdf, BLA 125785 Amendment 001 2017)

The details of the healthy donor-derived HSPCs used in hybrid capture sequencing experiment to test 171 *in silico* nominated loci is presented in **Table 2**. Three of the hybrid capture sequencing samples were from the individuals of Hispanic ancestry and one from African American ancestry. Statistical significance for confirmatory testing was set at ≥0.2% indel frequency difference between edited and control samples.

Table 2. Metadata for Samples Used for Hybrid Capture Experiments

Donor ID	Disease Status	Age (Years)	Sex	Race/Ethnicity
Donor 1	Healthy	29	Female	Hispanic
Donor 2	Healthy	27	Male	Hispanic/Latino
Donor 3	Healthy	31	Male	Hispanic
Donor 4	Healthy	35	Male	African American

Source: Table 1: Metadata for Hybrid Capture Experiments (responses-to-fda-bioinformatics-ir5.pdf, Amendment 0036, 2023)

In this analysis, there were no statistically significant off-target editing events observed at any of the off-targets nominated using *in silico* analysis. The summary of the hybrid-capture assessment of off-target candidate sites nominated from the *in silico* analysis is presented in **Table 3**.

Table 3. Hybrid Capture Characteristics and Results for Sites Identified Through Sequence Homology

Number of Mismatches: PAM Types	Number of Off-Target Sites Identified	Total Sites Confirmed by Hybrid Capture
3: NGG, NAG, NGA, NAA, NCG, NGC, NTG, NGT	171	0

Source: Table 7: Hybrid Capture Characteristics and Results for Sites Identified through Sequence Homology (res-ind-042.pdf, BLA 125785 Amendment 001 2018)

Reviewer Comment: The Applicant's assessment of in silico nominated off-target loci using hybrid capture is sufficient. However, to account for heterogeneity, Applicant will use a lenient allele frequency cutoff that would allow them to assess more variants and repeat the in silico off-target analysis in the PMR (see Bioinformatics PMR in the review recommendation section above)

Section IV. – SPY101 Off-Target Assessment by GUIDE-seq (HSPCs from healthy donors)

In the **section 4.2.3.3.1** of the BLA, the Applicant described the cellular Genome-wide Unbiased Identification of DSBs Enabled by Sequencing (GUIDE-Seq) they implemented to identify candidate off-targets of the SPY101 in CD34+ HSPCs. The Applicant used 6 samples from 1 healthy donor (Donor 1) consisting of CD34+ HSPCs from mobilized peripheral blood (mPB) and tested the effects of various double-stranded oligodeoxynucleotide (dsODN) concentrations, incubation time on number of off targets detected, and cell viability [ctxsr-16.pdf]. Three different dsODN concentrations in the range of (b) (4) μM were examined. After (b) (4) of incubation, genomic DNA was then purified from the cells and used to construct sequencing libraries (**Table 4**). These sequencing libraries were analyzed by NGS and computational tools at CRISPR Therapeutics.

Table 4. Sites Identified in *SPY101* GUIDE-seq

dsODN concentration	Treatment time	Cell viability (%)	Total off-target sites
(b)(4) micromolar	(b) (4)	88	(b) (4)
(b)(4) micromolar	(b) (4)	94	(b) (4)
(b) (4) micromolar	(b) (4)	33	(b) (4)
(b) (4) micromolar	(b) (4)	36	(b) (4)
(b)(4) micromolar	(b) (4)	29	(b) (4)
(b)(4) micromolar	(b) (4)	30	(b) (4)

Source: Table 7: Sites Identified in SPY101 GUIDE-seq (ctxsr-016.pdf, Amendment 001 2017)

In cells incubated with (b) (4) μM and (b)(4) μM dsODN, a higher number of double strand breaks (DSBs) was observed compared to (b)(4) μM dsODN. Higher concentrations of dsODN resulted in increased cellular toxicity measured by (b) (4) . The Applicant observed low cell viability (<50%) at dsODN concentrations (b) (4) μM. Hence, they reported only off-target loci identified with (b)(4) μM dsODN (see **Table 5** below).

Table 5. Sites Identified in SPY101 GUIDE-seq ((b)(4) μM dsODN)

dsODN Concentration	Treatment Time	Cell Viability (%)	Total Off-Target Sites
(b)(4) micromolar	(b) (4)	88	(b) (4)
(b)(4) micromolar	(b) (4)	94	(b) (4)

Source: Table 7: Sites Identified in SPY101 GUIDE-seq (ctxsr-016.pdf, Amendment 001 2017)

Subsequently, the applicant repeated this experiment with lower dsODN concentrations ((b) (4)) using HSPCs from 3 healthy donors that were different from Donor 1 (described above) and assessed cell viability (see **Table 6** below).

