Roadmap for the use of Base Editors to Decipher Drug Mechanism of Action

Estel Aparicio Prat (✉️ estel.aparicioprat@pfizer.com)  
Pfizer Inc  https://orcid.org/0000-0003-4477-5839

Dong Yan  
Pfizer Inc

Marco Mariotti  
Brigham and Women's Hospital Department of Medicine

Michael Bassik  
Stanford University

Gaelen Hess  
Stanford

Jean-Philippe Fortin  
Pfizer Inc

Andrea Weston  
Pfizer Inc

Hualin S. Xi  
Pfizer Inc

Robert Stanton  
Pfizer Inc

Research article

Keywords: CRISPR, base editors, mechanism of action, B2AR, GPCR

DOI: https://doi.org/10.21203/rs.3.rs-119015/v1

License: ☝️ This work is licensed under a Creative Commons Attribution 4.0 International License.  
Read Full License
Abstract

Background: CRISPR base editors are powerful tools for large-scale mutagenesis studies. This kind of approach can elucidate the mechanism of action of compounds, a key process in drug discovery. Here, we explore the utility of base editors in an early drug discovery context, and we focus on G-protein coupled receptors.

Results: We set up a pooled mutagenesis screening framework based on a modified version of the CRISPR-X base editor system. We determine optimized experimental conditions for mutagenesis where sgRNAs are delivered by cell transfection or viral infection over extended time periods (>14 days), resulting in high mutagenesis produced in a short region located at -4/+8 nucleotides with respect to the sgRNA match. We thus target the Beta 2 Adrenergic Receptor (B2AR) and employ a 6xCRE-mCherry reporter system to monitor its activity. The results of our screening indicate that residue 184 of B2AR is crucial for its activation. Based on our experience, we then outline the crucial points to consider when designing and performing CRISPR-based pooled mutagenesis screening, including the typical technical hurdles encountered when studying compound pharmacology.

Conclusions: The base editing technology has a great potential to help deciphering the mechanism of action of drugs, and it is a very powerful tool in drug discovery. Here we show an application of pooled mutagenesis screening to study B2AR, and we provide a roadmap for successfully applying this approach to other target proteins.

Background

The use of CRISPR (Clustered Regularly Interspaced Short Palindromic Repeats) with Cas9 (CRISPR associated protein 9) has proven to be a revolutionary technology enabling the rapid introduction of genetic perturbations at targeted genomic regions. Initially, CRISPR was primarily used as a tool for specific and efficient gene knockout [1]–[3]. Since then, CRISPR-based methods have been developed for multiple research purposes including: gene activation/inhibition studies [4], genetic screens [5], cell line engineering [6], and imaging [7] among others. Moreover, various applications of CRISPR emerged both for diagnostics [8], and for therapy (reviewed in [9]), which recently yielded the first positive results in clinical trials [10].

Base editors were initially developed in 2016 fusing a base editor enzyme (such as deaminases) to the inactive form of Cas9 (dCas9) or alternatively to a nickase (creating single strand breaks) [11]–[14]. Cytosine base editors and adenine base editors allow to introduce point mutations without producing a double strand break on the DNA, creating a potential for mutagenesis studies on endogenous proteins. The mutagenesis location is determined by the sgRNA used. The efficiency of base editors has improved greatly over time, but remains dependent on the base editor and the cell types used (reviewed in [15]). Recent advancements in the base editing field have included narrowing of the base-pair window within the gene that undergoes mutagenesis [16], reducing the off-targets [17], and increasing the protospacer
adjacent motif (PAM) compatibility for greater coverage of DNA that is amenable to base editing [18]. Additionally, “prime editing” has been introduced, which directly rewrites information on a target DNA site: here, the effector system consists of an impaired Cas9 fused with a reverse transcriptase, and a peculiar guide RNA specifies both the target site and the desired edit [19].

Mechanism of action (MoA) studies play key roles in drug discovery. They are important for finding safety liabilities and can assist in the approval of new drugs initially identified through phenotypic screening. Knowing the MoA of a small molecule can also help with drug optimization, as specific receptor interactions can be targeted. Traditional MoA studies are performed by exogenous expression of multiple plasmids, including the protein of interest and not conducted on the endogenous protein and residues are mutated one at a time (e.g., by alanine scanning [20]). Recent advances in base editing may enable novel high throughput, targeted and multiplexed mutagenesis workflow to efficiently probe the role of specific residues within virtually any protein families.

Here, we show how a CRISPR-based screening approach can improve MoA studies, reducing time and costs and also allowing to directly target endogenous proteins. We set up a high-throughput targeted multiplexed mutagenesis workflow, in which a reporter system is employed with fluorescence-activated cell sorting and Next Generation Sequencing (NGS) to monitor the functional implications of mutations.

Here, we explore the utility of base editors in a typical drug discovery setting using a pooled CRISPR approach to identify key residues driving small molecule pharmacology of a G-protein coupled receptor (GPCR). GPCRs represent the largest transmembrane receptor family encoded by the human genome and a successful drug target class [21]. Various types of GPCRs modulators including agonists, antagonists and allosteric modulators are being developed as therapeutics. Despite successes in developing several GPCR-based drugs, the identification of safe and efficacious modulators remains a challenging process. The development of novel strategies to dissect the structure-function relationships underlying the complex mechanisms by which ligands bind and activate GPCRs could markedly accelerate the development of effective therapeutics. The fact that emerging base editors’ technologies allow manipulation of endogenously-expressed proteins offers a unique opportunity to study GPCRs in a more native context with more appropriate expression levels and physiological context [22].

Here we use base editors in a pooled format, coupled with a functional selection method and next generation sequencing (NGS) to help elucidate the mechanisms of GPCRs pharmacology and ligand-protein interactions in an endogenous cellular context. Additionally, we built on our experience to outline the driving principles, challenges and limitations of this approach, laying out a roadmap for future MoA studies.

