Genome Engineering with Designer Programmable Proteins

In the beginning:
Recombinant DNA

Transfer and Cloning of the Insulin Gene

Need for precise engineering of biological systems

Targeted DNA breaks facilitate alteration of the genome

Need for:
- Functional testing of genetic mutations
- Treatment of human diseases
- Engineering of synthetic biological systems
Targeted nucleases can facilitate genome editing

Customizing DNA Binding Proteins

Helix-turn-helix
Zinc finger
Leucine zipper
Winged helix
Winged helix turn helix
Helix loop helix
HMG-box
TAL effector
B3 domain
Immunoglobulin fold

Step 1: Targeting the genome
1. Zinc finger domains
2. TAL effectors
3. CRISPR/Cas9

ABSTRACT A long-term goal in the field of restriction-modification enzymes has been to generate restriction endonucleases with novel sequence specificities by mutating or engineering existing enzymes. This will avoid the increasingly onerous task of extensive screening of bacteria and other microorganisms for new enzymes. Here, we report the deliberate creation of novel site-specific endonucleases by linking two different zinc finger proteins to the change domain of Fok I endonuclease. Both fusion proteins are active and under optimal conditions cleave DNA in a sequence-specific manner. Thus, the modular structure of Fok I endonuclease and the zinc finger motif makes it possible to create "artificial" endonucleases that will cut DNA near a predetermined site. This opens the way to generate many new enzymes with tailor-made sequence specificities desirable for various applications.

Since their discovery nearly 25 years ago (1), type II restriction enzymes have played a crucial role in the development of the recombinant DNA technology and the field of molecular biology. A long-term goal in the field of restriction-modification enzymes has been to generate restriction endonucleases with novel sequence specificities by mutating or engineering existing enzymes (1). Our efforts on prototypic fragments of Fok I endonuclease (from Salmonella minnesota, belonging to the type II restriction enzyme family) have resulted in the development of a novel DNA-binding domain and a C-terminal domain with non-specific DNA-cleavage activity (2-7). The modular structure of Fok I endonuclease suggested that it might be feasible to construct hybrid endonucleases with novel sequence specificities by linking other DNA-binding proteins to the digestion domain. Recently, we reported the construction of the first "artificial" restriction endonuclease gene by linking the Drosophila E. coli homodomain to the digestion domain of Fok I (8).

Unlike the homodomain, the zinc finger proteins, because of their modular structure, offer as attractive frames for designing chimeric restriction enzymes with tailor-made sequence specificities. The CHIVR zinc finger proteins are a class of DNA-binding proteins that contain sequences of the
Biochemical analysis of double-strand break sites

How do zinc finger nucleases cleave the target substrate?

The modern ZFN dimer strategy improves specificity

Spectacular success in modifying the genomes of multiple organisms

Beyond cutting DNA

> Uromov et al., Nat. Rev. Genetics 2010

> Papworth et al., Gene 2006
Step 2: Effector Domains

1. Efficient DNA cleavage
2. Transcriptional Control
3. Epigenetic Modulation
4. Inducibility
5. …?

Newest ZF Selection System

Addresses some of the modularity challenges associated with monomeric ZF assembly processes.

Step 3: Applications

1. Medical therapeutics
2. Cellular engineering
3. Generating model organisms for disease
4. …?
In New Ways to Edit DNA, Hope for Treating Disease

For years, doctors have used tools like RNAi to  delete specific genes in a patient’s cells. To treat a disease, the transgenic or engineered zinc-finger nucleases (ZFNs) engineered as described (fig. S4) for a broad range of sequences. To accomplish this, ZFNs can be engineered for helpful suggestions; D. Smoller and et al. for technical assistance; and Caliper Life Sciences, for valuable tools to an increasingly technical assistance; and Caliper Life Sciences, 2007. Against a broad range of sequences, ZFNs can be engineered 

ES cell route

Progenitor cell route

Somatic nuclear transfer route


Efficient targeting of expressed and silenced genes in human ESCs and iPSCs using zinc-finger nucleases

Human hematopoietic stem cells modified by zinc-finger nucleases target to CDS context

Zinc Finger Nuclease (ZFN) Functional Genomics & RNA

Compass® Zinc Finger Nuclease Technology

ZFN technology uses a pair of zif-like proteins engineered to function as enzyme-linked transcriptional activators of the targeted gene. The nuclease then cleaves the targeted gene, insertional editing, and eventually disrupting the function of the gene. Finally, the ZFN technology provides a powerful tool for generating genetically modified animals, which can be used to study human diseases and drug development. The nuclease can be designed to target specific sequences, allowing for precise gene editing.

