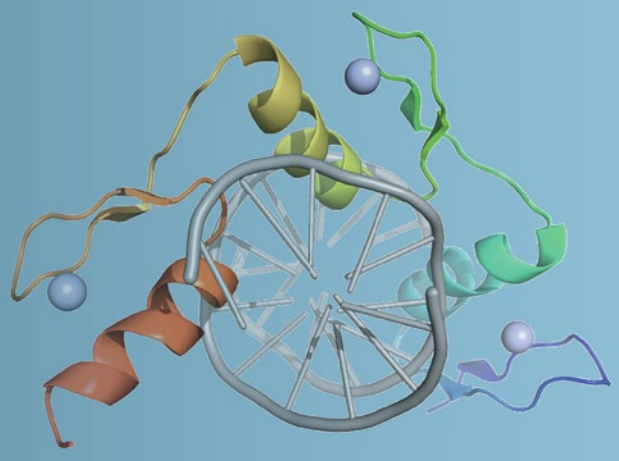
Methods in Molecular Biology 1867

Springer Protocols



Jia Liu *Editor*

Zinc Finger Proteins

Methods and Protocols

Humana Press

METHODS IN MOLECULAR BIOLOGY

Series Editor
John M. Walker
School of Life and Medical Sciences
University of Hertfordshire
Hatfield, Hertfordshire, AL10 9AB, UK

For further volumes
http://www.springer.com/series/7651

Zinc Finger Proteins

Methods and Protocols

Edited by

Jia Liu

Shanghai Institute for Advanced Immunochemical Studies, ShanghaiTech University, Shanghai, China

Humana Press

Editor
Jia Liu
Shanghai Institute for Advanced Immunochemical Studies
ShanghaiTech University
Shanghai, China

ISSN 1064-3745 ISSN 1940-6029 (electronic)
Methods in Molecular Biology
ISBN 978-1-4939-8798-6 ISBN 978-1-4939-8799-3 (eBook)
https://doi.org/10.1007/978-1-4939-8799-3

Library of Congress Control Number: 2018953313

This Humana Press imprint is published by the registered company Springer Science+Business Media, LLC part of Springer Nature.
The registered company address is: 233 Spring Street, New York, NY 10013, U.S.A.

Preface

Zinc finger proteins (ZFPs) represent the largest family of eukaryotic transcription factors (TFs). ZFPs are engaged in a wide range of biological functions including transcription, DNA methylation, cell cycling, cell division, tumorigenesis, and others. There are distinct classes of ZFPs such as C_2H_2 fingers, C_4 fingers, and CCHC fingers. Despite different sequence organization, all ZFPs are characterized by the tetrahedral coordination of zinc ions (Zn^{2+}) with cysteines and histidines. The DNA-binding preferences of ZFPs have been extensively studied for decades. ZFP can recognize DNA in a relatively simple mode. For example, each C_2H_2 finger recognizes three or four DNA base pairs via residues −1, +2, +3, and +6 on the α helix of ZFP. Elucidation of the "recognition code" enables the design of custom ZFPs for targeting specific sequences in the genome. These artificial ZFPs have proven versatile tools to manipulate gene expression that, along with the recent addition of other DNA-binding systems such as TAL effector and clustered regularly interspaced short palindromic repeats (CRISPR), open the era of genome engineering.

By fusing with different effector proteins, ZFPs can be reprogrammed to upregulate, downregulate, or knock out targeted genes. When co-delivered with a repair template, engineered ZFPs can mediate targeted nucleotide substitution or gene integration. In addition to serving as powerful research tools, ZFPs have demonstrated remarkable therapeutic potential in human trials for treating various diseases including HIV infection, hemophilia, and Hurler syndrome. Moreover, ZFPs can be engineered to bind RNA sequence or as zinc ion sensor. *Zinc Finger Proteins: Methods and Protocols* is the second volume of the *Methods in Molecular Biology* series devoted to ZFPs. Since the first ZFP volume in 2010 (eds. Joel Mackay and David Segal), notable progress has been made on understanding the biological functions of ZFPs and on the applications of ZFPs. The current book contains 19 chapters presenting the widely used methods and protocols in ZFP research.

Part I of this book presents methods and reviews for the basic biology of ZFPs and design and applications of custom ZFPs. Chapter 1 by Taniguchi et al. provides a protocol on investigating the role of ZFPs in cancer stemness. Chapter 2 by Najafabadi et al. presents a method to combine in silico prediction and ChIP-seq data to identify the in vivo DNA-binding sites of C_2H_2 ZFPs. Chapter 3 by Jeltsch et al. reviews the background, design, and application of methylation-specific ZFPs. Chapters 4 and 5 by Minczuk et al. and Vandevenne et al. describe strategies for the design and applications of mitochondrial DNA-targeting ZFPs and single-stranded RNA-targeting ZFPs, respectively. Chapter 6 by Miyamoto et al. presents a unique method for the design of artificial RING fingers as E3 enzyme for protein ubiquitination. Chapter 7 by Shannon et al. provides a method for ZFP fusions with the catalytic domain of modular recombinases for targeted DNA recombination. Chapters 8 and 9 by Qin et al. and Ma et al. streamline the utilities of ZFPs as zinc ion sensors or as protein transduction domains, respectively. Chapter 10 by Manotham et al. describes a protocol for editing patient-derived mesenchymal stem cells (MSCs) using zinc finger nucleases (ZFNs).

Part II provides methods for the evaluation and prevention of ZFN-mediated cytotoxicity. Chapter 11 by Sivalingam et al. presents a comprehensive strategy for the examination of ZFN-induced genotoxicity in primary human cells. The chapters by Ikebukuro et al. and Zhang et al. describe methods for controlled ZFN delivery using DNA aptamers against Fok I nuclease or suicidal expression system respectively for reducing the off-target effects and cytotoxicity of ZFNs.

Recent studies have highlighted the delivery methods of engineered ZFPs. *Part III* provides a collection of the representative methods of ZFN delivery. Chapters 14–16 by Hilioti et al., Minczuk et al., and Dunham et al. present the utilities of electroporation for delivery of ZFNs into plant (tomato seeds), fish (catfish sperms and embryos), and mouse (mitochondria of mouse embryos), respectively. Chapter 17 by Heller et al. presents a stereotaxic surgical approach for the viral delivery of ZFP epigenetic editors in rodent brains. The last two chapters by Segal et al. and Gaj et al. provide protocols for protein-based delivery of ZFP transcription factors and nucleases, respectively.

Tremendous progress has been made on ZFPs in the past decade, eventually leading to the transformation of ZFPs from a laboratory concept to clinical applications. The editor greatly appreciates the time and efforts of all contributing authors and sincerely hopes that the insights and methods provided in the book can prompt both the understanding of ZFP biology and the development of next-generation ZFP therapeutics.

Shanghai, China ***Jia Liu***

Contents

Contributors

Alexa N. Adams • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA*

Barbara J. Bailus • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA; The Buck Institute for Research on Aging, Novato, CA, USA*

Ray Black • *Department of Biological Sciences, University of Denver, Denver, CO, USA*

Supreecha Chattong • *Molecular and Cellular Biology Unit, Department of Medicine, Lerdsin General Hospital, Bangkok, Thailand; EST Laboratory, S.S. Manufacturing, Nonthaburi, Thailand*

Andy Cox • *Mammalian Embryo and Stem Cell Group, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK*

Simona De Franco • *Center for Protein Engineering, University of Liège, Liège, Belgium*

Berat Doğan • *Department of Human Genetics, McGill University, Montreal, QC, Canada; McGill University and Genome Quebec Innovation Centre, Montreal, QC, Canada; Department of Biomedical Engineering, İnönü University, Malatya, Turkey*

Rex A. Dunham • *School of Fisheries Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL, USA*

Ahmed Elaswad • *School of Fisheries Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL, USA; Department of Animal Wealth Development, Faculty of Veterinary Medicine, Suez Canal University, Ismailia, Egypt*

Dylan H. Fudge • *Department of Biological Sciences, University of Denver, Denver, CO, USA*

Thomas Gaj • *Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA*

Payam A. Gammage • *Mitochondrial Genetics Group, MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, UK*

Peter J. Hamilton • *The Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA*

Elizabeth A. Heller • *Department of Pharmacology and Penn Epigenetics Institute, The University of Pennsylvania, Philadelphia, PA, USA*

Zoe Hilioti • *Institute of Applied Biosciences (INAB), CERTH, Thessaloniki, Greece*

Linyong Hu • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*

Kazunori Ikebukuro • *Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo, Japan*

Kohzoh Imai • *The Institute of Medical Science, The University of Tokyo, Tokyo, Japan*

Albert Jeltsch • *Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, University Stuttgart, Stuttgart, Germany*

Yoshio Kato • *Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan*

Dimitar Kenanov • *Bioinformatics Institute, Agency for Science, Technology and Research, Singapore, Republic of Singapore*
Oi Lian Kon • *Division of Medical Sciences, Laboratory of Applied Human Genetics, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore, Republic of Singapore; Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore*
Sze Sing Lee • *Division of Medical Sciences, Laboratory of Applied Human Genetics, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore, Republic of Singapore*
Jiangmei Li • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Jie Li • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Ya-jun Liang • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Carissa J. Lim • *Department of Pharmacology and Penn Epigenetics Institute, The University of Pennsylvania, Philadelphia, PA, USA*
Jia Liu • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Peixiang Ma • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Johannes A. H. Maier • *Department of Biochemistry, Institute of Biochemistry and Technical Biochemistry, University Stuttgart, Stuttgart, Germany*
Krissanapong Manotham • *Molecular and Cellular Biology Unit, Department of Medicine, Lerdsin General Hospital, Bangkok, Thailand*
Sebastian Maurer-Stroh • *Bioinformatics Institute, Agency for Science, Technology and Research, Singapore, Republic of Singapore; Department of Biological Sciences, National University of Singapore, Singapore, Republic of Singapore*
Beverly J. McCann • *Mitochondrial Genetics Group, MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, UK*
Michal Minczuk • *Mitochondrial Genetics Group, MRC Mitochondrial Biology Unit, University of Cambridge, Cambridge, UK*
Kazuhide Miyamoto • *Faculty of Pharmaceutical Sciences, Department of Pharmaceutical Health Care, Himeji Dokkyo University, Hyogo, Japan*
Hamed S. Najafabadi • *Department of Human Genetics, McGill University, Montreal, QC, Canada; McGill University and Genome Quebec Innovation Centre, Montreal, QC, Canada*
Chikashi Nakamura • *Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo, Japan; Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan*
Eric J. Nestler • *The Friedman Brain Institute, Icahn School of Medicine at Mount Sinai, New York, NY, USA*
Wai Har Ng • *Division of Medical Sciences, Laboratory of Applied Human Genetics, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore, Republic of Singapore*
Maui Nishio • *Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo, Japan*

Mitchell R. O'Connell • *Department of Biochemistry and Biophysics, School of Medicine and Dentistry, University of Rochester, Rochester, NY, USA; Center for RNA Biology, University of Rochester, Rochester, NY, USA*
Henriette O'Geen • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA*
Toan Thang Phan • *Department of Surgery, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore; CellResearch Corporation, Singapore, Republic of Singapore*
Benjamin Pyles • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA*
Yan Qin • *Department of Biological Sciences, University of Denver, Denver, CO, USA*
Zhenkui Qin • *School of Fisheries Aquaculture and Aquatic Sciences, Auburn University, Auburn, AL, USA; Key Laboratory of Marine Genetics and Breeding (Ocean University of China), Ministry of Education, Qingdao, China*
Chonghua Ren • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA; Guangzhou Key Laboratory of Insect Development Regulation and Application Research, South China Normal University, Guangzhou, China*
Pei-ling Ren • *Shanghai Advanced Research Institute, Chinese Academy of Science, Shanghai, China; University of Chinese Academy of Sciences, Beijing, China*
Kazuki Saito • *Faculty of Pharmaceutical Sciences, Department of Pharmaceutical Health Care, Himeji Dokkyo University, Hyogo, Japan*
David J. Segal • *Genome Center and Department of Biochemistry and Molecular Medicine, University of California, Davis, CA, USA*
Shannon J. Sirk • *Department of Bioengineering, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Carl R. Woese Institute for Genomic Biology, University of Illinois at Urbana-Champaign, Urbana, IL, USA; Carle-Illinois College of Medicine, University of Illinois at Urbana-Champaign, Urbana, IL, USA*
Jaichandran Sivalingam • *Division of Medical Sciences, Laboratory of Applied Human Genetics, Humphrey Oei Institute of Cancer Research, National Cancer Centre, Singapore, Republic of Singapore; Department of Biochemistry, Yong Loo Lin School of Medicine, National University of Singapore, Singapore, Republic of Singapore; Bioprocessing Technology Institute, Agency for Science, Technology and Research, Singapore, Republic of Singapore*
James B. Stewart • *Department of Mitochondrial Biology, Max Planck Institute for Biology of Ageing, Cologne, Germany*
Hiroaki Taniguchi • *The Center for Antibody and Vaccine Therapy, Research Hospital, The Institute of Medical Science, The University of Tokyo, Tokyo, Japan*
Kaori Tsukakoshi • *Department of Biotechnology and Life Science, Tokyo University of Agriculture and Technology, Tokyo, Japan*
Marylène Vandevenne • *Center for Protein Engineering, University of Liège, Liege, Belgium*
Ling Wang • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Shaojie Wang • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*