Table 6: dsODN Dose Titration for the indicated donor samples

Donor ID	Treatment Time	dsODN Concentration	Cell Viability (%)
(b) (6) Healthy	(b) (4)	(b)(4) micromolar	74.5
Healthy		(b) (4) micromolar	68
Healthy		(b) (4) micromolar	64
Healthy		(b)(4) micromolar	58.5
Healthy		(b)(4) micromolar	29.5
Healthy		(b)(4) micromolar	73
Healthy		(b) (4) micromolar	73
Healthy		(b) (4) micromolar	68.5
Healthy		(b)(4) micromolar	57
Healthy		(b)(4) micromolar	37.5
Healthy		(b)(4) micromolar	79.5
Healthy		(b) (4) micromolar	76
Healthy		(b) (4) micromolar	72
Healthy		(b)(4) micromolar	63
Healthy		(b)(4) micromolar	36

Source: Table 2: dsODN Dose Titration (nonclin-info-amend.pdf, BLA125785 Amendment 013)

Based on the observation made in this experiment (**Table 6**), cell viability was set at a threshold of ≥70% to determine the optimal dsODN concentration for downstream studies. Using data from this experiment, they concluded that either (b) (4) µM or (b)(4) µM dsODN should be used for future studies. The Applicant performed GUIDE-seq on two of the three healthy donor-derived CD34+ HSPCs presented in **Table 7**. The cells were incubated with (b) (4) µM dsODN concentration and the Applicant reported sixteen and five off-target loci in the presence of sufficient on-target read count.

Table 7: Sites Identified in SPY101-RNP treated GUIDE-seq samples across two donors

Donor ID	Disease Status	Cell Viability (%)	dsODN Concentration	On-Target Read Count	Total Off-Target Sites
(b) (6) Healthy	Healthy	71	(b) (4) micromolar	12,095	16
(b) (6) Healthy	Healthy	75	(b) (4) micromolar	11,336	5

Source: Table 8: Sufficient dsODN Incorporation by On-Target Read Count; Table 9: Sites Identified in SPY101-RNP Treated GUIDE-seq Samples Across Two Donors (res-ind-041.pdf, BLA 125785 Amendment 001).

Section V – Confirmatory Testing of Off-Targets Identified in GUIDE-seq by Hybrid Capture Sequencing on Healthy Donor Cells

The four healthy donor samples used in hybrid capture sequencing experiment were different from the donors used in the GUIDE-seq experiment. Details for the four healthy donors used in hybrid capture are presented in **Table 2**. Combining the data from all the healthy-donor-derived HSPC studies, a total of 57 off-target loci were identified from the GUIDE-seq analyses (**Tables 5 & 7**). After removing the redundant loci, confirmatory testing using hybrid capture sequencing was performed on the remaining 52 off-target loci using genomic DNA samples from SPY101 edited HSPCs from 4 healthy donors

(**Tables 5 & 7** that is summarized in **Table 8**). Statistical significance for confirmatory testing was set at ≥0.2% indel frequency difference between edited/treated and control samples. From this analysis, none of the potential 52 off-target loci tested showed statistically significant off-target editing (see **Table 8** below).

Table 8. Hybrid Capture Characteristics and Results for Regions Identified Through GUIDE-seq

Samples Used for GUIDE-seq & dsODN Concentrations	Number of Off-Target Sites Identified from All Experiments	Total Sites Confirmed by Hybrid Capture
(b) (6) Healthy ((b) (4) micromolar) (b) (6) Healthy ((b) (4) micromolar) Donor 1 ((b)(4) micromolar)	52	0

Source: Table 8: Hybrid Capture Characteristics and Results for Regions Identified through GUIDE-seq (res-ind-042.pdf, BLA 125785 Amendment 001)

A summary of all the non-overlapping loci the Applicant identified in their prior experiments (*in silico* and cellular experiments) is presented in **Table 9**. All the sites listed in **Table 9** were tested in four healthy donor samples. Three of the donors were of Hispanic ethnicity and one donor was of African American ethnicity (**Table 2**).

Table 9. Summary of Regions Included in Hybrid Capture Analysis

Cellular Off-Target Studies	Total Sites Identified
GUIDE-seq in healthy donor derived HSPCs (**Table 8**)	52
In silico off-target nomination (**Table 3**)	171

Source: Table 4: Summary of Regions Included in Hybrid Capture Analysis for Each Patient (r264.pdf, BLA125785 Amendment 001); Table 4: GUIDE-seq in Cells from Six Patient Samples (nonclin-info-amend.pdf, Amendment 13, 2023)

The Applicant reported that none of the 223 (171+52) candidate off-target sites were found to have statistically significant off-target editing at the 0.2% threshold difference between treated and control (untreated) samples. Three candidate off-target sites demonstrated >0.2% indels in at least one sample but were not statistically significant across the four donors at that site. In the Study Report CTxSR-036.pdf, the Applicant performed confirmatory testing of 2,094 GUIDE-seq-identified off-target sites for Donor 1. Of this, 1,675 had sufficient sequencing quality to allow analysis. However, none of them displayed a differential editing threshold of >1% between experiment and control samples (see **Table 10**).