**Results**

**Testing modified CRISPR-X at low MOI targeting GFP**
A crucial feature for pooled screening is using a low multiplicity of infection (MOI) for plasmid delivery, as it is essential that each cell receives only a single guide RNA (sgRNA). Therefore, we first assessed whether low MOI could be successfully employed in our experimental framework. We employed the CRISPR-X base editing system for all our experiments [13], with some variations compared to the original paper, as explained in Methods. CRISPR-X couples the inactive form of Cas9 (dCas9) to AIDΔ (Activation-induced cytidine deaminase) which converts cytidine to uracil. In our experiments to validate low MOI sgRNA delivery, we targeted a Green Fluorescent Protein (GFP) construct. We employed HEK293T cells with stable expression of GFP and the red fluorescent protein mCherry (as a negative control that should not be impacted by the GFP-specific sgRNAs), and targeted GFP using the two most efficient sgRNAs identified in Hess et al. (sgGFP10 and sgGFP1). Additionally, we included two negative control sgRNA (sgNegCtrl1 and sgNegCtrl2) (Fig. 1A). Plasmids were designed for expression of the sgRNA, packaged into a lentivirus, and infected into cells at 5% of infection rate (0.05 MOI). Infected cells were selected for 14 days with puromycin (selection marker in the sgRNA plasmid), then GFP levels were analyzed by flow cytometry. For sgGFP10, 7% of cells lost the GFP fluorescence signal, indicating loss of function mutations (Fig. 1B and C). For sgGFP1, only 2% of the cells showed a diminished GFP signal. These results were in agreement with those obtained in [13] despite the different experimental systems (Hess et al. used K562 cells and an MS2 system). Next, cells were collected, and their genome were analyzed by NGS to assess mutagenesis. We sequenced unsorted cells for every sample, and we additionally analyzed sorted (GFP- mCherry+) cells for sgGFP10 (Fig. 1D). While the flow cytometry analysis indicated the overall proportion of cells with loss of function mutations, the analysis of sequencing data allowed us to determine the precise mutation rate per position at the DNA level.

We found that the sgGFP10 sample had an overall higher mutation rate than sgGFP1 (Fig. 2), consistent with its greater loss of GFP signal. The most highly mutated positions were positions 581 and 586, corresponding to residues 21 and 23 of GFP, with about 6% mutation rate in the unsorted sgGFP10 sample. For the sgGFP10 sorted fraction (GFP negative), mutation rates of 77% and 64% were seen for the same positions, confirming the functional consequence of the mutations at these sites. Additionally, significant mutagenesis was seen at the neighboring residues. For the sgGFP1 unsorted sample, the mutagenesis rate per position was around 1% in the expected region, significantly higher than the background mutation rate (< 0.1%). These results showed that our experimental system could be successfully used at low MOI, so that we proceeded to apply it to study a GPCR.

**Targeting B2AR as a model GPCR**

We decided to study GPCRs due to the considerable interest in this protein family in drug discovery. We picked the Beta 2 Adrenergic Receptor (B2AR) as model GPCR as it has been studied extensively, and studied the MoA of isoproterenol, an agonist molecule known to activate B2AR [23]–[25]. To enable a phenotypic readout of the base editing, it is essential to link the target protein activity to a fluorescent signal that can be employed to sort cells. Here, we monitored GPCR activity through an mCherry reporter whose promoter contained six copies of the cAMP response element (6xCRE-mCherry), so that it becomes expressed upon B2AR activation.
Beforehand, we generated stable HEK293T cell lines with dCas9-AID*Δ by single cell cloning, selecting the clone with the highest levels of the Cas9 protein as probed by Western Blot (Fig. 3A). The 6xCRE-mCherry reporter system was thus infected in HEK293T dCas9-AID*Δ cells, and single cell clones were grown. Next, we used flow cytometry to evaluate various clones and select one with a homogenous response to isoproterenol (i.e., a clone for which supplying isoproterenol resulted in a clear peak shift in fluorescent signal). Furthermore, mCherry fluorescence was evaluated over time using high content imaging (Fig. 3B). This led us to choose a 24 hours isoproterenol treatment to further experiments, to have high expression of mCherry. The intensity in mCherry fluorescence allowed to distinguish the cell populations with and without isoproterenol stimulation, allowing to sort cells based on GPCR activity (Fig. 3C).

Next, we evaluated the specificity of the 6xCRE reporter system for B2AR, since its homologs B3AR and B1AR might potentially be activated by isoproterenol. HEK293T cells do not endogenously express B3AR and the expression of B1AR is very low [26]. We assessed B2AR and B1AR mRNA levels by RT-qPCR, and we knocked them down using siRNAs. 72 hours after siRNA transfection, the mRNA level of B2AR was reduced to 19.2%, while the B1AR siRNA resulted in 22.1% of the normal expression. Supporting specificity of the 6xCRE system for B2AR, the siRNA against B1AR did not downregulate the mCherry signal. On the other hand, the signal was reduced upon B2AR siRNA treatment, although we noted it was not eliminated completely (Fig. 3D). We attributed this resilience to the natural signal amplification in the B2AR pathway, so that even a low level of B2AR expression results in full pathway activation. For a conclusive test of the specificity of 6xCRE reporter system, we also generated single cell clone knockouts (KO) for B2AR using a CRISPR-Cas9 system. Two different KO clones were tested, yielding the same result: after stimulation with isoproterenol, B2AR KO cells did not express any mCherry (Fig. 3E).

Altogether, our results confirmed the specificity of the 6xCRE mCherry reporter system for B2AR activity. They also indicated that an almost complete removal of B2AR may be necessary to eliminate the mCherry expression.