Xuan Wei • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*
Jiani Xing • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Kun Xu • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Ayana Yamagishi • *Biomedical Research Institute, National Institute of Advanced Industrial Science and Technology (AIST), Ibaraki, Japan*
Magdalena Zernicka-Goetz • *Mammalian Embryo and Stem Cell Group, Department of Physiology, Development and Neuroscience, University of Cambridge, Cambridge, UK*
Cunfang Zhang • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Tingting Zhang • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Zhiying Zhang • *College of Animal Science and Technology, Northwest A&F University, Yangling, Shaanxi, China*
Lianhui Zhu • *Shanghai Institute for Advanced Immunochemical Studies (SIAIS), ShanghaiTech University, Shanghai, China*

Part I

Basic Biology and Engineering of ZFPs

Chapter 1

PRDM14, a Zinc Finger Protein, Regulates Cancer Stemness

Hiroaki Taniguchi and Kohzoh Imai

Abstract

PRDI-BF1 and RIZ homology (PR) domain zinc finger protein 14 (PRDM14) contains a PR domain related to the SET methyltransferase domain and zinc finger motifs. PRDM14 maintains stemness in embryonic stem cells and primordial germ cells via epigenetic mechanisms. PRDM14, however, is not expressed in normal differentiated tissues. We and other groups previously reported that PRDM14 expression is markedly higher in some types of cancers compared to the corresponding normal tissues. PRDM14 confers stem cell-like characteristics upon cancer cells, such as sphere formation, dye efflux, chemotherapy resistance, proliferation, and distant metastasis. Cancer stem cells (CSCs) are thought to be responsible for tumor initiation, drug and radiation resistance, invasive growth, metastasis, and tumor relapse, which are the primary causes of cancer-related deaths. Because CSCs are also thought to be resistant to conventional therapies, an effective and novel therapeutic approach for CSCs is imperative.

RNAi silencing of *PRDM14* expressed by breast and pancreatic cancer cells reduced tumor size and distant metastasis of these cells in nude mice. Inhibition of *PRDM14* expression by cancer cells may be an effective and radical therapy for solid cancers. In this chapter, we discuss methods for studying CSC-like properties in cancer cells and describe the use of siRNA with a drug delivery system by systemic injection in vivo.

Key words PRDM14, Zinc finger protein, Cancer stemness, Breast cancer, siRNA therapy

1 Introduction

PRDM14 is a member of the PRDI-BF1 and RIZ homology (PR) domain containing (PRDM) family of transcription regulators with six Cys2His2 (C2H2)-type zinc finger (Znf) domains. The PR domain is related to the SET methyltransferase domain, but no histone methyltransferase activity has been reported [1]. C2H2 Znfs are the most common DNA-binding motifs found in transcription factors [2] and can bind to RNA and protein targets [3]. PRDM14 binds to the consensus sequence 5′-GGTC[TC]CTAA-3′ [4].

PRDM14 is specifically expressed in embryonic stem (ES) cells and primordial germ cells and promotes ES cell pluripotency [5,

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_1,

6]. PRDM14 directly upregulates the expression of the transcription factor OCT4 and localizes with other master regulators of pluripotency in ES cells [4]. PRDM14 was markedly expressed in many different types of cancer such as breast cancer, non-small cell lung carcinoma, testicular cancer, esophageal cancer, pancreatic cancer, ovarian cancer, renal cancer, and lymphoma [7–10]. Given the high level of PRDM14 expression in tumors and the ability of PRDM14 to mediate pluripotency in ES cells, we hypothesized that PRDM14 contributes to acquisition of the cancer stem cell (CSC) phenotype.

CSCs are thought to be responsible for tumor initiation, drug and radiation resistance, invasive growth, metastasis, and tumor relapse, which are the main causes of cancer-related deaths [11]. In solid tumors, the presence of CSCs was first demonstrated in the $CD44^{+}CD24^{-/low}$ fraction of breast cancer cells [12]. CSCs express unique surface markers, exist in a side-population (SP) fraction possessing increased Hoechst-33342 efflux capacity, exhibit high aldehyde dehydrogenase-1 (ALDH1) activity, and form spheres when cultured under non-adherent conditions [13]. CSCs also show high tumorigenicity on xenografting into immunocompromised mice [12, 13].

We previously reported that PRDM14 promotes tumor growth and metastasis in vivo and confers resistance to anticancer drugs and apoptosis, all of which characterize the CSC phenotype [9, 10]. Using an in vivo therapeutic model with the PRDM14 chimera siRNA [14] in combination with a novel drug delivery system (DDS)-calcium phosphate (CaP) hybrid micelles [15] or poly(ethylene glycol)-poly(L-lysine) [16], the growth of established tumors and distant metastasis can be inhibited by downregulating PRDM14 expression [9, 10]. This chapter reviews available methods for analyzing stem cell-like properties in cancer cells and the use of an innovative siRNA system against grafted tumor cells by systemic injection in vivo.

2 Materials

2.1 Tumorsphere Assay and Primary Tumor Cell Culture

1. Tumor cell lines obtained from the American Type Culture Collection.
2. An adequate culture medium containing 10% fetal bovine serum (FBS).
3. Phosphate-buffered saline (PBS).
4. Antibiotic antimycotic solution (100×), stabilized.
5. Liberase™ TL research grade.
6. Costar® 6-well clear flat-bottom ultralow attachment multiple well plates, individually wrapped, sterile.

7. DMEM/F12 medium.
8. Human epidermal growth factor and basic fibroblast growth factor.
9. B27 supplement.
10. PBS containing 4% paraformaldehyde.
11. Bovine serum albumin (BSA).
12. Triton X-100.
13. Normal donkey serum.
14. Fluorescent-labeled antibodies.

2.2 Analysis of Cancer Cell SP Fractions and Surface Markers

1. Dulbecco's modified Eagle's medium (DMEM).
2. Accutase.
3. Reserpine, verapamil, or fumitremorgin C.
4. Hoechst 33342 dye.
5. PBS and FBS.
6. Propidium iodide.
7. 35 mm Cell strainer.

2.3 Analysis of Aldehyde Dehydrogenase-1 (ALDH1) Activity

1. ALDEFLUOR™ Kit (#01700; StemCell Technologies) including Dry ALDEFLUOR reagent, diethylaminobenzaldehyde (DEAB) 1.5 mM in 95% ethanol, 2 N HCl, dimethylsulfoxide (DMSO), and ALDEFLUOR Assay Buffer.
2. Propidium iodide.

2.4 siRNA Therapeutic Model for Orthotopic Graft and Lung Metastasis of Breast Cancer Cells

1. Isoflurane.
2. Female, 6-week-old BALB/c-nu mice.
3. BD matrigel matrix growth factor reduced (BD Biosciences), stored as single-use aliquots at −20 °C, and thawed on ice.
4. PBS.
5. Custom siRNA or small interfering RNA/DNA chimera (so-called chimera siRNA).
6. In vivo-jetPEI (polyethyleneimine; Polyplus Transfection) including 5% glucose solution.
7. Doxorubicin hydrochloride, dissolved in sterile 0.9% saline, and sterilized by filtration.
8. Docetaxel, diluted in PBS/DMSO (ratio 1:1) and sterilized by filtration.

3 Methods

3.1 Tumorsphere Assay and Primary Tumor Cell Culture

Sphere formation assays are conducted to isolate cell subsets enriched with CSCs present in solid tumors and assess their self-renewal ability [17]. We applied the same method in a tumorsphere assay of cancer cell lines and primary culture using clinical tumor tissues.

1. Culture tumor cell lines obtained from the American Type Culture Collection in an adequate culture medium containing 10% FBS at 5% CO_2 at 37 °C.
2. For clinical tumor tissues, wash tumor samples three times with PBS supplemented with 1× antibiotics and an antimycotic agent and homogenize them into small fragments, using Liberase TL research grade.
3. Culture prepared tumor cells in ultralow attachment 6-well plates.
4. Add 20 ng/mL epidermal growth factor and 20 ng/mL basic fibroblast growth factor to 50 mL of DMEM/F12 medium.
5. After adding 0.5 mL of B27 supplement, culture the cells at 37 °C in 5% CO_2.
6. Culture the tumor cells for 2 weeks, after which is evaluated spheroid formation based on the number and diameter of the spheres. Tumorspheres are sometimes fused with each other, and therefore separating them is important (*see* **Note 1**).
7. (Optional) Fix spheres for 20 min at room temperature in PBS containing 4% paraformaldehyde and then wash them three times with 1% BSA in PBS for 5 min.
8. (Optional) Permeabilize and block spheres with 1% Triton X-100, 1% BSA, and 10% normal donkey serum in PBS at room temperature for 60–120 min (*see* **Note 2**).
9. (Optional) After blocking, incubate the spheres with 10 μg/mL fluorescent-labeled antibodies in the dark overnight at 4 °C.
10. (Optional) After washing the spheres three times with 1% BSA in PBS for 5 min, visualize the spheres under a confocal microscope.

3.2 Analysis of Cancer Cell SP Fractions and Surface Markers

The SP phenotype is characteristic of CSCs and associated with drug resistance [18]. SP cells can strongly efflux Hoechst 33342 dye because of the action of ATP-binding cassette (ABC) transporters [18, 19] (Fig. 1).

1. Culture tumor cells at 37 °C in an adequate culture medium supplemented with 5% FBS.

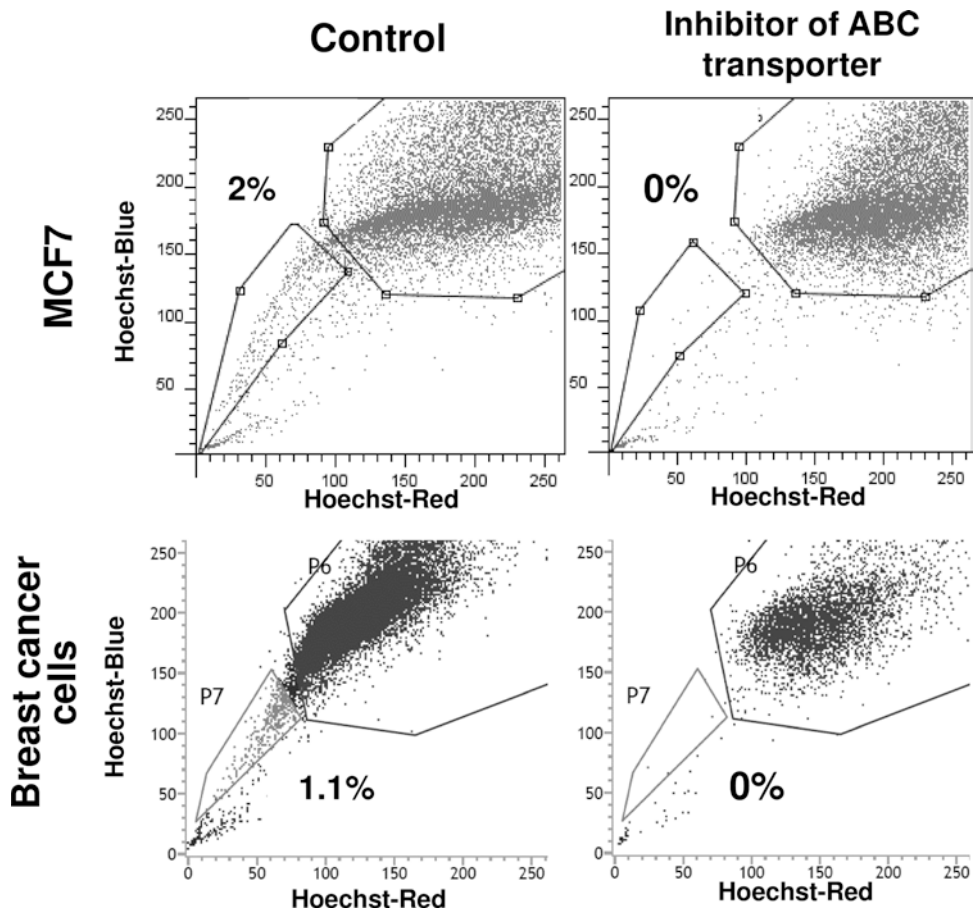


Fig. 1 Side population of breast cancer cells. The side-population fractions of breast cancer cell lines (MCF7) [11] and human breast cancer cell lines derived from poorly differentiated adenocarcinoma. Reserpine, which blocks ABC transporters, was added as a negative control

2. Suspend these cells in 4 mL of 5% FBS-DMEM (1×10^6 cells/mL), after removing the cells from the dishes using Accutase and neutralize them with culture medium.
3. Add 20 μM reserpine to 1 mL aliquots of the cell suspension. Then, incubate cells for 10 min at 37 °C. The ABC transporter can be specifically inhibited by drugs such as reserpine, verapamil, and fumitremorgin C [20] (*see* **Note 3**).
4. Add 5 μM Hoechst 33342 dye to both cell populations.
5. Incubate tumor cells with shaking for 90 min at 37 °C, immediately cool on ice, centrifuge for 5 min at 300 × *g* at 4 °C, and resuspend them in cooled 1× PBS-5% FBS solution. Repeat this procedure three times (*see* **Note 4**).
6. (Optional) Suspend tumor cells in 50–100 μL of cooled 1× PBS-5% FBS solution.

