Trends in Genetics

CellPress
REVIEWS

Spotlight

Base Editing Landscape Extends to Perform Transversion Mutation

Kutubuddin A. Molla [1,*,@]
Yiping Qi [2,3]
Subhasis Karmakar,[1]
and Mirza J. Baig[1]

Base editors have drawn considerable academic and industrial attention in recent years because of their ability to alter single DNA bases with precision. However, the existing cytosine and adenine base editors can only install transition mutations. Three recent studies (Kurt *et al.*, Zhao *et al.*, and Chen *et al.*) expand the base editing toolbox by developing cytosine transversion base editors.

Why We Need a Transversion Base Editor?

The CRISPR-Cas9 system has transformed the genome engineering field. This system is highly efficient in causing targeted knockouts by generating small indels in the genome. Precise modification from one nucleotide to another needs an adequate supply of a donor template and induction of the homology-directed repair (HDR) pathway [1]. The invention of cytosine base editors (CBEs) and adenine base editors (ABEs) has enabled us to perform targeted C-to-T and A-to-G conversion in DNA or RNA without donor templates [2–5]. Both CBEs and ABEs have been used extensively in diverse organisms to create or correct point mutations for different applications [5,6]. However, CBEs and ABEs only catalyze base transition (purine to purine or pyrimidine to pyrimidine) and can be applied to achieve only four out of 12 possible base substitutions. Nonetheless, many biological, therapeutic, and crop improvement applications require installation of transversion mutations (purine to pyrimidine or vice versa). For example, correcting nearly half of human pathogenic point mutations would require base transversions [6]. The scientific community has been seeking a transversion base editor for a long time and it remained elusive until the recent development of base editing platforms that can convert C-to-G in mammalian cells [7–9] and C-to-A in bacteria [8].

How Does the C-to-G Base Editor (CGBE) Convert a C Base to a G Base?

The CGBE consists of a Cas9 nickase (nCas9-D10A) fused to a cytidine deaminase and a uracil DNA glycosylase (UNG) (Figure 1A). Guided by RNA, cytidine deaminase converts a target C to U. UNG eliminates U from the DNA and creates an apurinic/apyrimidinic (AP) site. AP site creation, coupled with nicking at the nonedited strand by nCas9, initiates DNA repair and replication that leads to favored insertion of G at the AP site. These events result in a C-to-G editing outcome (Figure 1C). However, it remains unknown how the introduction of G is preferred over the other two bases. By contrast, CBEs contain a UNG inhibitor (UGI) that prevents the removal of U and subsequently increases the likelihood of a C-to-T editing outcome [2,3]. Structurally, CBE and CGBE are similar except that CGBE contains UNG rather than UGI.

To achieve C-to-G editing in mammalian cells, Zhao *et al.* fused rat APOBEC1 cytidine deaminase to the amino (N) terminus of nCas9 and UNG to the C terminus (APOBEC1-nCas9-UNG) [8], while Kurt *et al.* linked both UNG and APOBEC1 at the N terminus (UNG-APOBEC1-nCas9) [7]. A similar hypothetical CGBE design that relies on translesion synthesis has also been developed [10]. Kurt *et al.* used a mutant version of rAPOBEC1 (R33A) [7]. They have also developed a miniCGBE, with comparable but moderately lower editing efficiency, by deleting UNG from the original CGBE [7]. By contrast, Chen *et al.* fused XRCC1, a base excision repair (BER) protein, at the C terminus of APOBEC1-nCas9 to construct a CGBE [9]. Hypothetically, this platform (rAPOBEC1-nCas9-XRCC1) works in a slightly distinct way from the other two (Figure 1A) [9]. Once APOBEC1 induces C-to-U conversion, cellular UNG creates an AP site by removing U. Then XRCC1 recruits other BER proteins to repair the AP site, resulting in G as the predominant product [9].

How Efficient Are They?

Editing frequencies were highly variable across the sites in all three studies [7–9]. Zhao *et al.* tested 30 loci with the target C at a suitable window and reported an editing efficiency of 5.3–53% [8]. Kurt *et al.* showed that the efficiency varies from 3–71.5% at 25 sites tested [7]. The CGBE platform developed by Chen *et al.* exhibited 15 ± 7% efficiency in human cells [9].

A Similar Platform Enables C-to-A Editing in Bacteria

Inspired by their initial findings on UNG initiated C-to-A conversion in *Escherichia coli* cells treated with nCas9-AID, Zhao *et al.* constructed a UNG-nCas9-AID platform (Figure 1B) [8]. Activation-induced deaminase (AID) here being a cytidine deaminase from sea lamprey. UNG-nCas9-AID yielded 80% or higher C-to-A editing efficiency at four different sites within the *lacZ* sequence in *E. coli*, with more than 90% specificity. Except for the position and type of cytidine deaminase, this platform is similar to the CGBE platform [8]. However, it is not clear why A is preferred in bacterial cells and G in mammalian cells, giving rise to different editing outcomes.