Table 10. Hybrid Capture Characteristics and Results for Sites Identified from GUIDE-seq of Donor 1

Number of Mismatches: PAM Types	Number of Off-Target Sites Identified	Number of Sites With Sufficient Quality	Total Sites Confirmed by Hybrid Capture
5: NGG, NAG 4: NGA, NAA, NCG, NGC, NTG, NGT	2094	1675	0

Source: Table 6: Hybrid Capture Characteristics and Results for Sites Identified through Sequence Homology (ctxsr-015.pdf, BLA 125785 Amendment 001 2017).

Reviewer Comment for Sections IV and V: No common off-target loci were identified between the healthy donor derived cell samples. The different concentrations of dsODN

used could potentially interfere with identification of common off-target loci. The Applicant reasoned that the use of different dsODN concentrations was acceptable since they were able to identify adequate numbers of on-target reads in all the experiments irrespective of the dsODN concentrations used. While the Applicant reported >70% cell viability for the dsODN concentration used, they however did not report the on-target editing rates at the indicated experimental parameters. Combining all the off-targets from the healthy donor studies, a total of 52 off-target loci were identified and were tested using hybrid capture sequencing. The Applicant's off-target assessment using GUIDE-seq did not identify overlapping off-target editing y occurring across the heathy donor derived HSPC samples tested. The GUIDE-seq experiment presented here and using SCD samples later is sufficient. However, concerns about off-target editing arising due to sequence heterogeneity remain and will be addressed in a PMR (see Bioinformatics PMR in the review recommendation section above)

The Applicant annotated the off-target loci as either exonic, intronic, or intergenic, using the NCBI RefSeq protein coding gene annotations. This NCBI RefSeq gene set excluded pseudogenes, other non-protein-coding genes, and omitted isoforms which did not have a defined CoDing Sequence (CDS) start and stop site. A gene was defined as the region between the transcription start and end coordinates. If the candidate off-target region did not overlap with genic regions, it was annotated as intergenic. If it overlapped with a gene, then it was determined whether it overlaps any exons of the gene. If an overlap was observed, then the corresponding HGNC gene symbol, Entrez gene name, and a hyperlink to the corresponding GeneCards webpage for function information were provided. Entrez gene names were provided by the R package "org.Hs.eg.db," using the most current Entrez source data available (version: Nov 6 2017). If the region did not overlap an exon, it was annotated as intronic. They provided the off-target loci read counts from the GUIDE-Seq experiment and the 4 independent donor samples that were tested for confirmation of the off-targets as represented in **Table 11**.

Table 11: A snapshot of Hybrid capture data for the SPY101 on-target region and 223 candidate off-target regions included in this study confirm no off-target editing.

Editing sites	Annotation	GUIDE-seq (b) (4) µM	GUIDE-seq (b) (4) µM	GUIDE-seq (b) (4) µM	GUIDE-seq (b) (4) µM	≤3mm	donor with >0.2%	Editing confirmed	p-value	Hybrid capture donor	Hybrid capture donor	Hybrid capture donor	Hybrid capture donor
On-target	Intron	8285	8479	12095	11336	on-target	Yes	Yes	1.10E-04	58.80%	61%	59.90%	49.40%
gs_ot1	Intergenic	ND	ND	ND	3451		No	No					
gs_ot2	Exon Intron (HBD9)	3319	1291	ND	ND		No	No					
gs_ot3	Intergenic	1645	2949	ND	ND		No	No					
gs_ot4	Intergenic	ND	ND	ND	682		No	No					
gs_ot5	Intron	ND	ND	578	ND		No	No					
gs_ot6	Intergenic	483	ND	ND	ND		No	No					
gs_ot7	Intron	ND	ND	445	ND		No	No					
gs_ot8	Intergenic	ND	ND	398	ND		No	No					

Source: Appendix Table 1 (res-ind-042.pdf, BLA 125785 Amendment 001)

Reviewer Comment: The Applicant's annotation of off-target loci identified in the indicated assays is acceptable.