7. (Optional) Add 0.1–10 μg/mL antibodies to the cell suspension according to datasheets, and incubate in the dark for 30–60 min on ice. Isolate CSCs by detecting unique surface markers such as CD24, CD26, CD44, CD90, CD133, and CD166 (Table 1) (*see* **Note 5**) [11].
8. (Optional) Wash tumor cells three times in cooled 1× PBS-5% FBS solution.
9. Add 2 μg/mL propidium iodide for discrimination of dead cells.

Table 1
Representative unique markers of CSCs

Tumor type	Representative unique markers
Leukemia	CD34+/CD38−
Breast cancer	CD44+/CD24−/lin−/ALDH1+
Colorectal cancer	CD133+/CD44+/ALDH1+
	EpCAM+/CD44+, CD166+
	CD44+/CD24+
	Lgr5/GPR49
Metastatic colon	CD133+/CD26+
Gastric cancer	CD44+
Head and neck cancer	CD44+
Liver cancer	CD133+/CD49f+
	CD90+/CD45-
	CD13+
	EpCAM+
Pancreatic cancer	CD133+/CD44+/CD24+/ESA+
	CXCR4+
Esophageal cancer	CD44+/ALDH1+
Lung cancer	CD133+/ABCG2high
Brain cancer	CD133+/BCRP1+/A2B5+/SSEA-1+
Multiple myeloma	CD138−
	CD19+
Prostate cancer	CD133+/CD44+/α2β1high
Melanoma	CD20+
	ABCB5+

10. Filter the samples using a 35 mm cell strainer to eliminate aggregated cells.
11. Excite Hoechst 33342 dye by UV irradiation and measure fluorescence using both 675/20 (Hoechst Red) and 424/44 (Hoechst Blue) filters by FACSAria (BD Biosciences, San Jose, CA, USA).
12. Analyze data using FlowJo software (FlowJo LLC, Ashland, OR, USA).

3.3 Analysis of Aldehyde Dehydrogenase-1 (ALDH1) Activity

CSCs exhibit high aldehyde dehydrogenase-1 (ALDH1) activity [13]. We performed an ALDEFLUOR flow cytometry-based assay to measure the ALDH1 activity of cancer cells.

1. Dissolve dry ALDEFLUOR™ reagent in 25 μL of DMSO for 1 min at room temperature. This is converted into the fluorescent-activated ALDEFLUOR™ reagent (containing the ALDH substrate BODIPY-aminoacetaldehyde (BAAA)) by treatment with 25 μL of 2 N HCl for 15 min. Dilute this reagent further with 360 μL of ALDEFLUOR assay buffer (*see* **Note 6**).
2. Suspend tumor cells in 1 mL of ALDEFLUOR assay buffer (1×10^6 cells/mL).
3. Add approximately 5 μL of activated ALDEFLUOR reagent to 1 mL of the cell suspension and mix well (*see* **Note 7**).
4. Immediately, add approximately 0.5 mL of the cell suspension to the tube containing 5 μL of the specific ALDH inhibitor, DEAB (1.5 mM in 95% ethanol).
5. Incubate the test and control samples for 30–45 min in a water bath at 37 °C.
6. Centrifuge the tubes at 4 °C for 5 min at 300 × *g*.
7. Aspirate the supernatant without disturbing the pellet, which is resuspended in 0.5 mL of ice-cold ALDEFLUOR® assay buffer; place the samples immediately on ice.
8. Add 2 μg/mL propidium iodide to detect dead cells.
9. Detect brightly fluorescent ALDH1-expressing cells ($ALDH1^{high}$) using the green fluorescence channel (520–540 nm) of FACSAria (BD Biosciences).
10. Analyze data using FlowJo software (FlowJo LLC).

3.4 siRNA Therapeutic Model for Orthotopic Graft and Lung Metastasis of Breast Cancer Cells

Significant barriers limit the use of in vivo therapy models with siRNA. To overcome these barriers, it is important to select target genes with limited expression in cancer cells and select the sequence of siRNA while considering off-target effects and DDSs that accumulates at tumor sites through the enhanced permeability and retention effect because of their narrow and uniform diameter. The

"siDirect" program (http://sidirect2.rnai.jp/) can be applied to select adequate siRNA sequences. We examined the effects of PRDM14-siRNA administration with several DDSs on breast or pancreatic cancer cell growth and lung or liver metastasis in vivo [9, 10].

3.4.1 Therapeutic Model for Orthotopically Grafted Breast Cancer Cells

1. Suspend breast tumor cells (1×10^6 cells/mouse) in 100 μL of PBS and mix with 100 μL of Matrigel on ice (*see* **Note 8**).
2. Orthotopically, inoculate the mixture into the mammary fat pad of female nude mice at the age of 6 weeks.
3. After tumor formation, measure the length and width of the tumor using a caliper. Tumor volume can be calculated using the formula: Tumor volume = length × $(\text{width})^2 \times 3.14/6$.
4. When the tumor volume exceeds 100 mm^3, initiate treatment with siRNA and/or anticancer drug. Inject siRNA and in vivo-jetPEI (polyethyleneimine, PEI) directly into the tumors three times per week. siRNA (20 μg/mouse) is complexed with in vivo-jetPEI (2.8 μL/mouse) at an N/P ratio of 8, mixed with 40 μL of 5% glucose. Administer doxorubicin (1 mg/kg) or docetaxel (10 mg/kg) intraperitoneally once per week.
5. Resect tumors immediately after the animals are sacrificed by CO_2 inhalation. After tumor tissues are formalin-fixed and embedded in paraffin, stain the slices with hematoxylin and eosin (H&E) and incubate them overnight at 4 °C with antibodies for immunohistological analyses.

3.4.2 Therapeutic Model for Lung Metastasis of Breast Cancer Cells

1. To create a mouse model for the treatment of pulmonary metastasis, inject MDA-MB-231 cells (1×10^6 cells/mouse) prepared in PBS into nude mice via the tail vein while under anesthesia (isoflurane).
2. Initiate treatment with siRNA (40 μg/mouse) and in vivo-jetPEI (8.0 μL/mouse) with 250 μL of 5% glucose 48 h after tumor cell challenge; administrate this complex into the tail vein or the orbital venous plexus under anesthesia two or three times per week.
3. Six weeks later, the animals are sacrificed by CO_2 inhalation and assessed for lung weight and pulmonary metastasis.
4. After the tumor tissues are fixed in formalin and embedded in paraffin, stain the slices with H&E and incubated overnight at 4 °C with antibodies for immunohistological analyses.

Both protocols involve the use of PEI as the DDS; however, PEI has relatively high cytotoxicity. Please refer to the publication discussing an in vivo treatment model using CaP hybrid micelles [15] or polyion complexes [16] with very low toxicity in vivo for intravenous delivery of chimera siRNA to inhibit PRDM14 expression [9, 10] (Fig. 2) (*see* **Note 9**).

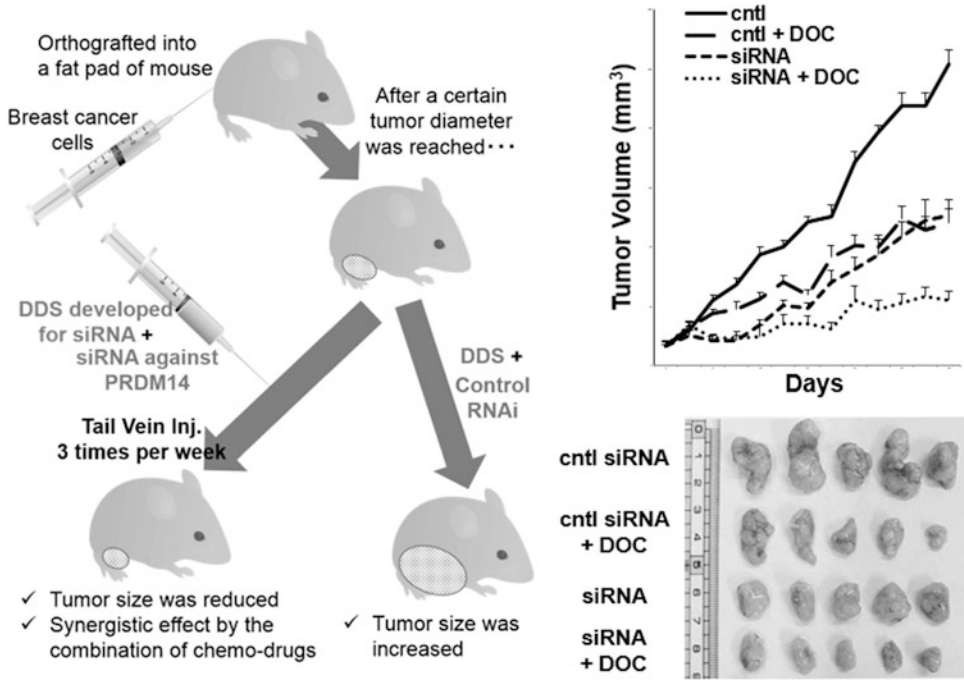


Fig. 2 PRDM14 siRNA reduces the tumorigenicity of breast cancer cells in vivo. Intravenous injection of PRDM14 siRNA/CaP hybrid micelles with or without DOC reduced the tumorigenicity of PRDM14$^+$ breast cancer cells [11]

4 Notes

1. Tumorspheres are sometimes fused with each other. In this case, please consider adding 1% methylcellulose to the medium.
2. 0.1–0.5% (v/v, in PBS) Triton X-100 is a commonly used detergent for permeabilizing cells. However, it is difficult to permeabilize the spheres with 0.1–0.5% detergents; therefore, spheres must be permeabilized with 1% Triton X-100.
3. The side-population fraction is altered by ABC transporter activity in cancer cells and Hoechst 33342 dye concentration. Please test several concentrations of ABC transporter inhibitors. First, try to add 50 μM reserpine, 50 μM verapamil, or 10 μM fumitremorgin C to the medium.
4. It is very important that the stained cells are always kept on ice and protected from light.
5. Many markers for CSCs are also expressed on normal stem cells (Table 2) [11]. Therefore, the best way to separate CSCs from the other fraction of cells is to identify the molecules responsible for the unique properties of CSCs, but not normal cells, such as Dclk1 in colorectal CSCs and PRDM14 in breast CSCs.