**Defining the optimal conditions for mutagenesis**

We performed a series of experiments to determine the optimal conditions for mutagenesis, evaluating duration of treatment (timepoint), delivery method, and location of mutagenesis. In this phase, we employed four sgRNAs targeting B2AR along with a scrambled control (Fig. 4A, Table 1 in Suppl Mat.) which were delivered to HEK293T cells containing dCas9-AID*Δ and 6xCRE mCherry reporter (clonal). We delivered plasmids either via infection, concentrated infection (concentrating the virus to have a higher titer), transfection or electroporation (see Methods). Genomic DNA was extracted at 4, 7, 9, 11 and 14 days after puromycin selection and sequenced by NGS.

Mutagenesis rates per sample and position are plotted in Fig. 4B, Fig. 5 and Supplementary Figs. 2 and 3. Plots of mutagenesis over time are shown in Fig. 4B and Supplementary Fig. 2 for the short region of the B2AR amplicon targeted by each sgRNA. We observed a consistent pattern of increasing mutagenesis over time, testifying the slow time scale of the mutational process. Notably, the mutagenesis window
along the DNA for a specific sgRNA did not increase over time: no additional sites of mutation appeared at later timepoints, instead we saw the same sites as early timepoints with increased mutagenesis. This allowed us to define a high mutagenesis region around each sgRNA. This region extended from position − 4 to position + 8 in respect to the 5'-end of the sequence matched by the sgRNA (Fig. 4C). When comparing the delivery methods for the sgRNA, similar efficiencies were seen for infection and transfection, while electroporation showed decreased mutagenesis rates (Fig. 5A and Supplementary Fig. 3A). Finally, as expected, we noted that different sgRNAs showed a wide range of mutagenesis rates (Fig. 5B and Supplementary Fig. 3B), ranging from 1% to more than 30%.

**Pooled screening for B2AR**

In an attempt to uncover residues of the B2AR receptor that are important in isoproterenol-mediated activation, we proceeded to perform a pooled screening of B2AR (Fig. 6A). We designed a sgRNA library tiling along this gene, using two separate algorithms to design sgRNAs on the B2AR coding sequence extended by 50 base pairs at both ends. The CHOPCHOP tool [27]–[29] and MIT CRISPR tool [30] gave 192 and 203 candidate sgRNA sequences, respectively. The results were combined taking the CHOPCHOP sequences as the reference and then adding the additional 12 unique sgRNAs from MIT. Single-guide RNAs were evaluated for off-target matches: we retained only those with no less than 3 mismatches with all genomic sites other than B2AR (and without the BbsI restriction enzyme target site, for cloning reasons). The final design consisted of 128 sgRNA along the 1341 base-pairs of the B2AR gene.

The sgRNA library was infected into the HEK293T-dCas9-AID*Δ/6xCRE mCherry reporter at a 0.05 MOI (5% infection rate). Three days after infection, cells were selected using puromycin for 14 days. On the last day, cells were treated with isoproterenol to stimulate B2AR activity. We then sorted cells by flow cytometry, selecting the mCherry-negative populations (cells not activated by isoproterenol). Cell pellets were obtained for genomic DNA extraction both before sorting (presort cells) and after (sorted cells), and non-infected cells were analyzed as negative control (Fig. 6A). The number of cells sorted into the mCherry negative gate was about 3% higher in the sample with the sgRNA library as compared to the no library control for the 3 replicates of the experiment (Fig. 6B), showing a small but consistent effect.

We next analyzed sequence data to assess mutagenesis. More than 95% of the mutations occurred at G and C bases, as anticipated given the use of the CRISPR-X base editor. We observed overall low mutation rates per genomic position (less than 1%, as shown in Fig. 7A). This was expected given that the library contained 128 sgRNAs, and the mutagenesis at any particular position is diluted by the library diversity (e.g. for a library of 100 sgRNAs, only 1/100 of the cells contain a particular sgRNA). Yet, we were surprised to see just a few differences between presort and sorted mCherry negative cells: we expected sorted cells to have a multitude of specific positions at higher mutagenesis rate than presort cells, pinpointing the functional sites of interest. Indeed, our strategy identified a mutation in position 764 corresponding to B2AR residue 184 (Fig. 7C) converting a cysteine to tyrosine (check “Bioinformatic analysis” in Methods for details on the analysis). This cysteine residue is known to form a key disulfide bond with residue 190 helping to form and stabilize the ligand binding site [31]. For the rest of the
mutagenesis observed, caused by the sgRNAs, not relevant differences were seen between the presorted sample and the one sorted for mCherry negative cells (Fig. 7B and D). Therefore, our experiment was successful in identifying functional sites, although with quite narrow extent. Next, we discuss the technical limitations that hindered functional identification at larger scale, which will be instrumental for further applications of this approach.

Discussion

The objective of this study was to evaluate the utility of base editors in early drug discovery, as a tool for understanding mechanisms of compound pharmacology and target biology. This has been traditionally done through alanine scanning and site directed mutagenesis, however the development of CRISPR-based gene editing has created the opportunity for a less time-consuming and more high-throughput approach. After designing an effective reporter system, the delivery a pool of sgRNA at a low MOI followed by flow cytometry allows to separate cells into responding and non-responding cohorts, and analysis of sequencing data pinpoints to those mutations with functional effects.

Here, we demonstrated the feasibility of low MOI infection for pooled editing, and we optimized the experimental conditions. Our results indicated that modified CRISPR-X is effective when sgRNA delivery is performed by transfection or infection, and puromycin selection is performed for 14 days, resulting in high mutagenesis at position −4/+8 from the sgRNA (with mutation rates ranging between 1% and 30% in non-pooled experiments). However, for this to translate functionally, the reporter system must have specific properties, otherwise leading to practical limitations as it occurred in our study.

One plausible confounding factor is the overall ploidy (number of complete sets of chromosomes) of the cell line used. In case of haploid cells, a single mutation effectively alters all protein products for the target gene. In diploid cells, however, there is the potential of synthesizing some number of wild-type products unless both alleles are mutated. This is a feasible scenario if mutation rates are high enough but may become limiting with low rates. The HEK293T cells used here were verified to be triploid (FISH showed 91% of the cells are triploid, data not shown), which diluted mutagenesis over 3 alleles. The cell line is an essential choice for this approach: if available, cells with fewer alleles as possible for the target gene should be used.