Section VI. – SPY101 Off-Target Assessment by GUIDE-seq (HSPCs from SCD and TDT donors)

The GUIDE-Seq protocol was applied to the six patient-derived HSPCs as described in module 4 of R263.pdf. For testing the suitability of the patient cells for GUIDE-Seq, the Applicant tested the cell viability after (b) (4) and prior to electroporation (EP). The viability data for TDT patients ((b) (6)) and SCD patients ((b) (6)) was reported and reproduced in the plot below (**Figure 4**).

Figure 4: Cell viability on the day of (b) (4) and after electroporation
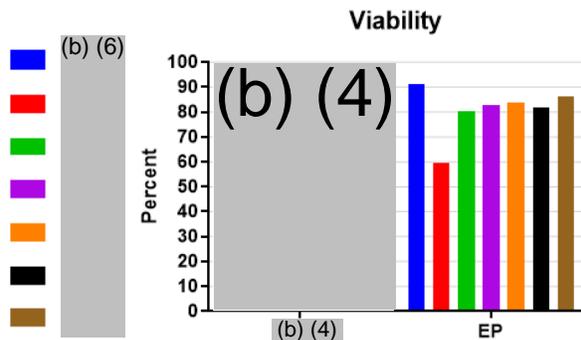
(source: r263.pdf. BLA 125785 Amendment 001)

Based on the cell viability data presented in Figure 4, all samples except (b) (6) had cell viability >70%. Sample (b) (6) was excluded from subsequent analysis. GUIDE-Seq experiments were performed on the remaining 6 patient samples in the presence of (b) (4) µM dsODN with or without SPY101 Cas9 RNP at the indicated concentration as shown in the **Table 12** (see below).

Table 12: Electroporation Condition

Condition	SpCas9 concentration	sgRNA concentration	dsODN concentration
Control	-		(b) (4)
Edited	(b) (4)	(b) (4)	(b) (4)

Source: Table 12 (r263.pdf, BLA 125785 Amendment 001)

In this experiment, the Applicant reported an adequate number of on-target reads and a high on-target editing rate for each sample (see **Table 13,** columns 3 and 5, respectively). A total of 64 off-targets were identified across 6 samples using GUIDE-seq analysis (see **Table 13**, column 4).

Table 13. Editing Rates and On-target Read Counts for GUIDE-Seq Experiments

Samples	dsODN Concentration	On-Target Read Count	Number of Off-Target Sites	On-Target Editing Rate (%)
SCD1	(b) (4) micromolar	16,508	12	83.8
SCD2	(b) (4) micromolar	28,879	13	93.5
SCD3	(b) (4) micromolar	20,857	17	93.6
TDT1	(b) (4) micromolar	23,468	5	92
TDT2	(b) (4) micromolar	23,938	11	92.6
TDT3	(b) (4) micromolar	18,807	6	87.7

Source: Table 6: Editing Rates and On-target Read Counts for GUIDE-Seq Experiments (nonclin-info-amend.pdf, Amendment 11, 2023); Table 4: GUIDE-seq in Cells from Six Patient Samples (nonclin-info-amend.pdf, Amendment 13, 2023)

The Applicant concluded that no common candidate off-targets were identified across all patient samples. Three regions (pgs_(b) (6)_ot5, pgs_(b) (6)_ot11, and pgs_(b) (6)_ot11) overlapped between some pairs of patient samples. There were two overlapping regions (pgs_(b) (6)_ot5 and pgs_(b) (6)_ot11) between this study and one of the 52 candidate GUIDE-Seq regions (gs_ot29) nominated by the previous GUIDE-Seq study (see **Tables 5 & 7**). All these overlapping GUIDE-Seq regions fell within the intergenic ~500bp window of chromosome 3 centromeric region. In total, 64 candidate off-target regions identified from the analysis on the 6 patient sample were included in the hybrid capture libraries.

Reviewer Comment: Applicant's off-target assessment using GUIDE-seq did not identify overlapping off-target loci between the six patient-derived HSPC samples. Applicant's off-target analysis in SCD and TDT cells is acceptable. However, additional in silico analysis will be performed by the Applicant to more comprehensively account for sequence heterogeneity in the PMR (see Bioinformatics PMR in the review recommendation section above)

Section VII. – SPY101 Off-Target Assessment by Hybrid Capture Sequencing on SCD and TDT donor cells

A total of 64 off-targets (see row 4 of **Table 14**) were identified across patient-derived HSPCs using GUIDE-seq analysis. The 223 loci included for testing in all 6 patient samples were loci identified from prior studies (**Sections II & IV**). Specifically, these 223 loci were reported in the *in silico* analysis (171 loci, **Section II: Table 3**) and in the GUIDE-seq analysis of healthy-donor-derived HSPCs (52 loci, **Section IV: Table 8**). The 9 additional off-targets were included after performing variant aware off-target nomination.