Narrow Activity Window and Sequence Context Preference

The activity window is a range of bases in the protospacer sequence, which is favorable for a base editor's editing activity

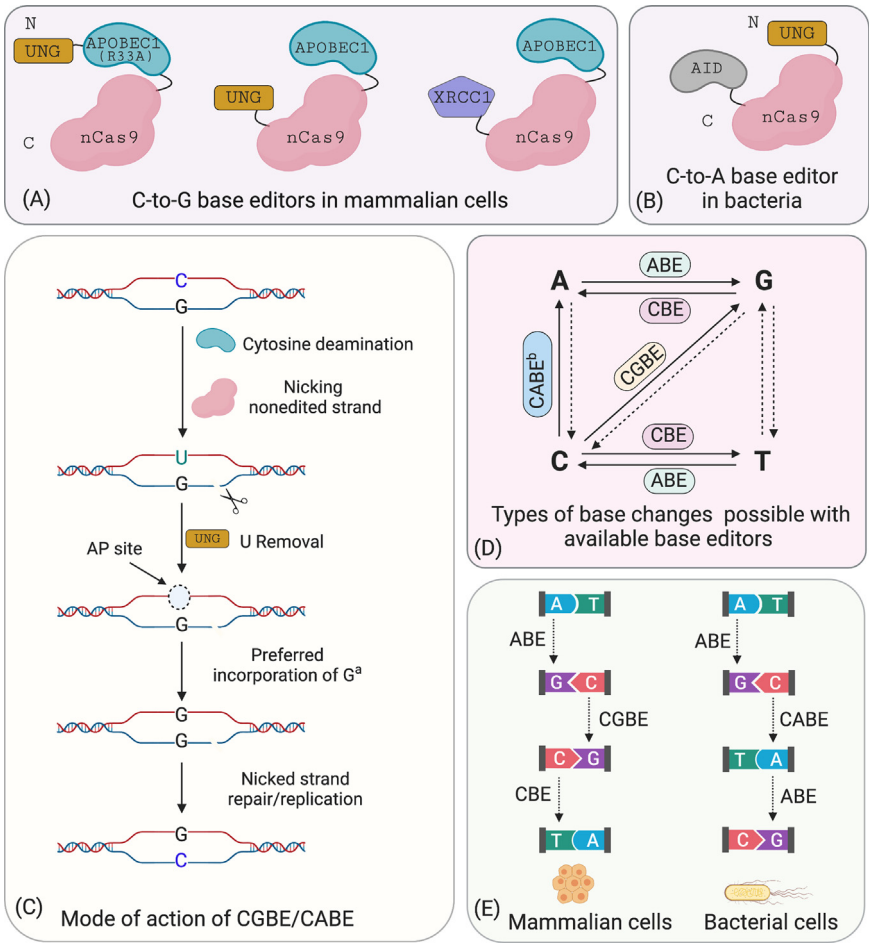

Figure 1. Structure and Mode of Action of Cytosine Transversion Base Editors. (A) Schematic of C-to-G base editor (CGBE) platforms developed by Kurt *et al.* [7], Zhao *et al.* [8], and Chen *et al.* [9], respectively (from left to right). C at position 6 of the protospacer, counting NGG protospacer adjacent motif (PAM) as 21–23, is edited with high efficiency. (B) C-to-A base editor (CABE) developed by Zhao *et al.* [8]. (C) Schematic diagram showing a probable mechanism of action of CGBE. In CABE, AID deaminates the target C. [a]In the bacterial cell, A is the most preferred base incorporated in an AP site causing C to A alteration. (D) Diagram showing types of achievable point mutations with the available base editors. [b]At present, CABE is feasible only in bacterial cells. Conversion showed in dotted arrow are not achievable in one step. (E) Hypothetical model showing multistep base editing to install T-to-G, T-to-A, A-to-C, and A-to-T in mammalian and bacterial cells (inspired by Zhao *et al.* [8]). Abbreviations: A, alanine; ABE, adenine base editor; AID, cytidine deaminase from sea lamprey; AP site, apurinic/apyrimidinic site; APOBEC1, rat APOBEC1 cytidine deaminase; APOBEC1 (R33A), mutated rat APOBEC1; C, carboxy terminus of nCas9; CBE, cytosine base editor; D, aspartic acid; N, amino terminus of nCas9; nCas9, nickase Cas9 (D10A); R, arginine; UNG, uracil DNA glycosylase; XRCC1, x-ray repair cross-complementing protein 1. This figure was created using BioRender.