The high percentage of disrupted chromosomes (fig. S1) indicates high fidelity for each target sequence. The single Rab38 mutant carried 25 to 100% disruption of target chromosomes (fig. S1). Sequence analysis of the single Rab38 mutant carried 25 to 100% disruption of target chromosomes (fig. S1). Sequence analysis of the single Rab38 mutant revealed successful mutagenesis of the three gene targets after multiple delivery methods and doses in three rat strains. The laboratory rat is a well-established model for studying many human diseases.

Table of injection data revealing successful mutagenesis of the three gene targets after multiple delivery methods and doses in three rat strains. The laboratory rat is a well-established model for studying many human diseases.
TALE: A New DNA Binding Protein

Cracking the TALE code

TALEs from nature

Engineering designer TAL effectors

Understanding and engineering important parameters
The TALE crystal structure

Works in plants!

How about human cells?

ARTICLES

A TALE nuclease architecture for efficient genome editing

LETTERS

Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription

Efficient cleavage in human cells
Other Genome Engineering Technologies

- Site-specific Recombinases
- Meganucleases
- Transposase
- Viral integration
- CRISPR-Cas

from yogurt to precise genome perturbation

programmable Genome Positioning System enables precise editing and modulation of the genome
**CRISPR: an RNA-guided nuclease**

Streptococcus thermophilus LMD-9 CRISPR1

- StCas9
- St-tracrRNA
- crRNA

Streptococcus pyogenes SF370 type II CRISPR locus

- crRNA
- protospacer
- PAM

**Cas9 homologues are functional in mammalian cells**

Streptococcus thermophilus LMD-9 CRISPR1

- crRNA
- protospacer
- PAM

**What’s next? How can we bring this beyond proof of concept?**

1. Understanding and improving specificity
2. Improving efficiency
3. Exploring delivery mechanisms
4. Inducibility
5. Novel effectors
6. Actual applications!
Optimizing chimeric guide RNA architecture

Targeting capability of S. pyogenes Cas9

Single-base specificity in Cas9 guide sequence

apid generation of genetically engineered mice

apid production of transgenic animals in less than 2 years of breeding

as mediates precise gene editing in human cells

HR efficiency (%): sense antisense

<table>
<thead>
<tr>
<th>frequency (x10^7)</th>
<th>distance (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>1.25</td>
<td>10</td>
</tr>
<tr>
<td>0.75</td>
<td>20</td>
</tr>
<tr>
<td>0.18</td>
<td>30</td>
</tr>
<tr>
<td>0.12</td>
<td>40</td>
</tr>
<tr>
<td>0.02</td>
<td>50</td>
</tr>
<tr>
<td>0.01</td>
<td>60</td>
</tr>
</tbody>
</table>

Median: 8bp
Mean: 12.7bp

Targeted Mutations (Deletion / Insertion)

Predefined Precise Mutations

Streptococcus pyogenes SF370 type I CRISPR
PAM occurrence in human genome (NGG)

HR efficiency (%):

- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –

HEK | hES

<table>
<thead>
<tr>
<th>template:</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>

Single-base specificity in Cas9 guide sequence

apid generation of genetically engineered mice

apid production of transgenic animals in less than 2 years of breeding

as mediates precise gene editing in human cells

HR efficiency (%): sense antisense

<table>
<thead>
<tr>
<th>frequency (x10^7)</th>
<th>distance (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>1.25</td>
<td>10</td>
</tr>
<tr>
<td>0.75</td>
<td>20</td>
</tr>
<tr>
<td>0.18</td>
<td>30</td>
</tr>
<tr>
<td>0.12</td>
<td>40</td>
</tr>
<tr>
<td>0.02</td>
<td>50</td>
</tr>
<tr>
<td>0.01</td>
<td>60</td>
</tr>
</tbody>
</table>