Table 14. Summary of Regions Included in Hybrid Capture Analysis

Cellular Off-Target Studies	Total Sites Identified
GUIDE-seq in healthy donor derived HSPCs (**Table 8**)	52
In silico off-target nomination (**Table 3**)	171
Off-targets nominated using variant-aware homology search (**Table 18**)	9
GUIDE-seq in patient HSPCs (**Table 14**)	64

Source: Table 4: Summary of Regions Included in Hybrid Capture Analysis for Each Patient (r264.pdf, BLA125785 Amendment 001); Table 4: GUIDE-seq in Cells from Six Patient Samples (nonclin-info-amend.pdf, Amendment 13, 2023)

The total number of loci identified in each patient using GUIDE-seq analysis is shown in **Table 13**. In this study, the applicant identified an acceptable number of on-target reads and 60-72% on-target editing rates. Specifically, they reported a median on-target reads range of 20,000 to 35,000, and on-target editing rates ranged from 60 to 72% in the edited/treated samples (shown in **Table 15**, columns 2 and 3, respectively). They also reported five loci with >0.2% indel frequency in treated experimental samples compared to controls (**Table 15**, column 4) that annotated to the centromeric region of chr3.

Table 15. Median On-Target Coverage for Each Sample and Sites Identified in Hybrid Capture Sequencing

Samples	Median On-Target Read Counts	Median On-Target Editing Frequency (%)	Number of Sites From Hybrid Capture
SCD1	20278.5	71.8	2
SCD2	22075.5	66.2	0
SCD3	22004.5	71.9	1
TDT1	30457.0	67.2	0
TDT2	34790.0	60.8	1
TDT3	26328.5	71.7	1

Source: Table 6: Median On- and Off-Target Coverage for Each Patient; Table 7: On-Target Editing Rates for Each Patient; Table 8: Summary of Significant Results for Each Patient (r264.pdf, BLA125785 Amendment 001)

In a separate study, the applicant generated control GUIDE-seq data from 8 samples: 6 patient-derived HSPCs (**Section VI: Table 13**) and 2 healthy-donor-derived HSPCs (**Section IV: Table 7**). From the data obtained in this study, the applicant determined that the 4 loci identified in the hybrid capture experiment were from the centromeric

region of chr3 and the 2 additional loci observed in just 2 patient samples did not possess the CRISPR-Cas9 editing features. Furthermore, an additional 13 loci also identified in the GUIDE-seq analyses were present in control samples. Thus, the applicant reasoned these DSBs were naturally occurring hotspots that are independent of CRISPR-Cas9 editing activity.

Finally, the Applicant attributed the off-targets identified in these samples as likely false positives (FPs). To support this conclusion, they reanalyzed the GUIDE-seq data using matched control sequencing data and applied the FP filtering step from GUIDE-seq. From this reanalysis, they reported that all the off-targets loci identified by GUIDE-seq across 6 samples were removed after FP filtering (**Table 16**).

Table 2. Results From GUIDE-seq With and Without "False Positive Filtering"

Samples	dsODN Concentration	Number of Off-Target Sites	Number of Filtered Off-Target Sites
SCD1	(b) (4) micromolar	12	0
SCD2	(b) (4) micromolar	13	0
SCD3	(b) (4) micromolar	17	0
TDT1	(b) (4) micromolar	5	0
TDT2	(b) (4) micromolar	11	0
TDT3	(b) (4) micromolar	6	0

Source: Table 1: Results from GUIDE-seq Nomination in Patient Off-target Study (R263) With and Without "False Positive Filtering" (scd-mcm-followup.pdf, BLA125785 Amendment 040)

Together, from the GUIDE-seq experiments, manual analysis of the off-target loci, and by applying FP filtering, the Applicant concluded that no evidence for *SPY101* off-target editing was identified.

Reviewer Comment: We noted that the on-target editing frequency in the hybrid capture experiment is lower than the on-target editing frequency observed in GUIDE-seq for the same six samples. The lower on-target editing frequency can interfere with optimal editing at off-target loci as they occur at a much lower rate. The Applicant stated that their hybrid capture sequencing was performed at high depth to enable detection of low frequency off-target edits. The Applicant did not identify off-target editing at any of the loci tested using hybrid capture sequencing.

Section VIII. – SPY101 Off-Target Assessment Accounting for Heterogeneity

To address the impact of human genetic variation on off-target activity, the Applicant used the variant information from the 1,000 Genomes Project. In particular, the Applicant used the database that contains WGS data from 2,504 individuals divided into five continental groups: Africa, Americas, East Asia, Europe, and South Asia. The total

number of samples present in each of these continental groups is presented in **Table 17**.