The mutagenesis rate is another important factor to be maximized for the success of this approach. This can be increased by modifying the experimental set up. One option is introducing multiple copies of the base editor, either beforehand while creating the cell line, within the sgRNA plasmid, or by co-transfection. Also, the more recently developed base editors like evoCDA1-BE4max, with reported mutagenesis efficiency reaches up to 80%, should be considered for this type of experiment [32].

Finally, we believe that the most important factor in pooled screenings is the choice of the reporter system. One important characteristic is the specificity of signal for the target gene, which was well validated in our case. Another aspect, which resulted most problematic in our work, is the resolution of response between functional and disrupted phenotypes, essential for the accurate sub-setting of
populations by flow cytometry. While the 6xCRE mCherry reporter system has more than sufficient resolution for traditional applications, for the subtle effects studied here its breadth of response makes it difficult to fully differentiate cell populations, as they have overlapping tails in the histograms plotting fluorescence for flow cytometry. Much of this may be due to cells still having at least one wild-type copy of the allele. Indeed, our siRNA experiments showed that only 19% of the mRNA level was sufficient for a significant activation of the B2AR pathway. This is likely attributed to cascade amplification as is typical of many signaling receptor pathways. An assay to monitor a more upstream element of the pathway could help reduce the problems of amplification but could require a tradeoff of sensitivity.

Our design consisted of screening for loss-of-function mutations with a turn-off reporter system (i.e., wherein signal disappears upon occurrence of functional mutations) (Fig. 8). There are alternative designs that may obviate some of the limitations we presented. A key point is that functional mutations should ideally have a dominant effect on the fluorescent signal. This may be obtained by employing a turn-on system, where signal is activated when the interaction between drug and target is disrupted by mutations. This framework may greatly facilitate meaningful sorting and thus functional identification, although designing and optimizing one such system is typically more challenging. To investigate B2AR, we performed early experiments to set up a turn-on system by employing a negative allosteric modulator (compound-15 [33]) together with isoproterenol (agonist) aiming to identify mutations which hindered compound-15’s effect. We also tried some weak/partial agonists for B2AR, such as salbutamol, dopamine, alprenolol and 1,2-dihydroxybenzene. Yet, we saw that the effects obtained were too weak to report significant differences in the 6xCRE mCherry system. Despite the impracticability of this particular system, we strongly recommend to preferentially seek a turn-on strategy for pooled editing MoA studies.

**Conclusions**

Base editing technology has a great potential to quickly help decipher the mechanism of action of multiple drugs. Here, we implemented a modified CRISPR-X based pool editing system for MoA studies, and we applied it to study the B2AR agonist isoproterenol. We successfully conducted a proof-of-principle use of this system, but various limitations restricted its discovery power. Thus, we build on this experience to outline a “roadmap” for future studies of this kind, defining the most important points to consider during study design, and where to spend effort in optimization (see Fig. 8):

1. **After choosing the target gene and drug to study, the most crucial factor is the design of the reporter.** Most compelling, its signal must clearly differentiate the active and inactive states of drug/target interaction. Moreover, the system should be activated uniquely by the protein of interest. Regardless of whether one seeks to find loss-of-function or gain-of-function mutations, it is recommended to set up a turn-on system, in which functional mutations activate (rather than deactivate) fluorescence. Finally, we recommend optimizing the signal resolution of drug response.

2. **Second, the decision of which cell line to use is essential.** This must be chosen in first place to be a suitable system to study the target pathway. Additionally, the ploidy for the target gene should be
considered, preferentially selecting cells with fewer sets of alleles. Ideally, a haploid state would facilitate the identification of functional recessive mutations.

3. Third, one must pick which **base editor** system to employ. The most important feature is the efficiency of mutagenesis: the discovery power of the screening is largely dependent on it, so that it is worth spending effort to optimize it as much as possible. The types of the mutations that the editor can introduce must also be considered, as it determines which sites can be targeted and assayed functionally. When comparing different options, one should evaluate which fraction of coding sequence can be interrogate through non-synonymous mutations.

4. Fourth, one must decide which **delivery** method to use, both for the sgRNA library and for the base editor components. Ideally, available options should be tested during project development, optimizing the signal response to drug and the mutagenesis rate.

Base editing technologies show great promise to decipher the MoA of drugs. Here, we explored optimal conditions of application, and we defined driving principles that will facilitate future studies of this type.

**Methods**

**Cell lines**

HEK293T cells with stable GFP and mCherry were obtained from the Bassik lab at Stanford [13]. HEK293T wild type cells were obtained from ATCC (CRL-3216). All cells were cultured with DMEM, 10% FBS, penicillin-streptomycin and L-glutamine. For drug selection 2 ng/µl was used for puromycin, 4 ng/µl for blasticidin and 300 µg/ml for zeocin.

Transfection was performed with Lipofectamine 2000 (ThermoFisher) or Lipofectamine CRISPRMAX Cas9 Transfection reagent (ThermoFisher). Electroporation was performed using Neon transfection system (ThermoFisher).

**Plasmids for modified CRISPR-X**

pGH389_AiD*-dCas9-BlastR was obtained from the Bassik lab at Stanford [13]. sgRNAs were cloned in plasmid #234 sgRNA-PuroR (no MS2) (pGH020 with Ef1-puro). The fact the sgRNA plasmid does not have MS2 changes the mutagenesis targeting region location. sgRNAs against GFP were sgGFP10 and sgGFP1 from [13]. Oligonucleotides with overhangs compatible with subsequent ligation were designed and annealed followed by ligation into the digested vector. The sequences for the sgRNAs are listed in Table 1 in Supplementary Material. All plasmid sequences were verified using Sanger sequencing.