Table 2
Markers for both normal and cancer stem cells

Marker	Cell types	
	Normal stem or progenitor cells	Cancer stem cells
ALDH1	AdSC (breast)	Many types of carcinoma
Bmi-1	HSC, NSC, AdSC (intestine, breast, prostate)	Many types of carcinoma, neuroblastoma, leukemia
CD29	AdSC (breast)	Breast cancer, colorectal cancer
CD34	HSC, MSC, HProgC, EnProgC	Leukemia, sarcomas
CD44	HSC, HProgC, PSC	Many types of carcinoma
CD90	ProgC (thymus), MSC	Breast cancer, glioblastoma, liver cancer
CD117	ProgC	Breast cancer, ovarian cancer, lung cancer, glioblastoma
CD133	HSC, NSC, AdSC (colon)	Many types of carcinoma, glioblastoma, melanoma
CDw338	ESC, HSC, AdSC	Breast cancer, lung cancer, glioblastoma, melanoma
Nestin	NSC, ProgC (brain), HProgC	Liver cancer, glioblastoma, melanoma
Oct4	ESC, iPSC	Many types of carcinoma

Abbreviations: *AdSC* adult stem cell, *EnProgC* endothelial progenitor cell, *ESC* embryonic stem cell, *HProgC* hematopoietic progenitor cell, *HSC* hematopoietic stem cell, *ProgC* progenitor cell, *PSC* pluripotent stem cell, and *iPSC* induced pluripotent stem cell

6. Aliquot the remaining activated ALDEFLUOR™ reagent and store at −20 °C for a maximum of 1 year.
7. BODIPY aminoacetaldehyde (BAAA), a fluorescent nontoxic substrate for ALDH, enters viable cells. Intracellular ALDH converts BAAA into negatively charged BODIPY-aminoacetate (BAA), leading to increased fluorescence depending on ALDH activity.
8. Dilute Matrigel with PBS containing cancer cells (1:1) and keep this solution on ice until injection into the mouse.
9. Mix 1 mg/kg chimera siRNA (RNAi, Inc., Tokyo, Japan) with calcium phosphate (CaP) hybrid micelles (kindly provided by Dr. Kazunori Kataoka) according to the reference [15] and injected into the tail vein three times each week. Administer 2.5 mg/kg docetaxel intraperitoneally once weekly.

Acknowledgments

This work was supported by the Department of Research Promotion, Practical Research for Innovative Cancer Control, Ministry of Health, Labour and Welfare and Japan Agency for Medical Research and Development (AMED). We wish to thank Prof. Kazunori Kataoka from the Innovation Center of NanoMedicine for DDSs.

References

1. Mzoughi S, Tan YX, Low D et al (2016) The role of PRDMs in cancer: one family, two sides. Curr Opin Genet Dev 36:83–91
2. Bouhouche N, Syvanen M, Kado CI (2000) The origin of prokaryotic C2H2 zinc finger regulators. Trends Microbiol 8:77–81
3. Brayer KJ, Segal DJ (2008) Keep your fingers off my DNA: protein-protein interactions mediated by C2H2 zinc finger domains. Cell Biochem Biophys 50:111–131
4. Chia NY, Chan YS, Feng B et al (2010) A genome-wide RNAi screen reveals determinants of human embryonic stem cell identity. Nature 468:316–320
5. Yamaji M et al (2008) Critical function of Prdm14 for the establishment of the germ cell lineage in mice. Nat Genet 40:1016–1022
6. Yamaji M, Seki Y, Kurimoto K et al (2013) PRDM14 ensures naive pluripotency through dual regulation of signaling and epigenetic pathways in mouse embryonic stem cells. Cell Stem Cell 12:368–382
7. Nishikawa N, Toyota M, Suzuki H et al (2007) Gene amplification and overexpression of PRDM14 in breast cancers. Cancer Res 67:9649–9657
8. Dettman EJ, Justice MJ (2008) The zinc finger SET domain gene Prdm14 is overexpressed in lymphoblastic lymphomas with retroviral insertions at Evi32. PLoS One 3:e3823
9. Taniguchi H, Hoshino D, Moriya C et al (2017) Silencing PRDM14 expression by an innovative RNAi therapy inhibits stemness, tumorigenicity, and metastasis of breast cancer. Oncotarget 8:46856–46874
10. Moriya C, Taniguchi H, Miyata K et al (2017) Inhibition of PRDM14 expression in pancreatic cancer suppresses cancer stem-like properties and liver metastasis in mice. Carcinogenesis 38:638–648
11. Taniguchi H, Moriya C, Igarashi H et al (2016) Cancer stem cells in human gastrointestinal cancer. Cancer Sci 107:1556–1562
12. Al-Hajj M, Wicha MS, Benito-Hernandez A et al (2003) Prospective identification of tumorigenic breast cancer cells. Proc Natl Acad Sci U S A 100:3983–3988
13. Visvader JE, Lindeman GJ (2008) Cancer stem cells in solid tumours: accumulating evidence and unresolved questions. Nat Rev Cancer 8:755–768
14. Ui-Tei K, Naito Y, Zenno S et al (2008) Functional dissection of siRNA sequence by systematic DNA substitution: modified siRNA with a DNA seed arm is a powerful tool for mammalian gene silencing with significantly reduced off-target effect. Nucleic Acids Res 36:2136–2151
15. Pittella F, Cabral H, Maeda Y et al (2014) Systemic siRNA delivery to a spontaneous pancreatic tumor model in transgenic mice by PEGylated calcium phosphate hybrid micelles. J Control Release 178:18–24
16. Shimizu H, Hori Y, Kaname S et al (2010) siRNA-based therapy ameliorates glomerulonephritis. J Am Soc Nephrol 21:622–633
17. Visvader JE, Lindeman GJ (2012) Cancer stem cells: current status and evolving complexities. Cell Stem Cell 10:717–728
18. Dean M, Fojo T, Bates S (2005) Tumour stem cells and drug resistance. Nat Rev Cancer 5:275–284
19. Zhou S, Schuetz JD, Bunting KD et al (2001) The ABC transporter Bcrp1/ABCG2 is expressed in a wide variety of stem cells and is a molecular determinant of the side-population phenotype. Nat Med 7:1028–1034
20. Doyle L, Ross DD (2003) Multidrug resistance mediated by the breast cancer resistance protein BCRP (ABCG2). Oncogene 22:7340–7358

Chapter 2

Computational Methods for Analysis of the DNA-Binding Preferences of Cys2His2 Zinc-Finger Proteins

Berat Doğan and Hamed S. Najafabadi

Abstract

Cys_2His_2 zinc-finger proteins (C2H2-ZFPs) constitute the largest class of human transcription factors (TFs) and also the least characterized one. Determining the DNA sequence preferences of C2H2-ZFPs is an important first step toward elucidating their roles in transcriptional regulation. Among the most promising approaches for obtaining the sequence preferences of C2H2-ZFPs are those that combine machine-learning predictions with in vivo binding maps of these proteins. Here, we provide a protocol and guidelines for predicting the DNA-binding preferences of C2H2-ZFPs from their amino acid sequences using a machine learning-based recognition code. This protocol also describes the tools and steps to combine these predictions with ChIP-seq data to remove inaccuracies, identify the zinc-finger domains within each C2H2-ZFP that engage with DNA in vivo, and pinpoint the genomic binding sites of the C2H2-ZFPs.

Key words Cys_2His_2 zinc-finger proteins, Transcription factors, DNA motifs, ChIP-seq, Machine learning, Recognition code

1 Introduction

Cys_2His_2 zinc-finger proteins (C2H2-ZFPs) are by far the largest class of human transcription factors (TFs), constituting roughly ~45% of all human TFs (~750 out of ~1700 TFs) [1, 2]. This class includes some of the most widely studied TFs with integral roles in gene regulation and genome organization. Nonetheless, C2H2-ZFPs overall represent the least characterized group of TFs—only ~20% of them have known sequence preferences, and even fewer have known roles in cell function and regulation [3].

Most C2H2-ZFPs contain an array of multiple zinc fingers (ZFs) (Fig. 1)—human C2H2-ZFPs on average contain ~10 ZFs per protein [4]. Each ZF domain adopts a single ββα fold and most often has the amino acid sequence motif "X_2-Cys-$X_{2–4}$-Cys-X_{12}-His-$X_{3–5}$-His", where X represents one of the 20 natural amino

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_2,

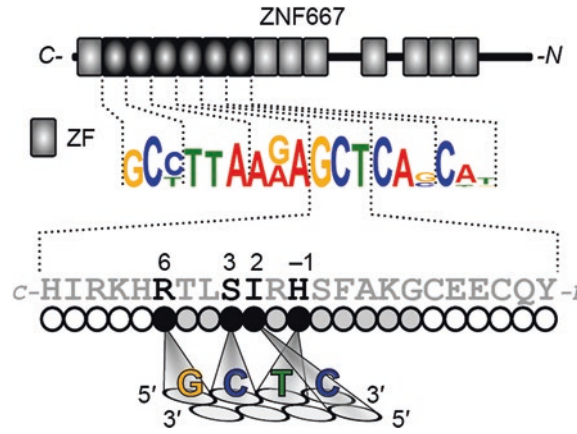


Fig. 1 DNA recognition by C2H2-ZFPs. The DNA-binding preference of ZNF667 is shown as an example—the ZFs that bind to this motif are shown in black. Each ZF recognizes three to four bases, mainly through four "specificity" residues (i.e., the numbered residues at the bottom part of the panel)

acids [5]. Each ZF typically recognizes three to four nucleotides [6], with DNA binding often limited to a subset of the fingers in a multi-ZF array. The DNA base-contacting residues of each ZF are most commonly defined as four canonical "specificity residues" in positions −1, +2, +3, and +6 relative to the start of the α-helix [7] (Fig. 1), although other residues may also contribute to sequence specificity [4].

The molecular principles that dictate the relationship between the amino acid sequences of TFs and their preferences for specific DNA sequences have been studied for decades [8]. For the C2H2-ZF class of TFs, the earlier studies mostly focused on the 3D structure of the mouse Egr1 protein [6] and a limited number of mutation analyses [9] to derive models of DNA recognition by ZFs, and understand how different amino acids in different ZF positions dictate the DNA-binding preference. Extensive in vitro binding data from thousands of ZFs has enabled recent studies to derive more complex "recognition codes" by correlating the amino acid sequence of the ZFs with their binding preferences [4, 10–12]. Despite their limitations [4], these machine-learning-based recognition models have enabled discovery of integral and unexpected roles for C2H2-ZFPs [13] and substantial new insights into their evolution [14, 15].

Importantly, these recognition codes have provided new means to analyze the in vivo binding maps of C2H2-ZFPs, in order to extract accurate models for the DNA sequence specificity of these proteins. A recent study [16] has introduced an algorithm called RCADE that combines predictions from a

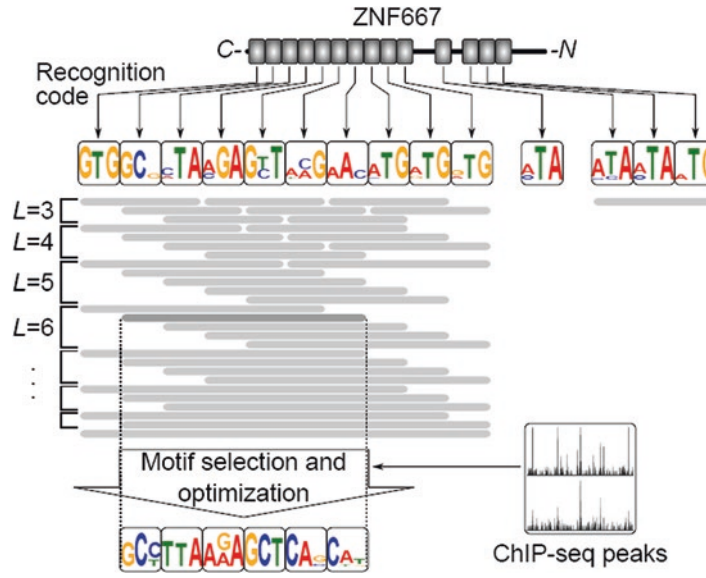


Fig. 2 Schematic illustration of the process of prediction and optimization of the motifs by RCADE. The individual three-nucleotide motifs for the ZFs are predicted using the recognition code, and then the best combination of the ZFs is selected and optimized based on enrichment in the ChIP-seq peaks

machine-learning-based recognition code [17] with the in vivo maps of the genomic binding sites of C2H2-ZFPs (Fig. 2), in order to (a) identify "motifs" that accurately represent the DNA-binding preferences of full-length C2H2-ZFPs, and (b) identify the ZFs in the multi-ZF array that engage with the DNA. RCADE starts by predicting the DNA-binding preferences of all possible combinations of adjacent ZFs using the recognition code. It then identifies the combinations that are enriched in the ChIP-seq peaks, and uses the peak sequences to optimize the predicted motif for each combination. Finally, it sorts the optimized motifs based on their ability to distinguish real peak sequences from their dinucleotide-shuffled counterparts—the top-scoring motif corresponds to the ZF domains that are most likely to engage in DNA binding (Fig. 2).