[5]. It determines the targetability of a nucleotide by a base editor. Unlike most CBEs and ABEs that prefer the target base at position 4–8 in the protospacer [5], CGBEs showed optimum activity at position 6 [7–9]. Although positions 5 and 7 exhibited appreciable editing [7,9], CGBEs act mainly on the C6 position. However, CGBEs are not equally effective at all C6 sites. Types of bases flanking C6 play an important role and may influence efficiency. Most efficiently edited C6 sites are those flanked by As and/or Ts [7]. Like CBEs, CGBE also showed very low to nil editing if a 5′G precedes the target C [2,7]. rAPOBEC1 deaminase, used in both the platforms, dislikes a GC context.

Concluding Remarks

The three studies, briefly discussed here, have enriched the base editing toolbox by enabling C-to-A editing in bacteria and C-to-G editing in mammalian cells [7–9] (Figure 1D). These developments are indeed a leap forward in precise genome editing for basic biology, therapeutics, and crop improvement. For example, CGBEs could potentially correct ~3700 human disease-causing mutations [6]. Many SNPs are associated with stress-resistant traits in plants. Once adopted for plants, CGBEs could be utilized to install a fraction of those SNPs in order to develop stress-tolerant crops. C-to-G editing enables us to substitute 18 different amino acids and generate two premature stop codons, that were not achievable with CBEs and ABEs. Similarly, C-to-A editing enables 19 amino acid substitutions and the generation of two stop codons in bacterial cells. It is evident that C-to-G and C-to-A editors have greatly expanded our ability to manipulate genomes at single-base resolution in mammalian and bacterial cells, respectively.

However, there are many challenges to overcome before realizing their full potential. The most notable limitation of CGBE is its extremely narrow activity window, mainly confined to protospacer's sixth position. Using orthologous or engineered Cas9 with alternative protospacer adjacent motif (PAM) requirements, deaminase engineering, and linker length variation may help broaden the editing window. Editing efficiency varies across cell types and reduces drastically in GC context [7]. More research on architectural fine-tuning is needed to overcome those limitations. If applied successively, base editors could perform other base transversions through multiple rounds

of editing (Figure 1E). Without doubt, these initial reports [7–9] will motivate researchers to optimize the platforms for improved targetability, efficiency, and specificity.

Acknowledgments

K.M. and M.J.B. would like to acknowledge funding from ICAR, New Delhi, in the form of the Plan Scheme-'Incentivizing Research in Agriculture' project and support from the Director, NRRI. S.K. is supported by the Department of Biotechnology, Government of India -RA program. Y.Q. acknowledges funding from the National Science Foundation, Plant Genome Research Program grants (IOS-1758745 and IOS-2029889), US Department of Agriculture, Biotechnology Risk Assessment Grant Program competitive grants (2018-33522-28789 and 2020-33522-32274), Foundation for Food and Agriculture Research grant (award no. 593603), and Syngenta.

[1]Indian Council of Agricultural Research (ICAR)-National Rice Research Institute (NRRI), Cuttack-753006, India
[2]Department of Plant Science and Landscape Architecture, University of Maryland, College Park, MD 20742, USA
[3]Institute for Bioscience and Biotechnology Research, University of Maryland, Rockville, MD 20850, USA

*Correspondence:
kutubuddin.molla@icar.gov.in (K.A. Molla).
@Twitter: @Kutub_joy (K.A. Molla).

https://doi.org/10.1016/j.tig.2020.09.001

References

1. Molla, K.A. and Yang, Y. (2020) Predicting CRISPR/Cas9-induced mutations for precise genome editing. *Trends Biotechnol.* 38, 136–141
2. Komor, A.C. *et al.* (2016) Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424
3. Nishida, K. *et al.* (2016) Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353, aaf8729
4. Gaudelli, N.M. *et al.* (2017) Programmable base editing of A*T to G*C in genomic DNA without DNA cleavage. *Nature* 551, 464–471
5. Molla, K.A. and Yang, Y. (2019) CRISPR/Cas-mediated base editing: technical considerations and practical applications. *Trends Biotechnol.* 37, 1121–1142
6. Rees, H.A. and Liu, D.R. (2018) Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788
7. Kurt, I.C. *et al.* (2020) CRISPR C-to-G base editors for inducing targeted DNA transversions in human cells. *Nat. Biotechnol.* Published online July 20, 2020. https://doi.org/10.1038/s41587-020-0609-x
8. Zhao, D. *et al.* (2020) Glycosylase base editors enable C-to-A and C-to-G base changes. *Nat. Biotechnol.* Published online July 20, 2020 https://doi.org/10.1038/s41587-020-0592-2
9. Chen, L. *et al.* (2020) Precise and programmable C: G to G: C base editing in genomic DNA. *bioRxiv.* Published online July 21, 2020 https://doi.org/10.1101/2020.07.21.213827
10. Gajula, K.S. (2019) Designing an elusive C• G→ G• C CRISPR base editor. *Trends Biochem. Sci.* 44, 91–94