**Plasmid for monitoring GCPR activity**

6xCRE mCherry plasmid in lentiviral backbone was cloned using lentiviral backbone plasmid pGH235 (gRNA-Zeo, from the Bassik lab) with the insert 6xCRE mCherry SV40 late Poly (A). Full plasmid sequences can be found in supplementary material.
**Isoproterenol treatment**

Cells were seeded and left for 24 hours in regular media. They were starved with media with 0.1% FBS overnight, before the isoproterenol treatment with 1uM stimulation for 24 hours.

**Creation of stable cell line with AID*-dCas9**

HEK293T and HEK293T GFP/mCherry were infected with virus for pGH389_AID*-dCas9-BlastR plasmid produced by Transomics with 10µ g/ml polybrene. 3 days after infection, cells were selected with blasticidin 4 ng/µl for at least 7 days.

**Single cell clones for HEK293T AID*-dCas9**

In order to optimize the homogeneity of the cell population in the study, single cell clones were created for dCas9-AID* using Sony Sorter. Clones were tested for Cas9 expression by Western Blot. One million cells were lysed with RIPA buffer + protease inhibitors + DNaseI. Cas9 antibody used was CRISPR CAS9 MAB 7A9 100UG (Epigentek).

**Generation of KO cell lines for B2AR**

HEK293T were transfected with Synthetic sgRNA AAGAAGGCGCTGCCGTTCCC (Synthego) and TrueCut Cas9 (RNP) (ThermoFisher), using Lipofectamine CRISPRMAX (ThermoFisher). Two days after transfection, genomic DNA was extracted to evaluate the knock-out efficiency for the pool of cells. The amplicon was amplified using oligos AACTCGCACCAGAAGTTGCC and GCACAGCACATCAATGGAAG, sequenced and analyzed using the ICE tool (Synthego). Single cell clones were plated with conditioned media and left to grow for 2 weeks. Genomic DNA was extracted from the single cell clones using QuickExtract DNA Extraction Solution (Lucigen). The same PCR as previously used was done again, using oligos AACTCGCACCAGAAGTTGCC and GCACAGCACATCAATGGAAG, sequenced and analyzed using the ICE tool (Synthego).

**Virus production for sgRNAs**

Takara (Clonetech) Lenti-X Packaging Single Shots (4th generation) were used following manufacture protocol. Supernatant was collected 48–72 hours post-transfection to infect cells (with 10µ g/ml polybrene). For the timepoints experiment, when concentrated infection is indicated, Lenti-X concentrator (Clonetech) was used to concentrate the virus.

**sgRNA library production**

sgRNAs were designed on the B2AR coding sequence along with an additional 50 base pairs at both ends of sequence using the CHOPCHOP tool [27]–[29] and the MIT tool [30]. Results were combined taking CHOPCHOP as reference and adding the 12 unique sgRNAs from MIT results. sgRNAs with off-targets with 3 or more mismatches and no BbsI targets (for cloning purposes) were kept. The final list contained 128 sgRNAs along the 1341 bp of the B2AR coding sequence +/- 50 bp (sgRNA list in Table 2, in
supplementary Excel file “B2AR library sgRNAs.xlsx”). The 128 sgRNAs conceptually target 248 of the 413 amino acids in the B2AR sequencing (60% of residues). To ensure diversity on the sgRNA infection, 9 million cells in a p10 plate were infected with the virus at low MOI (0.05) to ensure that at least 1000 cells were infected with the same sgRNA.

**Generation of mutants**

2–3 days after transfection/infection/electroporation of sgRNA, sustained selection with puromycin 2 ng/µl containing media was used for 14 days, unless shorter times are specified.

**siRNAs**

siRNA ADRB2 - assay ID s1122 (ref. 4427037), Silencer™ Select Negative Control No. 1 siRNA (ref. 4390843) from Thermo Fisher were used at 10 pmol/well in 24-well plate. ON-TARGETplus Human ADRB1 (153) siRNA – SMARTpool (ref L-005425-00-0005), ON-TARGETplus Non-targeting Control Pool (ref. D-001810-10-05) from Dharmacon were used at 50 nM. RNA was extracted 72 h after transfection using RNeasy Plus Mini Kit (Qiagen).

**RT-qPCR**

Retrotranscription was performed with QuantiTect Rev. Transcription Kit (Qiagen) using 500 ng of RNA. qPCR was performed with the Taqman assay (ThermoFisher). Taqman probes used were: Hs00240532_s1 ADRB2, Hs03003658_s1 ADRB2, Hs02330048_s1 ADRB1, Hs00265096_s1 ADRB1, Hs00609046_m1 ADRB3, Hs02786624_g1 GAPDH, Hs04420632_g1 GAPDH (ThermoFisher).

**Flow Cytometry analysis and Sorting**

The Fortessa cytometer was used for analysis and a Sony sorter was used for sorting cells. The sample buffer was PBS-CMF + 10 mM HEPES + 0.5% FBS and the Collection media growth media + 25 mM HEPES.

**High content imaging**

Red fluorescence was measured in a 96-well plate using IncuCyte Zoom (ThermoFisher Scientific).

**Genomic extraction and NGS**

Genomic DNA was extracted from 0.5-2 million cells using the Quick-DNA micro prep kit (Zymo). The targeted loci were PCR amplified from 50 ng of genomic DNA using the primers shown in Table 3 in Supplementary Materials with Kapa HiFi Hotstart ready mix (Kapa Biosystems). The PCR product was run for quality control in an Agilent TapeStation 4200. The concentration was measured using Qubit dsDNA HS Assay kit (ThermoFisher Scientific). Libraries were prepared using Taqmentation and the Nextera XT kit protocol (Illumina). Libraries were sequenced through NextSeq 500 (Illumina) with paired-end reads of length 76 bp.
Bioinformatic analysis

BCL files from the sequencing were converted to FASTQ format using BCL2FASTQ (Illumina). SamTools was then used for alignment to the amplicon sequence using a quality score of 30. Mutagenesis rate per position in the amplicon was calculated as “reads of non-reference base allele/total reads”. Mutagenesis rate of 1 is the maximum and means all alleles have been edited. We also normalized the mutagenesis values per sample obtaining mutagenesis z-scores per genomic position. We focused on those positions where the difference of z-scores was greater than 2 between the mCherry negative sorted sample compared to presorted sample.