This chapter aims to provide a guide for understanding the basic steps necessary to use the C2H2-ZF recognition code, analyze ChIP-seq datasets using RCADE, and find the DNA-binding motifs of C2H2-ZFPs as well as the ZF domains within each C2H2-ZFP that engage with those motifs. The protocol has been implemented as a semiautomated workflow in open-source software, with each step documented in detail to provide reproducible and easily interpretable analyses, as outlined below.

2 Materials

2.1 Software

1. Unix-compatible operating system: This protocol has been tested on Linux and macOS. While we have not tested our pipeline in Windows, in principle it should be possible to perform these analyses on Windows machines that support a Unix-like environment, such as Cygwin (https://www.cygwin.com/).
2. In this protocol, commands that should be entered in the Linux console or the macOS Terminal are denoted with a "$" sign at the beginning of the line. This "$" sign, which separates the command prompt from what you type in the console, is not part of the command itself, and should not be typed. In contrast, the "$" signs that appear in the middle of the lines are in fact part of the command. In the following example, the part that should be typed is underlined:

   ```
   $ mkdir $HOME/tools
   ```

3. Python: Some of the tools used in this protocol rely on Python scripts. The latest version of Python and installation instructions can be found at https://www.python.org/. Most Linux distributions come with a preinstalled version of Python.
4. R and the required libraries: The core components of this protocol use the statistical computing language R. The latest version of R along with installation instructions can be found at https://www.r-project.org/.
   (a) Download and install R from a suitable mirror: https://cran.r-project.org/mirrors.html. You can download a precompiled binary distribution.
   (b) Install the "randomForest" library, which is required for this protocol. To do so, in the Linux console (or Terminal if you are using macOS), execute the command:

   ```
   $ R
   ```

   This will take you to the R environment, where you can install packages. Type and execute the command:

   ```
   install.packages("randomForest")
   ```

   Once the package is installed, type and execute the following command to exit the R environment:

   ```
   quit()
   ```

5. BEDTools: The BEDTools [18] is a software suite for handling BED files, which are files that contain genomic coordinates of a feature of interest, e.g., the in vivo binding sites of a TF. The most up-to-date instructions for installing BEDTools

can be found at http://bedtools.readthedocs.io/. For the current version (v 2.25.0):

(a) Create a folder where the tools required for this protocol will be installed. For example:

```
$ mkdir $HOME/tools
```

(b) Go to the tools folder:

```
$ cd $HOME/tools
```

(c) Download the source package (*see* **Note 1**):

```
$ wget --no-check-certificate
https://github.com/arq5x/bedtools2/releases/
download/v2.25.0/bedtools-2.25.0.tar.gz
```

(d) Extract the package:

```
$ tar -zxvf bedtools-2.25.0.tar.gz
```

(e) Go to the extracted folder:

```
$ cd bedtools2
```

(f) Compile the source code:

```
$ make
```

(g) Copy the executable files from the "`./bin`" folder to "`/usr/local/bin`", or another folder that is included in the "PATH" environmental variable.

6. MEME suite: The MEME suite [19] has several tools that are used in this protocol, including tools for manipulating FASTA files (such as extracting the central regions of a set of sequences, or creating dinucleotide-shuffled sequences). The latest version of MEME suite can be found at http://meme-suite.org/doc/download.html. For the current version (v 4.12.0):

(a) Go to the tools folder that you have created:

```
$ cd $HOME/tools
```

(b) Download the source package:

```
$ wget http://meme-suite.org/meme-software/
4.12.0/meme_4.12.0.tar.gz
```

(c) Extract:

```
$ tar zxf meme_4.12.0.tar.gz
```

(d) Go to the extracted folder:

```
$ cd meme_4.12.0
```

(e) Configure the settings for installation:

```
$ ./configure --prefix=$HOME/tools/meme --with-
url=http://meme-suite.org --enable-build-
libxml2 --enable-build-libxslt
```

(f) Compile the code:

```
$ make
```

(g) Test the compilation:

```
$ make test
```

(h) Install the binaries:

```
$ make install
```

7. RCADE: RCADE [16] is the core algorithm that analyzes the peak sequences in order to identify C2H2-ZFP motifs. The most up-to-date version of RCADE along with installation instructions can be found at https://github.com/hsnajafabadi/RCADE:

(a) Go to the tools folder that you have created:

```
$ cd $HOME/tools
```

(b) Download the package using Git (https://git-scm.com/):

```
$ git clone https://github.com/hsnajafabadi/RCADE
```

(c) Go to the downloaded folder:

```
$ cd RCADE
```

(d) Compile the code:

```
$ make
```

(e) Change the value of line 7 of the "`RCOpt.sh`" script to the path that contains the executable files for MEME suite on your computer. For example, if you installed MEME in "`$HOME/tools/meme`", then line 7 of "`RCOpt.sh`" should be:

```
memebin=$HOME/tools/meme/bin
```

2.2 Data Files

1. The C2H2-ZF protein sequence: The recognition code requires the protein sequence of the C2H2-ZFP (in FASTA format) in order to predict the binding preference of each individual ZF. The `examples` folder in the RCADE directory contains protein FASTA files for a few C2H2-ZFPs (*see* **Note 2**). We will use CTCF as an example in this protocol:

```
$HOME/tools/RCADE/examples/CTCF/CTCF.fasta
```

2. The ChIP-seq BED file: The main output from the analysis of ChIP-seq raw data is often a BED file (Fig. 3), which contains the genomic coordinates of the peaks and their associated scores (*see* **Note 3**). An array of analysis tools exists in order to generate peak coordinates from the raw ChIP-seq data. The complete protocol for one such tool, MACS [20], can be found in [21]. The "examples" folder in the RCADE directory contains BED files for a few C2H2-ZFPs, based on ChIP-seq data described in [22]. We will use CTCF as an example in this protocol—The following file in the RCADE package contains the genomic coordinates for the 500 bp region around the summits of CTCF

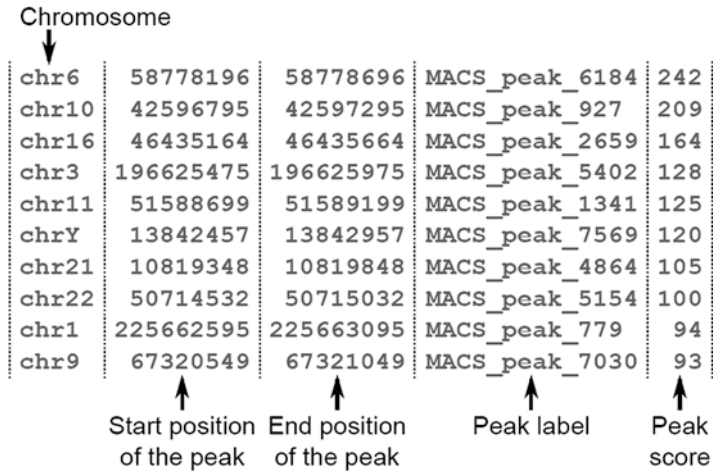


Fig. 3 The protocol that is explained in this chapter requires a peak BED file that has at least five columns, as shown in this figure

peaks in HEK293T cells (the coordinates are based on the hg19 assembly of the human genome):

```
$HOME/tools/RCADE/examples/CTCF/GSM1407629.
summit_500bp.bed
```

3. Genome sequence: RCADE requires the DNA sequence of the peaks in order to optimize the recognition code predictions and identify the ZFs that engage with DNA. To obtain these sequences, a copy of the human genome sequence needs to be downloaded (in FASTA format). The hg19 assembly of the human genome can be downloaded from a variety of sources, including the GATK Resource Bundle:
   (a) Create a folder where you want to store the human genome sequence. For example:

   ```
   $ mkdir -p $HOME/resources/hg19
   ```

   (b) Go to the folder:

   ```
   cd $HOME/resources/hg19
   ```

   (c) Download the FASTA file of the genome sequence:

   ```
   $ wget ftp://gsapubftp-anonymous@ftp.broadin-
   stitute.org/bundle/hg19/ucsc.hg19.fasta.gz
   ```

   (d) Download the index file of the genome sequence:

   ```
   $ wget ftp://gsapubftp-anonymous@ftp.broadin-
   stitute.org/bundle/hg19/ucsc.hg19.fasta.fai.gz
   ```

   (e) Unpack the files:

   ```
   $ gunzip ucsc.hg19.fasta.gz ucsc.hg19.fasta.fai.gz
   ```

3 Methods

3.1 Predicting the DNA-Binding Preference of C2H2-ZFPs from Protein Sequence

The scripts provided in the RCADE package can be used to directly predict the DNA-binding preference of different ZFs in a given C2H2-ZFP. Here, we use CTCF as an example:

1. Go to the directory where RCADE was installed:
   ```
   $ cd $HOME/tools/RCADE
   ```
2. Run the "`RC.sh`" script (*see* **Note 4**):
   ```
   $ bash RC.sh CTCF_predictions ./examples/CTCF/CTCF.fasta
   ```
3. The output files will be located in "`./out/CTCF_predictions`":
   (a) The file "`results.RF_out.txt`" contains the PFMs predicted for each individual ZF.
   (b) The file "`results.PFM.txt`" will contain the concatenation of the position frequency matrices (PFMs) that are predicted for the ZFs, in a format similar to what is used in the CisBP database (http://cisbp.ccbr.utoronto.ca/). This PFM reflects the predicted binding preference of the protein if it binds to the DNA using all of its ZFs (*see* **Note 5**).
   (c) The file "`results.ps`" is a PostScript file that contains the motif logo presentation of the concatenated PFM.
   (d) The files "`log.info.txt`, `log.step1.txt`" and "`log.step2.txt`" will contain the information messages produced during the analysis.

3.2 Recognition Code-Assisted Analysis of ChIP-Seq Data

A more accurate approach for inferring the DNA-binding preferences of C2H2-ZFPs is to combine the recognition code predictions with the in vivo binding site information, such as those obtained from ChIP-seq experiments. By combining ChIP-seq data and recognition code predictions, we can correct the inaccuracies in the predictions, identify the ZFs that contribute to DNA recognition, and pinpoint the in vivo binding locations of the protein.

1. Identify the peaks that should be used for motif discovery: It is often preferable to use only the most reliable peaks for motif discovery, as these are less likely to include false-positive hits, and often show stronger enrichment of the preferred sequence of the TF (*see* **Note 6**).
   (a) Go to the folder where RCADE was installed:
   ```
   $ cd $HOME/tools/RCADE
   ```

(b) Run the following command to sort the BED file (in this example the summits of the CTCF peaks) according to the scores in the fifth column, keeping only the top 500 peaks (*see* **Note 7**):

```
$ sort -r -g -k5 ./examples/CTCF/GSM1407629.
summit_500bp.bed | head -n 500 > ./examples/
CTCF/GSM1407629.summit_500bp.top500.bed
```

2. Get the FASTA sequences of the selected peaks:

```
$ bedtools getfasta -fi $HOME/resources/hg19/
ucsc.hg19.fasta -bed ./examples/CTCF/GSM1407629.
summit_500bp.top500.bed -fo ./examples/CTCF/
GSM1407629.summit_500bp.top500.fasta
```

3. Run the "`RCOpt.sh`" script to predict the binding preferences of individual ZFs and concatenate/optimize them based on the ChIP-seq peaks (*see* **Note 8**):

```
$ bash RCOpt.sh CTCF_ChIP ./examples/CTCF/CTCF.
fasta ./examples/CTCF/GSM1407629.summit_500bp.
top500.fasta
```

4. The output files will be located in "`./out/CTCF_ChIP`":

(a) The file "`results.opt.PFM.txt`" contains the PFM of the top-scoring optimized motif, in CisBP format.

(b) The file "`results.opt.PFM.meme.txt`" contains the PFM of the top-scoring optimized motif in a format compatible with the MEME suite (http://meme-suite.org/).

(c) The file "`results.PFM.txt`" contains all seed motifs (predicted directly from the protein sequence using the recognition code) and their optimized versions (the optimized motif names end with the phrase "opt"). The motifs are in CisBP format.

(d) The file "`results.ps`" is a PostScript file that visualizes all the optimized motifs, sorted by their area under the ROC curve (AUROC) for distinguishing ChIP-seq peaks from dinucleotide-shuffled sequences.