Table 17. Population Ancestry Breakdown of 2,504 Individuals in the 1000 Genomes Project

Continental Group	Total Samples
Africa	661
Americas	347
East Asia	504
Europe	503
South Asia	489

Source: Table 4: Population Ancestry Breakdown of 2504 Individuals in Phase 3 of the 1000 Genomes Project (response-to-bioinformatics-ir-6.pdf, BLA 125785 Amendment 039 2023)

The sponsors stated that they used single nucleotide polymorphism (SNP) variants from dbSNP version 151 that overlapped with the 1000 Genomes Project database and included ~83 million single nucleotide polymorphisms, insertions, and deletions. The Applicant used two different allele frequency cutoffs while performing variant aware off-target nomination. First, they tested ~7 million variants that have a global allele frequency of >10% in the 1,000 Genomes Project database. They then implemented a variant-aware off-target search that used sites with 4 mismatches that would turn into a 3-mismatch site upon inclusion of the variant nucleotide. Alternately, they included sites that, upon inclusion of the variant nucleotide, would result in creation of a PAM sequence with either 3 mismatches or up to 2 mismatches and 1 gap. A schema of the variant-aware homology search is provided in **Figure 5** (see below).

Figure 5. Schematic of Variant-Aware Homology Search
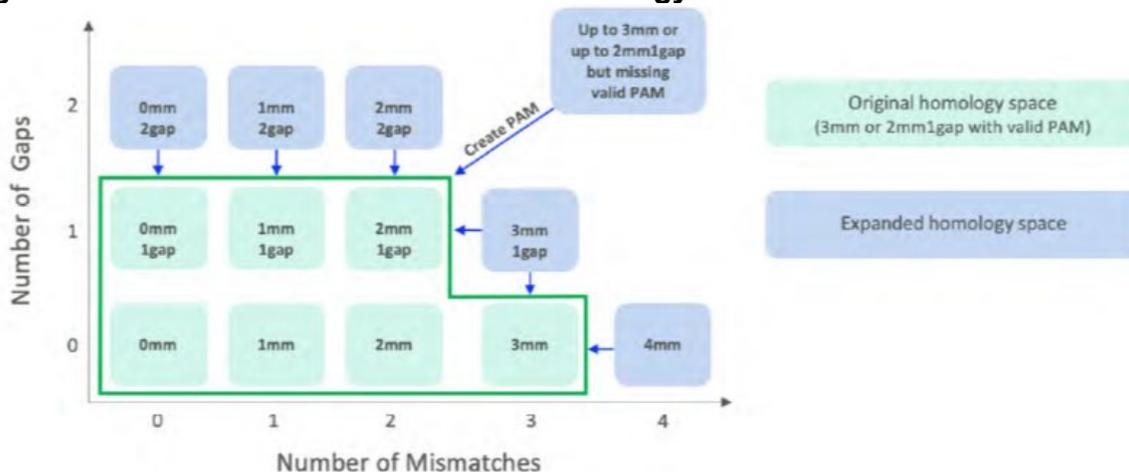
Source: r264.pdf, BLA125785 Amendment 001

From this analysis, they identified nine additional off-target loci with non-canonical PAM sequences. Eight of these off-targets were intergenic and one was intronic. The off-

targets and the associated variant allele frequency and annotation are reported in **Table 18**.

Table 18: Candidate Off-Target Regions Identified Through Computational Variant-Aware Homology Search at 10% Minor Allele Frequency Threshold

Off-Target ID	Genomic Location	Variant Allele Frequency
homva_ot1	Intergenic	42.2%
homva_ot2	Intron	46.2%
homva_ot3	Intergenic	37.7%
homva_ot4	Intergenic	97.1%
homva_ot5	Intergenic	73.7%
homva_ot6	Intergenic	21.5%
homva_ot7	Intergenic	41.0%
homva_ot8	Intergenic	43.4%
homva_ot9	Intergenic	12.4%

Source: Table 10: Candidate Off-Target Regions Identified Through Computational Variant-Aware Homology Search at 10% Minor Allele Frequency Threshold (r264.pdf, BLA 125785 Amendment 001 2022)

In the next analysis, they included ~21 million variants from the 1,000 Genomes Project database with >1% allele frequency cutoff in at least one of the five continental groups. From this analysis, they identified 41 additional candidate off-target loci. **Table 19** lists a subset of 20 off-target loci that were either intronic or exonic, their respective annotations, and the continental group that has an allele frequency >1%. Seven of these off-target loci arose from variants present at frequency >3% in the African continental group.