Abbreviations

B2AR - Beta 2 Adrenergic Receptor  
Cas9 - CRISPR Associated Protein 9  
CRISPR - Clustered Regularly Interspaced Short Palindromic Repeats  
DNA - Deoxyribonucleic Acid  
GFP – Green Fluorescent Protein  
GPCR - G-protein coupled receptor  
MoA - Mechanism of Action  
MOI - Multiplicity of Infection  
NGS - Next Generation Sequencing  
sgRNA - single guide Ribonucleic Acid

Declarations

Acknowledgements

We thank Hanna Sobon from the Flow Cytometry facility and Justin Boyd from the High Content facility at Pfizer Cambridge for their support.

Authors’ contributions

All authors have read and approved this manuscript. EAP and HSX conceptualized the study. EAP performed most of the experiments and analysis. DY performed RT-qPCRs for siRNA experiments. MM contributed scientific input into experimental design and data interpretation. JPF provided useful guidance in knowledge on GPCRs. AW, HSX and RS oversaw the scientific team and provided crucial
guidance for experimental design, data analysis, data interpretation, and general scientific method. MB and GH designed the base editor technology used in this study.

**Funding**

The entire study, including the design of the study and collection, analysis, and interpretation of data was funded by Pfizer, Inc. The funding bodies played no role in the design of the study and collection, analysis, and interpretation of the data and in writing the manuscript.

**Availability of data and materials**

Sequencing data are available at Sequence Read Archive accession number SUB7836316 under BioProject ID PRJNA649296. http://www.ncbi.nlm.nih.gov/bioproject/649296

**Ethics approval and consent to participate**

There was no animal data presented in this paper, and ethics approval and consent are therefore not applicable.

**Consent for publication**

There was no human subject’s data presented in this paper, and consent is therefore not applicable.

**Competing interests**

The authors declare that they have no competing interests.

**References**

1. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E, “A Programmable Dual-RNA–Guided DNA Endonuclease in Adaptive Bacterial Immunity,” vol. 337, no. August, pp. 816–822, 2012.

2. Cong L, et al., “Multiplex Genome Engineering Using CRISPR/Cas Systems,” no. February, pp. 819–824, 2013.

3. Mali P, et al., “RNA-Guided Human Genome Engineering via Cas9,” vol. 339, pp. 823–827, 2013, doi: 10.1126/science.1232033.

4. Gilbert L, et al. Genome-Scale CRISPR-Mediated Control of Gene Repression and Activation. Cell. 2014;159(3):647–61. doi:10.1016/j.cell.2014.09.029.

5. Wang T, Wei JJ, Sabatini DM, Lander ES. Genetic screens in human cells using the CRISPR-Cas9 system. “ Science. 2014;343(6166):80–4. doi:10.1126/science.1246981.

6. Ran FA, Hsu PPD, Wright J, Agarwala V, Scott D, Zhang F. Genome engineering using the CRISPR-Cas9 system. “ Nat Protoc. 2013;8(11):2281–308. doi:10.1038/nprot.2013.143.
7. Chen B, et al. Dynamic imaging of genomic loci in living human cells by an optimized CRISPR/Cas system. “ Cell. Dec. 2013;155(7):1479–91. doi:10.1016/j.cell.2013.12.001.

8. Broughton JP, et al., “CRISPR – Cas12-based detection of SARS-CoV-2,” Nat. Biotechnol., 2020.

9. Uppada V, Gokara M, Rasineni GK, “Diagnosis and therapy with CRISPR advanced CRISPR based tools for point of care diagnostics and early therapies,” Gene, vol. 656, no. November 2017, pp. 22–29, 2018, doi: 10.1016/j.gene.2018.02.066.

10. Stadtmauer EA, et al., “CRISPR-engineered T cells in patients with refractory cancer,” Science (80-.), 367, 6481, 2020, doi:10.1126/science.aba7365.

11. Komor AC, Kim YB, Packer MS, Zuris Ja, Liu DR. Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. Nature. 2016;533(7603):420–4. doi:10.1038/nature17946.

12. Nishida K, et al., “Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems,” Science (80-.), 353, 6305, 2016, doi:10.1126/science.aaf8729.

13. Hess GT, et al. Directed evolution using dCas9-targeted somatic hypermutation in mammalian cells. “ Nat Methods. 2016;13:12, pp. 1036–42. doi:10.1038/nmeth.4038., ", , no.

14. Gaudelli NM, et al., “Programmable base editing of A•T to G•C in genomic DNA without DNA cleavage,” Nature, vol. 551, no. 7681, pp. 464–471, Nov. 2017, doi: 10.1038/nature24644.

15. Rees HA, Liu DR. Base editing: precision chemistry on the genome and transcriptome of living cells. Nat Rev Genet. 2018;19(12):770–88. doi:10.1038/s41576-018-0059-1.

16. Kim YB, Komor AC, Levy JM, Packer MS, Zhao KT, Liu DR. Increasing the genome-targeting scope and precision of base editing with engineered Cas9-cytidine deaminase fusions. Nat Biotechnol. 2017;35(4):371–6. doi:10.1038/nbt.3803.

17. Akcakaya P, et al. In vivo CRISPR editing with no detectable genome-wide off-target mutations. Nature. 2018;561(7723):416–9. doi:10.1038/s41586-018-0500-9.

18. Hu JH, et al., “Evolved Cas9 variants with broad PAM compatibility and high DNA specificity,” Nature, 1–24, 2018, doi:10.1038/nature26155.