(e) The file "`results.opt.ps`" is similar to `results.ps`, except that it only includes the top-scoring optimized motif (Fig. 4).

(f) The file "`results.report.txt`" contains a tab-delimited table summarizing the optimization results for the motifs.

(g) The files "`log.info.txt`, `log.step1.txt`" and "`log.step2.txt`" will contain the information messages produced during the analysis.

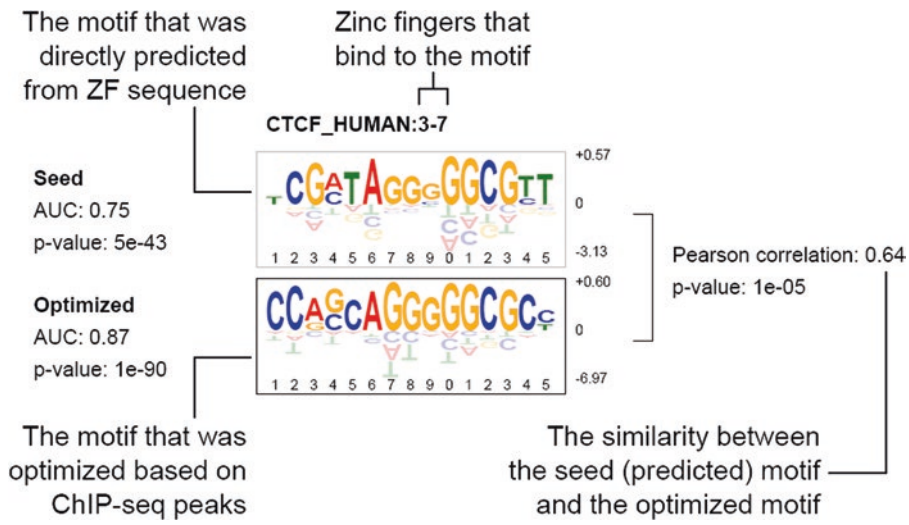


Fig. 4 An annotated example of the output of RCADE

3.3 Using RCADE Motifs in Downstream Analyses

The output of RCADE can be used for a variety of downstream analyses, including identification of potential binding sites within a given set of sequences. In this example, we use the RCADE motif for CTCF to pinpoint the binding site of CTCF within the ChIP-seq peaks.

1. Go the folder where RCADE was installed:
   ```
   $ cd $HOME/tools/RCADE
   ```
2. Get the FASTA sequences of all the CTCF peaks:
   ```
   $ bedtools getfasta -fi $HOME/resources/hg19/
   ucsc.hg19.fasta -bed ./examples/CTCF/GSM1407629.
   summit_500bp.bed -fo ./examples/CTCF/GSM1407629.
   summit_500bp.fasta
   ```
3. Use FIMO from the MEME suite to scan peak sequences for instances of the CTCF motif identified by RCADE (*see* **Note 9**):
   ```
   $HOME/tools/meme/bin/fimo --o ./out/CTCF_ChIP/fimo_
   out ./out/CTCF_ChIP/results.opt.PFM.meme.txt ./ex-
   amples/CTCF/GSM1407629.summit_500bp.fasta
   ```
4. The output files will be located in "`./out/CTCF_ChIP/fimo_out`":
   (a) The file "`fimo.html`" is an HTML file containing the list of motif instances in the provided sequences.
   (b) The file "`fimo.txt`" contains the motif instances in tab-delimited text format.
   (c) The file "`fimo.gff`" contains the motif instances in GFF3 format (*see* **Note 10**). The coordinates are relative to the peaks.

5. Run the following command to convert the GFF3 file coordinates to hg19-based coordinates:

```
$ awk -v FS='\t' -v OFS='\t' 'NR==1 { print $0 }
NR>1 { split( $1, chr, ":" ); split( chr[2], coor,
"-" ); $1=chr[1]; $4+=coor[1]; $5+=coor[1]; print
$0; }' ./out/CTCF_ChIP/fimo_out/fimo.gff > ./out/
CTCF_ChIP/fimo_out/fimo.hg19.gff
```

6. Upload the resulting GFF3 file on the UCSC Genome Browser to visualize and navigate the motif hits:
   (a) Go to https://genome.ucsc.edu/index.html using a Web browser.
   (b) From the "Genomes" menu on top, select "Human GRCh37/hg19".
   (c) On the new page, from the "My Data" menu on top, select "Custom Tracks".
   (d) On the new page, in the section titled "Paste URLs or data", click on "Choose File", and select the file "`./out/CTCF_ChIP/fimo_out/fimo.hg19.gff`" that was created in the previous step. Then click "Submit".
   (e) On the new page, select the option to view in "Genome Browser" and click on the "go" button.
   (f) You can now use the UCSC Genome Browser navigation tools to browse the CTCF-binding sites.

4 Notes

1. On systems that do not support the command "`wget`", use "`curl`" with the "`-kLO`" options. For example:

```
$ curl -kLO https://github.com/arq5x/bedtools2/re-
leases/download/v2.25.0/bedtools-2.25.0.tar.gz
```

2. The amino acid sequences of most proteins can be downloaded in the FASTA format from UniProt [23] (http://www.uniprot.org/).
3. The BED file specifications are explained here: https://www.ensembl.org/info/website/upload/bed.html.
4. The general format for calling the `RC.sh` script in the RCADE package is as follows:

```
$ RC.sh <job_name> <C2H2_protein_fasta>
```

   (a) The parameter `job_name` is a label that is used to create the output and temporary directories.
   (b) The parameter `C2H2_protein_fasta` contains the path to the file that contains the amino acid sequence of the C2H2-ZFP of interest in the FASTA format.

5. In the output of RC.sh, the files `results.PFM.txt` and `results.ps` may contain multiple PFMs—this happens if the provided C2H2-ZFP sequence contains multiple "arrays", i.e., multi-ZF stretches that are separated from each other by long linkers in the protein sequence (such as ZNF667 in Fig. 2).
6. C2H2-ZFP-binding sites are often within repeat elements of the genome. These repeat elements may make it problematic to infer the DNA-binding preference of the C2H2-ZFP from its in vivo binding sites, given that these elements may share a substantial amount of sequence similarity. The identified motifs are often more likely to represent the *bona fide* binding preference of the protein if peaks that overlap repeat elements are removed from the analysis:
   (a) Download a BED file containing the genomic coordinates of repeat elements, e.g., from the RepeatMasker track of the UCSC Genome Browser using the Table Browser tool: https://genome.ucsc.edu/cgi-bin/hgTables).
   (b) Use the intersect tool from BEDTools to remove peaks that overlap repeat elements. For example:

```
$ bedtools intersect -a peaks.bed -b repeats.
bed -v > peaks_without_repeats.bed
```

7. The number of top peaks that will be used for motif optimization can be changed by changing the parameter passed to the `head` command. Usually, the top 500 peaks provide a good starting point for motif finding.
8. The general format for calling the `RCOpt.sh` script in the RCADE package is as follows:

```
$ RCOpt.sh <job_name> <C2H2_protein_fasta>
<ChIPseq_peaks_fasta>
```

   (a) The parameter `job_name` is a label that is used to create the output and temporary directories.
   (b) The parameter `C2H2_protein_fasta` contains the path to the file that contains the amino acid sequence of the C2H2-ZFP of interest in the FASTA format.
   (c) The parameter `ChIPseq_peaks_fasta` contains the path to the file that contains the nucleotide sequence of the ChIP-seq peaks that will be used for motif optimization. We recommend to use the 500 bp sequence surrounding the peak summits.
9. The general format for calling FIMO in the MEME suite is as follows:

```
$ fimo [options] <motif_file> <sequence_file >
```

The full list of options for running FIMO can be found at http://meme-suite.org/doc/fimo.html.

10. The GFF3 file specifications can be found at http://useast.ensembl.org/info/website/upload/gff3.html.

References

1. Weirauch MT, Hughes TR (2011) A catalogue of eukaryotic transcription factor types, their evolutionary origin, and species distribution. Subcell Biochem 52:25–73
2. Vaquerizas JM, Kummerfeld SK, Teichmann SA, Luscombe NM (2009) A census of human transcription factors: function, expression and evolution. Nat Rev Genet 10:252–263
3. Lambert SA, Jolma A, Campitelli LF, Das PK, Yin Y, Albu M, Chen X, Taipale J, Hughes TR, Weirauch MT (2018) The human transcription factors. Cell 172:650–665
4. Garton M, Najafabadi HS, Schmitges FW, Radovani E, Hughes TR, Kim PM (2015) A structural approach reveals how neighbouring C2H2 zinc fingers influence DNA binding specificity. Nucleic Acids Res 43:9147–9157
5. Pabo CO, Peisach E, Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. Annu Rev Biochem 70:313–340
6. Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. Science 252:809–817
7. Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys2His2 zinc finger proteins. Annu Rev Biophys Biomol Struct 29:183–212
8. Seeman NC, Rosenberg JM, Rich A (1976) Sequence-specific recognition of double helical nucleic acids by proteins. Proc Natl Acad Sci U S A 73:804–808
9. Nardelli J, Gibson TJ, Vesque C, Charnay P (1991) Base sequence discrimination by zinc-finger DNA-binding domains. Nature 349:175–178
10. Gupta A, Christensen RG, Bell HA, Goodwin M, Patel RY, Pandey M, Enuameh MS, Rayla AL, Zhu C, Thibodeau-Beganny S, Brodsky MH, Joung JK, Wolfe SA, Stormo GD (2014) An improved predictive recognition model for Cys(2)-His(2) zinc finger proteins. Nucleic Acids Res 42:4800–4812
11. Persikov AV, Wetzel JL, Rowland EF, Oakes BL, Xu DJ, Singh M, Noyes MB (2015) A systematic survey of the Cys2His2 zinc finger DNA-binding landscape. Nucleic Acids Res 43:1965–1984
12. Persikov AV, Osada R, Singh M (2009) Predicting DNA recognition by Cys2His2 zinc finger proteins. Bioinformatics 25:22–29
13. Myers S, Bowden R, Tumian A, Bontrop RE, Freeman C, MacFie TS, McVean G, Donnelly P (2010) Drive against hotspot motifs in primates implicates the PRDM9 gene in meiotic recombination. Science 327:876–879
14. Jacobs FM, Greenberg D, Nguyen N, Haeussler M, Ewing AD, Katzman S, Paten B, Salama SR, Haussler D (2014) An evolutionary arms race between KRAB zinc-finger genes ZNF91/93 and SVA/L1 retrotransposons. Nature 516:242–245
15. Najafabadi HS, Garton M, Weirauch MT, Mnaimneh S, Yang A, Kim PM, Hughes TR (2017) Non-base-contacting residues enable kaleidoscopic evolution of metazoan C2H2 zinc finger DNA binding. Genome Biol 18:167
16. Najafabadi HS, Albu M, Hughes TR (2015) Identification of C2H2-ZF binding preferences from ChIP-seq data using RCADE. Bioinformatics 31:2879–2881
17. Najafabadi HS, Mnaimneh S, Schmitges FW, Garton M, Lam KN, Yang A, Albu M, Weirauch MT, Radovani E, Kim PM, Greenblatt J, Frey BJ, Hughes TR (2015) C2H2 zinc finger proteins greatly expand the human regulatory lexicon. Nat Biotechnol 33:555–562
18. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842
19. Bailey TL, Boden M, Buske FA, Frith M, Grant CE, Clementi L, Ren J, Li WW, Noble WS (2009) MEME SUITE: tools for motif discovery and searching. Nucleic Acids Res 37:W202–W208
20. Zhang Y, Liu T, Meyer CA, Eeckhoute J, Johnson DS, Bernstein BE, Nusbaum C, Myers RM, Brown M, Li W, Liu XS (2008) Model-based analysis of ChIP-Seq (MACS). Genome Biol 9:R137
21. Feng J, Liu T, Qin B, Zhang Y, Liu XS (2012) Identifying ChIP-seq enrichment using MACS. Nat Protoc 7:1728–1740

22. Schmitges FW, Radovani E, Najafabadi HS, Barazandeh M, Campitelli LF, Yin Y, Jolma A, Zhong G, Guo H, Kanagalingam T, Dai WF, Taipale J, Emili A, Greenblatt JF, Hughes TR (2016) Multiparameter functional diversity of human C2H2 zinc finger proteins. Genome Res 26:1742–1752
23. The UniProt C (2017) UniProt: the universal protein knowledgebase. Nucleic Acids Res 45:D158–D169

Chapter 3

Design and Application of 6mA-Specific Zinc-Finger Proteins for the Readout of DNA Methylation

Johannes A. H. Maier and Albert Jeltsch

Abstract

Designed zinc-finger (ZnF) proteins can recognize AT base pairs by H-bonds in the major groove, which are disrupted, if the adenine base is methylated at the N6 position. Based on this principle, we have recently designed a ZnF protein, which does not bind to DNA, if its recognition site is methylated. In this review, we summarize the principles of the recognition of methylated DNA by proteins and describe the design steps starting with the initial bacterial two-hybrid screening of three-domain ZnF proteins that do not bind to CcrM methylated target sites, followed by their di- and tetramerization to improve binding affinity and specificity. One of the 6mA-specific ZnF proteins was used as repressor to generate a methylation-sensitive promoter/repressor system. This artificial promoter/repressor system was employed to regulate the expression of a CcrM DNA methyltransferase gene, thereby generating an epigenetic system with positive feedback, which can exist in two stable states, an off-state with unmethylated promoter, bound ZnF and repressed gene expression, and an on-state with methylated promoter and active gene expression. This system can memorize transient signals approaching bacterial cells and store the input in the form of DNA methylation patterns. More generally, the ability to bind to DNA in a methylation-dependent manner gives ZnF and TAL proteins an advantage over CRISPR/Cas as DNA-targeting device by allowing methylation-dependent genome or epigenome editing.