Table 19. Candidate Off-Target Regions Identified Through Computational Variant-Aware Homology Search at 1% Minor Allele Frequency Threshold

Site Coordinates	Annotation	Gene Name	Homology	Global Variant Frequency	Population-Specific Variant Frequency	Population With Variant Frequency >1%
chr1:100152737-100152758	Intron	LRRC39	2mm	0.44%	1.49%	EUR
chr1:100152738-100152758	Intron	LRRC39	2mm	0.44%	1.49%	EUR
chr10:17367608-17367629	Intron	ST8SIA6	2mm,1gap	0.80%	3.03%	AFR
chr11:108331858-108331879	ExonIntron	ATM1	2mm,1gap	0.46%	1.74%	AFR
chr11:112196597-112196618	Intron	BCO2	2mm,1gap	0.24%	1.19%	EAS
chr12:4651564-4651586	Intron	NDUFA9	2mm,1gap	0.30%	1.13%	AFR
chr12:4651563-4651586	Intron	NDUFA9	2mm,1gap	0.30%	1.13%	AFR
chr12:61748091-61748111	Intron	FAM19A2	2mm,1gap	36.30%|0.12%	47.28%|0.45%	Global,|None
chr12:64064594-64064615	Intron	SRGAP1	2mm,1gap	1.44%|1.44%	5.3%|5.3%	Global,AFR|Global,AFR
chr12:104862519-104862540	Intron	SLC41A2	2mm,1gap	1.38%	4.99%	Global,AFR
chr13:46667776-46667797	Intron	LRCH1	2mm,1gap	0.72%	3.27%	SAS
chr5:131306754-131306775	Intron	CDC42SE2	2mm,1gap	0.20%	1.02%	SAS
chr5:147417885-147417906	Intron	DPYSL3	2mm,1gap	2.34%	8.62%	Global,AFR
chr6:86944935-86944957	Intron	HTR1E	2mm,1gap	25.06%|6.81%	37.22%|13.84%	Global|Global
chr8:53720818-53720839	Intron	ATP6V1H	2mm,1gap	0.44%	1.66%	AFR
chr9:16618034-16618056	Intron	BNC2	2mm,1gap	0.84%|0.10%	3.03%|0.38%	AFR|None
chrX:85346594-85346615	Intron	POF1B	2mm,1gap	0.53%	2.09%	EUR
chr5:148425480-148425502	Intron	FBXO38	3mm	0.92%	3.48%	AFR

Site Coordinates	Annotation	Gene Name	Homology	Global Variant Frequency	Population-Specific Variant Frequency	Population With Variant Frequency >1%
chr7:33483352-33483374	Intron	BBS9	3mm	0.50%	1.82%	AFR
chr7:70395236-70395258	Intron	AUTS2	3mm	0.20%	1.01%	AMR

Source: Table 13-1: 41 Candidate Off-target Sites Identified Through Computational Variant-aware Homology Search at 1% Minor Allele Frequency Threshold (response-to-bioinformatics-ir-6.pdf, BLA 125785 Amendment 039 2023)
Note: "ExonIntron" refers to an off-target cut site located 5-bp away from the intron-exon boundary.
For off-target loci where two genetic variants were included for homology, the frequency and population information are separated by "|."
Abbreviations: AFR, Africa; AMR, Americas; EAS, East Asia; EUR, Europe; SAS, South Asia.

The Applicant reported 50 new off-target loci by including variant information from the 1,000 Genomes Project database. Of these, 20 loci mapped to 18 genes whose intronic/exonic locations were identified as potential off-target loci. One of these off-target loci was 5-bp from the intron-exon junction of *ATM* gene that is known to cause a rare neurodegenerative disease in subjects homozygous for null mutations and is associated with increased cancer risk in subjects with heterozygous disease-causing mutations. Since an edit near the exon-intron junction can potentially disrupt *ATM* function, the Applicant provided a risk assessment for this locus. They stated that the concerning off-target locus uses a non-canonical PAM with a gap that has been shown to have very low off-target editing. Additionally, they stated that this variant has an allele frequency of 1.74% in the target population and would likely be present in individuals as one copy and rarely as two copies. The remaining 17 genes encode for proteins involved in GTPase signaling, mitochondria, DNA repair, etc. Since CRISPR/Cas9 editing has been shown to result in large deletions,[5] such editing events in the intronic loci can potentially disrupt the function of each gene. However, the impact of the loss of any of the 17 genes function on HSPCs has not been evaluated. The Applicant stated that they performed confirmatory testing of these newly nominated 50 off-targets in 1 SCD and 2 TDT donor cells. No significant off-target editing at these loci was observed.