19. Anzalone AV, et al. Search-and-replace genome editing without double-strand breaks or donor DNA. Nature. 2019;576(7785):149–57. doi:10.1038/s41586-019-1711-4.

20. Morrison KL, Weiss GA. Combinatorial alanine-scanning. Curr Opin Chem Biol. 2001;5(3):302–7. doi:10.1016/S1367-5931(00)00206-4.

21. Salon JA, Lodowski DT, Palczewski K. The Significance of G Protein-Coupled Receptor Crystallography for Drug Discovery. Pharmacol Rev. 2011;63(4):901–37. doi:10.1124/pr.110.003350.901.

22. Insel PA, et al. G protein-coupled receptor (GPCR) expression in native cells: ‘Novel’ endoGPCRs as physiologic regulators and therapeutic targets. Mol Pharmacol. 2015;88(1):181–7. doi:10.1124/mol.115.098129.
23. Rasmussen SGF, et al. Crystal structure of the human β2 adrenergic G-protein-coupled receptor. Nature. 2007;450:7168, pp. 383–7. doi:10.1038/nature06325., “,”, no.

24. Rosenbaum DM, et al., “GPCR Engineering Yields High-Resolution Structural Insights into b2-Adrenergic Receptor Function,” Science (80-.), vol. 318, no. November, pp. 1266–1273, 2007.

25. Cherezov V, et al., ”High Resolution Crystal Structure of an Engineered Human β2 Adrenergic G protein-Coupled Receptor,” Science (80-.), vol. 318, no. 5854, pp. 1258–1265, 2007, doi: 10.1126/science.1150577.

26. Atwood BK, Lopez J, Wager-Miller J, Mackie K, Straiker A. Expression of G protein-coupled receptors and related proteins in HEK293, AtT20, BV2, and N18 cell lines as revealed by microarray analysis. BMC Genom. 2011;12(1):14. doi:10.1186/1471-2164-12-14.

27. Labun K, Montague TG, Krause M, Torres Cleuren YN, Tjeldnes H, Valen E. CHOPCHOP v3: Expanding the CRISPR web toolbox beyond genome editing. Nucleic Acids Res. 2019;47:no. W1, pp. W171–4. doi:10.1093/nar/gkz365.

28. Labun K, Montague TG, Gagnon Ja, Thyme SB, Valen E. CHOPCHOP v2: a web tool for the next generation of CRISPR genome engineering. Nucleic Acids Res. 2016;44:no. May, p. gkw398. doi:10.1093/nar/gkw398.

29. Montague TG, Cruz JM, Gagnon JA, Church GM, Valen E, “CHOPCHOP: A CRISPR/Cas9 and TALEN web tool for genome editing,” Nucleic Acids Res., vol. 42, no. W1, pp. 401–407, 2014, doi: 10.1093/nar/gku410.

30. Hsu PD, et al. DNA targeting specificity of RNA-guided Cas9 nucleases. Nat Biotechnol. 2013;31(9):827–32. doi:10.1038/nbt.2647.

31. Isin B, Estiu G, Wiest O, Oltvai ZN, “Identifying Ligand Binding Conformations of the β2-Adrenergic Receptor by Using Its Agonists as Computational Probes,” PLoS One, vol. 7, no. 12, p. e50186, Dec. 2012, doi: 10.1371/journal.pone.0050186.

32. Thuronyi BW, et al. Continuous evolution of base editors with expanded target compatibility and improved activity. Nat Biotechnol. 2019. doi:10.1038/s41587-019-0193-0.

33. Liu X, et al., “Mechanism of intracellular allosteric β2AR antagonist revealed by X-ray crystal structure,” Nature, vol. 548, no. 7668, pp. 480–484, Aug. 2017, doi: 10.1038/nature23652.

Figures
Figure 1

Testing modified CRISPR-X at low MOI targeting GFP. A. Design of the experiment. B. Flow Cytometry analysis plotting GFP and mCherry fluorescent intensities. C. The percentages of cells in Q4 quadrant for figure 1B are graphed. The mean of three independent replicates with the standard deviation is shown. D. In purple, the GFP-negative cells that were sorted for sample sgGFP10.
Figure 1

Testing modified CRISPR-X at low MOI targeting GFP. A. Design of the experiment. B. Flow Cytometry analysis plotting GFP and mCherry fluorescent intensities. C. The percentages of cells in Q4 quadrant for figure 1B are graphed. The mean of three independent replicates with the standard deviation is shown. D. In purple, the GFP-negative cells that were sorted for sample sgGFP10.
Figure 2

Mutagenesis rates with modified CRISPR-X at low MOI targeting GFP Mutagenesis rate per position for GFP amplicon region targeted with sgRNAs. Mutagenesis rate is calculated by ratio of “reads of non-reference base allele/total reads”, with a ratio of 1 indicating a complete edition of all alleles. The mean and standard deviation of three independent replicates are shown. A zoom-in of the selected region is shown below for more details.
Figure 2

Mutagenesis rates with modified CRISPR-X at low MOI targeting GFP. Mutagenesis rate per position for GFP amplicon region targeted with sgRNAs. Mutagenesis rate is calculated by ratio of "reads of non-reference base allele/total reads", with a ratio of 1 indicating a complete edition of all alleles. The mean and standard deviation of three independent replicates are shown. A zoom-in of the selected region is shown below for more details.
Figure 3

Targeting B2AR as a model GPCR. A. Western Blot for Cas9 expression levels on HEK293T wild type cell line, HEK293T cell line + dCas9-AID* as a pool and the single cell clone selected with the highest expression of Cas9. Cropped blot where arrow indicates (last lane). Full uncropped blot found in Supplementary Figure 1. B. Clone HEK293T dCas9-AID* was infected with 6xCRE-mCherry reporter system and single cell clones were evaluated. The most homogeneous clones for mCherry expression was selected. mCherry red fluorescence was measured with high content imaging using a cell incubator imaging system. Cells treated with and without isoproterenol were monitored for 30 hours after stimulation. C. mCherry fluorescence intensity histogram measured by Flow Cytometry 24 hours after isoproterenol treatment or its absence on the same single cell clone as in B. D. Left, mCherry fluorescence intensity histogram measured by Flow Cytometry on cells transfected with control siRNA or siRNA against B2AR. After siRNA treatment, cells were stimulated with isoproterenol. Right, mCherry
fluorescence intensity histogram measured by Flow Cytometry on cells transfected with control siRNA or siRNA against B1AR. After siRNA treatment, cells were stimulated with isoproterenol. E. mCherry fluorescence intensity histogram measured by Flow Cytometry on knock-out clone for B2AR, with and without stimulation of isoproterenol.