Key words Zinc-finger protein, DNA methylation, Adenine-N6 methylation, Protein design, Bacterial gene expression, Epigenetic circuit

1 DNA Methylation and Methylation Readout

Epigenetic processes encode inherited information that control chromatin organization for gene expression regulation [1, 2]. One important epigenetic modification is DNA methylation, found in prokaryotes and eukaryotes. It does not interfere with base pairing and adds additional information to the DNA without altering the DNA sequence [3].

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_3,

1.1 Role of DNA Methylation in Mammals

In mammals, DNA methylation is mainly restricted to methylation of cytosines at the C5 position in 5′-CpG-3′ sequences, but it also occurs at non-CpG sites [4, 5]. Recently, N6-methyladenine has been discovered in mammalian DNA as well, but its biological role is still unclear [6]. Cytosine-C5 methylation plays crucial roles in cellular differentiation and often is involved in the onset and progression of diseases [7]. In particular, promoter methylation leads to gene repression, whereas methylation in gene bodies is observed in transcribed genes [8, 9]. Furthermore, DNA methylation is involved in genomic imprinting, X-chromosome inactivation, and silencing of transposable elements.

1.2 DNA Methylation in Prokaryotes

In bacteria, three types of DNA methylation occur, 5-methylcytosine (5mC), N4-methylcytosine (4mC), and N6-methyladenine (6mA) (Fig. 1). In each of these modified bases, the methyl group projects into the major groove of the DNA, where it can be recognized by DNA-binding proteins. Most of the known bacterial DNA methyltransferases (MTases) are part of restriction/modification (RM) systems [3, 10–12]. These systems serve to protect prokaryotes from bacteriophage infections. RM systems employ a restriction endonuclease and a DNA MTase, which both recognize the same target DNA sequence, usually a 4–8 base pair palindromic site [13]. The restriction endonuclease cleaves the DNA at the target sequence only if it is in an unmethylated state, as found on incoming phage DNA during the early steps of infection. The host cell DNA, however, is kept in a methylated state by the corresponding DNA MTase and hence is protected from cleavage. These systems are complemented by the adaptable CRISPR-Cas systems that

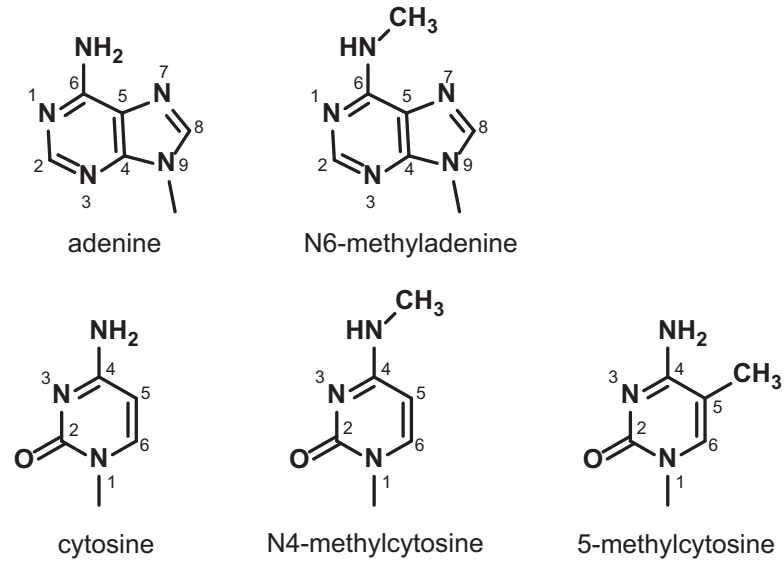


Fig. 1 DNA bases and their naturally occurring methylated variants

have been discovered only recently as an additional epigenetic defense system of bacteria and archaea against invading phages, which is independent of DNA methylation [14].

Aside from MTases being part of RM systems, there are so-called solitary or orphaned MTases that do not serve in protection of DNA from cognate endonucleases. The *Escherichia coli* deoxyadenosine DNA methyltransferase (Dam) and the *Caulobacter crescentus* cell cycle-regulated methyltransferase (CcrM) are two well-characterized enzymes of this kind [3, 10–12]. *E. coli* Dam methylates adenine residues in 5′-G**A**TC-3′ sequences and is involved in DNA mismatch repair, initiation of chromosome replication, and regulation of gene expression [11]. *Caulobacter crescentus* CcrM methylates adenine residues in 5′-G**A**NTC-3′ sequences and it is essential for *C. crescentus*, at least under certain growth conditions, and plays a central role in regulation of the cell cycle [10, 15]. Moreover, there are certain phenomena in bacteria that resemble epigenetic mechanisms in higher organisms like heritable DNA methylation patterns involved in regulation of gene expression at the *pap* operon and the *agn43* gene, as well as IS10 and traJ transposons [10–12].

2 Natural Readout of DNA Methylation

DNA-binding proteins recognize the DNA sequence by two processes, so-called direct and indirect readout [16]. In direct readout, the DNA sequence is identified by hydrogen bonds and van der Waals interactions between the side chains of critical amino acids and the edges of the target base pairs, mostly in the major groove of the DNA (Fig. 2). The direct readout mechanism had been predicted more than 40 years ago in a seminal paper by Seeman and colleagues [17] and validated in innumerable structures of proteins in complex with their specific DNA till now. Later it has been recognized that DNA adopts specific conformations including DNA bending, twisting, or unwinding. As the accessibility of specific conformations depends on the DNA sequence, this process can lead to an indirect readout of the sequence, in which proteins contact the backbone of the DNA, thereby enforcing a specific DNA conformation which is possible only on DNA with a defined sequence [18].

Readout of DNA methylation in general is based on the sequence- and methylation-specific DNA binding of proteins, which then cause downstream effects on gene expression and chromatin structure. In principle, one has to distinguish proteins that prefer binding to methylated DNA and others that prefer binding to unmethylated DNA.

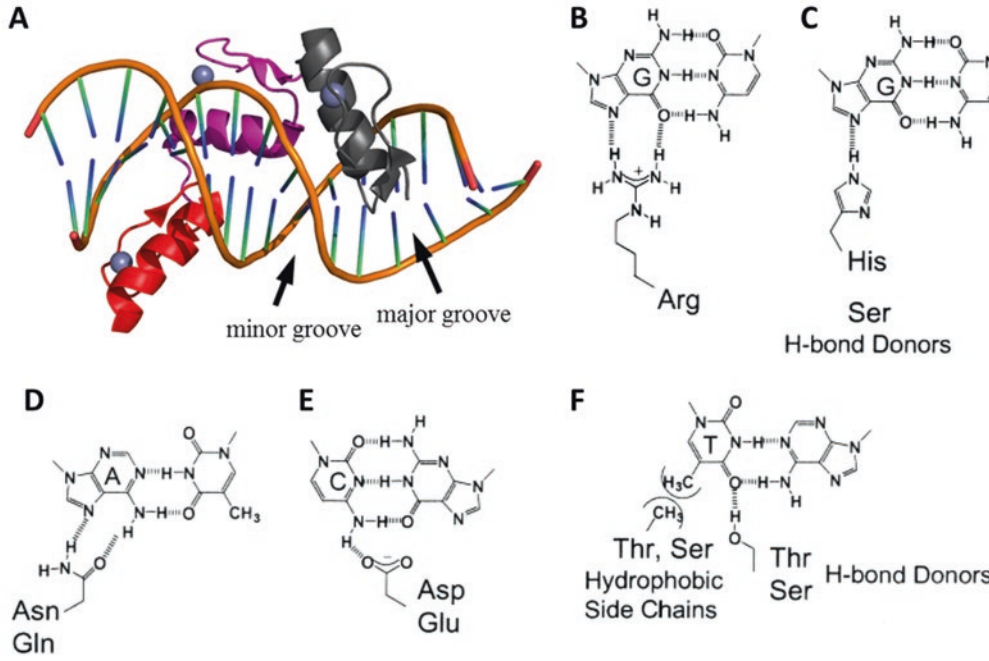


Fig. 2 DNA binding and sequence recognition by zinc-finger proteins. (**a**) Structure of the Gfi-1 zinc fingers 3–5 in complex with DNA (PDB ID: 2KMK, [49]). The individual zinc-finger domains are colored red, violet, and grey. The complexed zinc ions are shown as light grey spheres. The zinc-finger protein binds to the major groove of the DNA. (**b-f**) Major groove interactions of amino acid side chains with DNA bases via H-bonds and hydrophobic interactions. Picture modified from [50]

2.1 Proteins that Specifically Bind to Methylated DNA

Two families of proteins which specifically bind to methylated CpG sites have been identified in mammals, the methyl-CpG-binding domain (MBD) proteins (MeCP2, MBD1-6, SETDB1-2, TIP5/BAZ2A, and BAZ2B) and a C2H2 zinc-finger domain transcription factor family including Kaiso, Zfp57, and KLF4 [19, 20]. A third family, the SET and RING finger-associated (SRA) domain proteins, recognizes hemimethylated CpG sites. They employ base flipping to identify the methylated cytosine [21] and will not be discussed here. MBD and Kaiso proteins play crucial cellular roles, since they are the link between the cytosine modifications and chromatin interpretation. They are involved in co-repressor complexes with chromatin remodelers, splicing factors, and RNA-processing proteins; recruit proteins such as DNMTs to chromatin; and participate in DNA demethylation and repair [19, 20].

The MBD family is composed of 11 proteins all containing a characteristic highly conserved MBD domain. This domain binds with a variable degree of selectivity to methylated DNA. The first identified member of this group was MeCP2, discovered in the group of Adrian Bird [22]. Today, several crystal and NMR structures have been determined for MBD domains in complex with

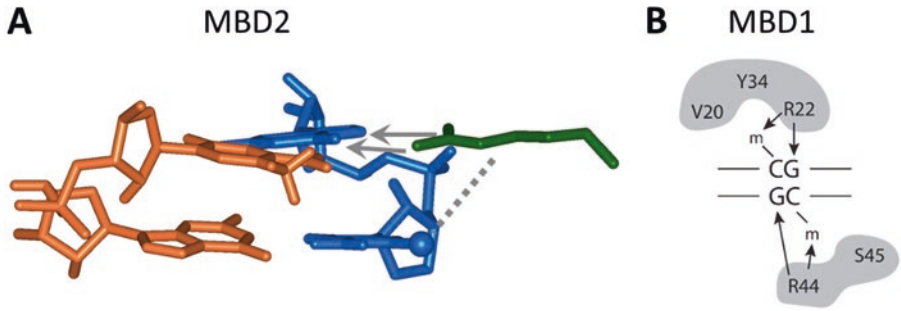


Fig. 3 Principles of methylation-specific binding to CpG sites by MBD domains and other proteins. (**a**) Schematic representation of the 5mC-Arg-Gua triad in MBD2 in which one critical arginine interacts with the methylated C and its adjacent G residue in the same DNA strand (pdb structure 2KY8). The methylene part of the Arg forms van der Waals contacts with the cytosine methyl group and the guanidinium group two classical hydrogen bonds with the guanine. Arg24 is shown in green, the DNA strands of the CpG site in blue and orange. The recognized methyl group is highlighted by a blue ball. (**b**) Schematic representation of the methyl CpG recognition by MBD1. The two 5mC bases in the fully methylated CpG are recognized by two separate protein recognition modules, formed by V20-R22-Y34 and R44-S45. Both arginine residues, R22 and R44, form Arg-triads. Figure adapted from [20]

methylated DNA showing that these domains bind with two loops asymmetrically to the symmetrically methylated CpGs in DNA (Fig. 3). The recognition of the methylated CpG site is based on a direct readout by Arg residues, in which the guanidinium group forms a classical double H-bond with the guanine base. In addition, the methylene part of the Arg and the methyl group of the 5mC in the same DNA strand interact, forming a highly conserved 5mC-Arg-G triad [23]. Depending on the protein, the recognition of the methyl group is further supported by additional aliphatic residues forming a small hydrophobic cavity.