Median: 8bp
Mean: 12.7bp

Targeted Mutations (Deletion / Insertion)

Predefined Precise Mutations

Streptococcus pyogenes SF370 type I CRISPR
PAM occurrence in human genome (NGG)

HR efficiency (%):

- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –

HEK | hES

<table>
<thead>
<tr>
<th>template:</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>

Single-base specificity in Cas9 guide sequence

apid generation of genetically engineered mice

apid production of transgenic animals in less than 2 years of breeding

as mediates precise gene editing in human cells

HR efficiency (%): sense antisense

<table>
<thead>
<tr>
<th>frequency (x10^7)</th>
<th>distance (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>1.25</td>
<td>10</td>
</tr>
<tr>
<td>0.75</td>
<td>20</td>
</tr>
<tr>
<td>0.18</td>
<td>30</td>
</tr>
<tr>
<td>0.12</td>
<td>40</td>
</tr>
<tr>
<td>0.02</td>
<td>50</td>
</tr>
<tr>
<td>0.01</td>
<td>60</td>
</tr>
</tbody>
</table>

Median: 8bp
Mean: 12.7bp

Targeted Mutations (Deletion / Insertion)

Predefined Precise Mutations

Streptococcus pyogenes SF370 type I CRISPR
PAM occurrence in human genome (NGG)

HR efficiency (%):

- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –

HEK | hES

<table>
<thead>
<tr>
<th>template:</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>

Single-base specificity in Cas9 guide sequence

apid generation of genetically engineered mice

apid production of transgenic animals in less than 2 years of breeding

as mediates precise gene editing in human cells

HR efficiency (%): sense antisense

<table>
<thead>
<tr>
<th>frequency (x10^7)</th>
<th>distance (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.5</td>
<td>0</td>
</tr>
<tr>
<td>1.25</td>
<td>10</td>
</tr>
<tr>
<td>0.75</td>
<td>20</td>
</tr>
<tr>
<td>0.18</td>
<td>30</td>
</tr>
<tr>
<td>0.12</td>
<td>40</td>
</tr>
<tr>
<td>0.02</td>
<td>50</td>
</tr>
<tr>
<td>0.01</td>
<td>60</td>
</tr>
</tbody>
</table>

Median: 8bp
Mean: 12.7bp

Targeted Mutations (Deletion / Insertion)

Predefined Precise Mutations

Streptococcus pyogenes SF370 type I CRISPR
PAM occurrence in human genome (NGG)

HR efficiency (%):

- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –
- Cas9 nickase –

HEK | hES

<table>
<thead>
<tr>
<th>template:</th>
<th>sense</th>
<th>antisense</th>
</tr>
</thead>
<tbody>
<tr>
<td>20</td>
<td>2.4</td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>1.0</td>
<td></td>
</tr>
<tr>
<td>2.2</td>
<td>6.0</td>
<td></td>
</tr>
</tbody>
</table>
Spacing requirements for efficient offset nicking

Optimal offset range: 0 to +30bp

Improved specificity over wildtype Cas9 DSB:

Over 1000X reduction in off-target activity

Double nicking works in mouse zygotes

Double nicking stimulates efficient HR
Modeling Autism Spectrum Disorders

Heterozygous Modification (UBE3a +/–)

Homozygous Modification (UBE3a –/–)

UBE3a +/+  UBE3a –/–  UBE3a +/–

Nanog  DAPI  Oct3/4  1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18

Modeling Autism Spectrum Disorders

http://tools.genome-engineering.org Hsu et al., 2013

Heterozygous Modification (UBE3a +/–)

Homozygous Modification (UBE3a –/–)

UBE3a +/+  UBE3a –/–  UBE3a +/–

Nanog  DAPI  Oct3/4  1  2  3  4  5  6  7  8  9  10  11  12  13  14  15  16  17  18

Modeling Autism Spectrum Disorders

http://tools.genome-engineering.org Hsu et al., 2013