Reviewer Comment: This information was submitted in Amendment 47, in response to Bioinformatics Information Request #6.
In their in silico analysis accounting for heterogeneity, the Applicant reported 50 additional off-target loci. Since the off-target loci were identified in the presence of a variant, confirmatory testing should be performed in samples that are known to carry variants of interest. The applicant reported the presence of 13 variants in at least one sample that was used for confirmatory testing. Of this, three variants were present in just one sample. Since the remaining 37 variants were not present in the samples, an

absence of editing in the confirmatory testing may not necessarily rule out off-target editing at these sites in individuals that harbor these alternate alleles.
*A closer look at the samples included in the African continental subgroup showed that the individuals that were sampled in these groups were predominantly from the western or eastern regions of the African continent. A small amount of the sequencing data was from individuals who were either from the southwest United States (N=61) or the Caribbean in Barbados (N=96). A breakdown of the number of samples from these regions is provided in **Table 20**.*

Table 20. Population Ancestry Breakdown of 661 African Continental Group Individuals in the 1000 Genomes Project

Population-Region	Population Code	Number of Individuals
Esan in Nigeria	AFR	99
Gambian in Western Division, Mandinka	AFR	113
Luhya in Webuye, Kenya	AFR	99
Mende in Sierra Leone	AFR	85
Yoruba in Ibadan, Nigeria	AFR	108
African Caribbean in Barbados	AFR/AMR	96
People with African ancestry in Southwest United States	AFR/AMR	61

Source: Table 4: Population Ancestry Breakdown of 2504 Individuals in Phase 3 of the 1000 Genomes Project (response-to-bioinformatics-ir-6.pdf, BLA 125785 Amendment 039 2023)
Abbreviations: AFR, Africa; AMR, Americas.

African American individuals make up 13.6% of the U.S. population,[6] which is about ~45 million people.[7] SCD is the most common genetic disorder affecting 1 in 500 African Americans and an estimated 100,000 people make up the patient population.[8] Given the large number of patients in United States who are the intended target population for this drug, it is not clear if the sequencing information from the limited number of individuals captured in the 1,000 Genomes Project reference database would sufficiently capture variants that may contribute to an off-target locus. To adequately assess potential variants contributing to off-target editing and empirically test the editing potential at the off-target loci additional studies will be requested as post-marketing requirement (See Bioinformatics PMR in the review recommendation section above).

---

[6] Bureau, USC, Quick Facts, accessed September 14, 2023, https://www.census.gov/quickfacts/fact/table/US/RHI225222.

[7] Christine Tamir, 2021, The Growing Diversity of Black America, Pew Research Center, accessed September 14, 2023, https://www.pewresearch.org/social-trends/2021/03/25/the-growing-diversity-of-black-america/#:~:text=46.8%20million%20people%20in%20the%20U.S.%20identify%20as%20Black,-How%20we%20did.

[8] Sedrak, A and NP Kondamudi, 2023, Sickle Cell Disease, StatPearls, Treasure Island (FL): StatPearls Publishing, Copyright © 2023, StatPearls Publishing LLC.

Module 3.2.R.1

Section I. Robustness Testing of Analytical Procedure to Determine Percent On-Target Editing

Background

The editing frequency in CTX001, was analyzed by using TIDE (Tracking of Indels by Decomposition), a computational approach in which Sanger sequence traces are analyzed from edited and unedited cells using a decomposition algorithm to estimate the editing frequency. The analytical procedure includes (i) the isolation if gDNA from non-edited (donor) and edited (edited pool) test materials, (ii) PCR amplification of the SPY101 locus, (iii) PCR clean-up and bi-directional Sanger sequencing and (iv) TIDE analysis. In addition, (b) (4) are used in the analytical procedure to facilitate the documentation and analysis.

TIDE Analysis Parameters

The Applicant stated that they used default parameters when evaluating on-target editing frequency from the Sanger sequencing data. However, the Applicant pointed out that changes in the settings may occur erroneously by the operator. To address this, the Applicant performed robustness testing to measure the impact of erroneous software parameter changes on the output.

To test the impact of erroneous changes to the default settings, the following settings were selected: (b) (4) . Values for the above settings were changed to analyze the effect on the output data.

Preparing Test and Control Materials

(b) (4)

(b) (4)

Results for Robustness Testing

The Applicant changed the default parameters of the TIDE software and analyzed samples listed in **Table 21**. Their goal was to assess deviations in estimation of indels from the baseline reading. They reported that the deviation from the baseline (Exp12) was (b) (4) in (b)(4) of (b)(4) comparisons showing acceptable robustness of the analytical procedure (**Table 22**).

(b) (4)

A graphical representation of the robustness testing results was also provided in the original submission in the document **ar-51013.pdf**. Technical replicates of all the derivatives showed low deviation from the baseline.

Reviewer Comment: The TIDE performance to detect the editing frequency in the mixed sample is acceptable.

11 pages have been determined to be not releasabl: (b)(4)