**Figure 3**

Targeting B2AR as a model GPCR A. Western Blot for Cas9 expression levels on HEK293T wild type cell line, HEK293T cell line + dCas9-AID as a pool and the single cell clone selected with the highest expression of Cas9. Cropped blot where arrow indicates (last lane). Full uncropped blot found in Supplementary Figure 1. B. Clone HEK293T dCas9-AID was infected with 6xCRE-mCherry reporter system and single cell clones were evaluated. The most homogeneous clones for mCherry expression was selected. mCherry red fluorescence was measured with high content imaging using a cell incubator imaging system. Cells treated with and without isoproterenol were monitored for 30 hours after
stimulation. C. mCherry fluorescence intensity histogram measured by Flow Cytometry 24 hours after isoproterenol treatment or its absence on the same single cell clone as in B. D. Left, mCherry fluorescence intensity histogram measured by Flow Cytometry on cells transfected with control siRNA or siRNA against B2AR. After siRNA treatment, cells were stimulated with isoproterenol. Right, mCherry fluorescence intensity histogram measured by Flow Cytometry on cells transfected with control siRNA or siRNA against B1AR. After siRNA treatment, cells were stimulated with isoproterenol. E. mCherry fluorescence intensity histogram measured by Flow Cytometry on knock-out clone for B2AR, with and without stimulation of isoproterenol.
Defining the optimal mutagenesis conditions

A. Experiment design: 4 sgRNA were designed along B2AR gene (P18-P21) and a negative control (P13) targeting a non-specific sequence. Positions on the B2AR amplicon are specified in the figure.

B. Mutagenesis rate per position for B2AR amplicon shown at different timepoints, per sgRNA.

C. Diagram of a sgRNA, with the PAM sequence after. Highlighted in red is the region with increased mutagenesis rate, indicating high activity of the base editor.
Figure 4

Defining the optimal mutagenesis conditions A. Experiment design: 4 sgRNA were designed along B2AR gene (P18-P21) and a negative control (P13) targeting a non-specific sequence. Positions on the B2AR amplicon are specified in the figure. B. Mutagenesis rate per position for B2AR amplicon shown at different timepoints, per sgRNA. C. Diagram of a sgRNA, with the PAM sequence after. Highlighted in red is the region with increased mutagenesis rate, indicating high activity of the base editor.
Figure 5

Defining the optimal mutagenesis conditions (continuation) A. Mutagenesis rate per position for B2AR amplicon showing different delivery methods of the sgRNA. B. Mutagenesis rate per position for B2AR amplicon showing different sgRNA efficiencies.
Figure 5

Defining the optimal mutagenesis conditions (continuation) A. Mutagenesis rate per position for B2AR amplicon showing different delivery methods of the sgRNA. B. Mutagenesis rate per position for B2AR amplicon showing different sgRNA efficiencies.
Figure 6

Pooled screening for B2AR A. Design of the experiment. The sgRNA library containing 128 sgRNAs was infected into clone HEK293T + dCas9-AID*Δ + 6xCRE mCherry reporter at about 0.05 MOI. Cells were selected with puromycin and treated with isoproterenol, to stimulate B2AR activity. Cell pellets were obtained for genomic extraction before sorting (presort cells) and sorted for the mCherry negative in both
populations (the infected with the B2AR library and the non-infected cells). B. mCherry fluorescence intensity histogram at the Sorter where the mCherry negative gates were sorted out.

**Figure 6**

Pooled screening for B2AR A. Design of the experiment. The sgRNA library containing 128 sgRNAs was infected into clone HEK293T + dCas9-AID* + 6xCRE mCherry reporter at about 0.05 MOI. Cells were selected with puromycin and treated with isoproterenol, to stimulate B2AR activity. Cell pellets were
obtained for genomic extraction before sorting (presort cells) and sorted for the mCherry negative in both populations (the infected with the B2AR library and the non-infected cells). B. mCherry fluorescence intensity histogram at the Sorter where the mCherry negative gates were sorted out.

Figure 7

Mutagenesis rates for pooled screening for B2AR. A. Mutagenesis rate per position for B2AR amplicon. The mean and standard deviation of three independent replicates are shown. B, C, D. Zoom-ins of the
regions in B2AR amplicon marked in A.

**Figure 7**

Mutagenesis rates for pooled screening for B2AR A. Mutagenesis rate per position for B2AR amplicon. The mean and standard deviation of three independent replicates are shown. B, C, D. Zoom-ins of the regions in B2AR amplicon marked in A.
Figure 8

Graphical conclusions Roadmap for future studies of this kind, defining the most important points to consider during study design, and where to spend effort in optimization.
Figure 8

Graphical conclusions Roadmap for future studies of this kind, defining the most important points to consider during study design, and where to spend effort in optimization.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- AparicioPratSuplFigures20200727.pdf
- AparicioPratSuplFigures20200727.pdf
- AparicioPratSupplMaterial20200727.pdf
- AparicioPratSupplMaterial20200727.pdf
- B2ARlibrarysgRNAs.xlsx
- B2ARlibrarysgRNAs.xlsx