Kaiso, the first identified member of the C2H2 zinc-finger domain family of methyl-binding proteins [24], contains three C2H2 zinc fingers. The preferential sequence determinants for DNA binding of Kaiso are two methylated CpG dinucleotides or a TG site. Kaiso recognizes the methyl group either in mCpG or TpG dinucleotides using a 5mC-Arg-Gua triad structure similarly as MBD proteins [20, 25]. Zfp57 belongs to the Krüppel-associated box (KRAB) zinc-finger family. It recognizes the TGCCGC motif containing a methylated CpG dinucleotide and also uses the 5mC-Arg-Gua triad for methylated cytosine recognition [26]. A similar mechanism has been observed with KLF4 [27].

In addition, a recent study has shown that many more human transcription factors bind to methylated DNA with some preference over unmethylated [28], although not with full specificity. Often they can bind to binding sites either containing a TA or mCG site and a hydrophobic pocket is used for the interaction with the methyl group in T or 5mC, making the interaction methylation specific.

2.2 Proteins that Specifically Bind to Unmethylated DNA

Based on the direct readout of DNA sequence in the major groove, which is employed by many families of transcription factors, many proteins are prevented from high-affinity interaction with their cognate DNA recognition sequences, when the DNA target site is methylated. Examples for this include transcription factors of the myc, CREB, and members of the E2F family. A recent study reported strong inhibition of DNA binding for 23% of all human transcription factors [28]. In addition, the CXXC domain proteins bind to DNA sites containing unmethylated CpG sites [29]. These domains are found in many chromatin proteins, where they contribute to the methylation-specific targeting of these factors to chromatin sites.

3 Design of an Artificial mA-Specific ZnF Protein

As described, many DNA-binding proteins interact with DNA in a sequence- and methylation-specific manner. However, so far only few examples of designed DNA-binding proteins with 5mC specificity have been reported for ZnF [30] and TAL proteins [31]. One recent example is our design of a 6mA-specific ZnF protein [32] that will be summarized and discussed in this section.

3.1 Design of a 6mA-Specific ZnF Protein

ZnF proteins can recognize AT base pairs via asparagine or glutamine residues and form specific H-bond contacts to the adenine base via the N6 and N7 position. Adenine-N6 methylation should prevent these H-bonds and hence strongly impair the binding of methylated AT base pairs (Fig. 4). In our previous work, we theoretically designed several 6mA-specific ZnF proteins consisting of three zinc-finger domains with target sequences that overlapped with the CcrM target sequence (5′-GANTC-3′) using the Zinc Finger Targeter (ZiFiT) [33, 34]. We then used a bacterial two-hybrid reporter assay [35] to test if the designed ZnF proteins specifically bind to their target sequences (Fig. 5). In this assay, expression of a *lacZ* reporter gene is driven by the recruitment of an RNA polymerase-Gal4 fusion protein to target promoters by a ZnF-Gal11p fusion protein binding to its target sequence placed within the promoter of the *lacZ*. Four designed ZnF proteins were tested and reduced reporter gene activation was observed after DNA methylation for all of them. The most promising zinc-finger protein candidate (ZnF_1012) was used for further experimental steps. This ZnF protein binds to a 5′-GGAGAAGAA-3′ sequence, which can be overlapped with the target sequence of CcrM (5′-G**A**NTC-3′) resulting in a 5′-$\underline{\text{GGAGAA}\,\overline{\text{G}\mathbf{A}\text{A}\text{TC}}}$-3′ sequence (the ZnF-binding site is underlined, the CcrM target site is overlined, the methylated A is bolded) (Fig. 5). Binding of the ZnF protein to this site was shown to be inhibited by DNA methylation.

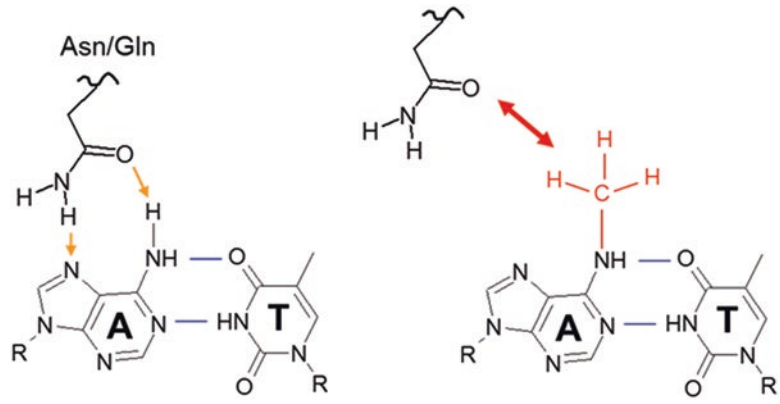


Fig. 4 AT base pair sequence direct readout by asparagine or glutamine residues with two hydrogen bonds. This interaction is blocked by adenine-N6 methylation. This figure has been taken from [32] with permission

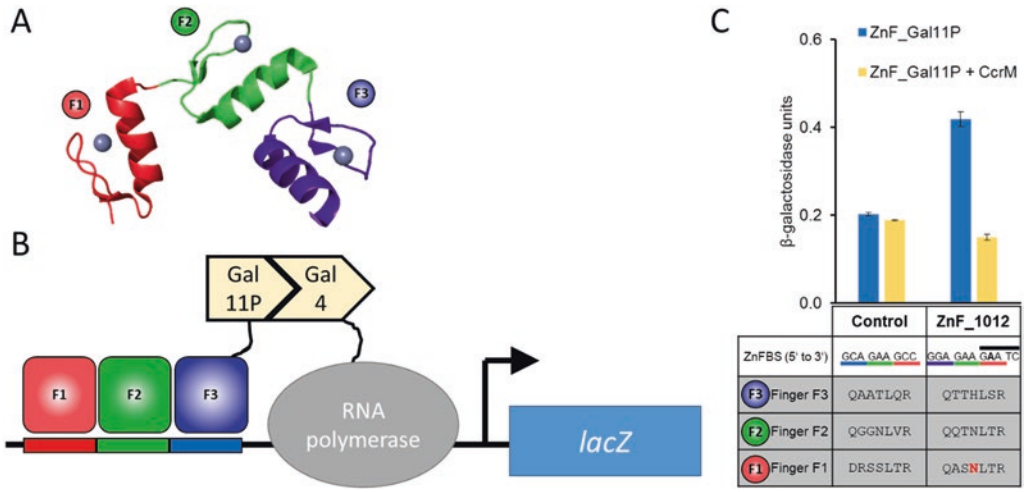


Fig. 5 Design of a ZnF module specifically interacting with unmethylated DNA. (**a**) Scheme of a three-finger ZnF protein able to bind to nine base pairs of DNA. (**b**) Scheme of the bacterial two-hybrid system used to select ZnF modules, which specifically bind the target sequence [35]. (**c**) Results of the screening with the best candidate protein ZnF_1102 [32]. A control ZnF protein binds its GCAGAAGCC target sequence, and this interaction is not altered by expression of the CcrM MTase in the cells. In contrast, ZnF_1102 binds more strongly to its GGAGAAGAA target sequence. When this site is overlapping with a CcrM site at the end, DNA binding is reduced if CcrM is expressed, indicating a methylation-specific interaction. Data were taken from [32]

3.2 Improvement of Methylation-Specific Binding by Multimerization

Although showing reduced binding in the presence of DNA methylation in the bacterial two-hybrid system, we realized that for this individual ZnF module the ΔG difference in binding to methylated and unmethylated DNA might be insufficient for biological applications. Hence, we followed classical design principles of nature and fused a coiled-coil domain for dimerization [36] to the

ZnF protein to induce its dimerization in the cell. We expected that the dimeric repressor would bind palindromic DNA-binding sites with two reading heads, similarly as most bacterial repressors (Fig. 6a, b). Importantly, DNA binding of both reading heads is regulated by DNA methylation leading to a doubling of the expected ΔG difference (ΔΔG) of binding to methylated and unmethylated DNA. However, later gene expression studies (described in Subheading 3.3) still showed an insufficient repression of gene expression by the dimeric repressor [32].

Taking natural repressor systems as paradigm, we next fused the synthetic zinc-finger protein to the coiled-coil domain from GCN4-p-LI, which mediates tetramerization [37]. The resulting tetrameric ZnF protein could simultaneously bind to two palindromic double-recognition sites, both regulated by DNA methylation (Fig. 6c). We introduced additional palindromic double-binding sites for the ZnF repressor upstream and downstream of the transcriptional start site. The first double-binding site was designed to surround the -35 region of the promoter as shown in Fig. 6b. The second double-binding site was placed 100 bps upstream of the first one in a manner resembling the arrangement of pseudo-operator sites in the lac operon [38]. A third double-binding site was placed downstream in the operon. All

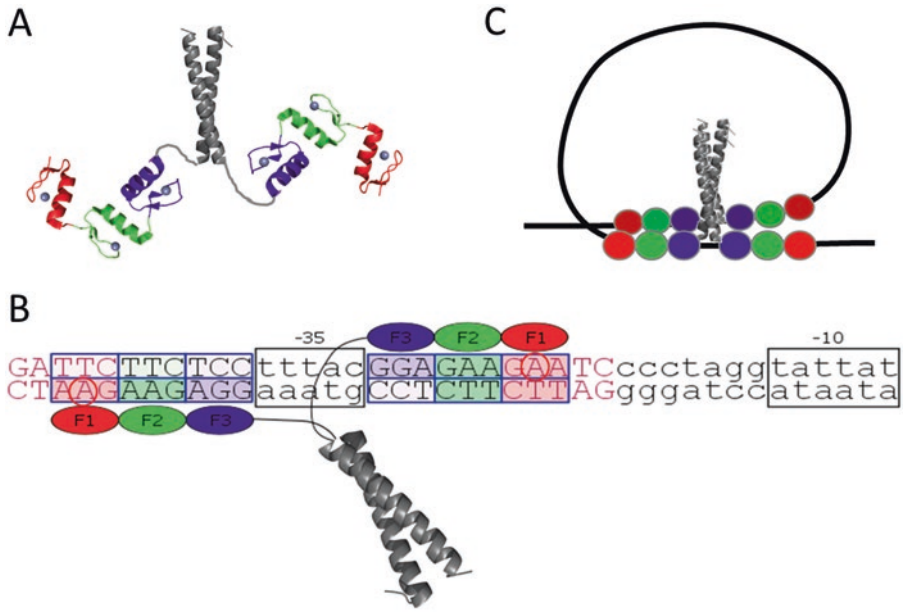


Fig. 6 Enhancement of methylation-specific DNA binding by multimerization of the ZnF protein. (**a**) Scheme of the dimerization of the ZnF module using a coiled-coil domain. (**b**) Scheme of the palindromic double-binding site recognized by the dimeric ZnF. The CcrM recognition site is printed in red and the adenine residue methylation of which prevents DNA binding of the F1 domain is highlighted by a red circle. The binding sites surround the -35 region of a bacterial gene promoter allowing to employ the ZnF protein as methylation-specific repressor. (**c**) Tetramerization of the ZnF protein allows simultaneous interaction with two palindromic recognition sites via DNA looping, similarly as observed in bacterial repressors, for example the lac repressor

these binding sites can be methylated by CcrM to regulate repressor binding, leading to a system which featured very tight inhibition in the absence of DNA methylation and strong reporter gene expression when the recognition sites were methylated [32].

3.3 Application of the mA-Specific ZnF in Synthetic Circuit Design

We designed an artificial expression system that is regulated by DNA methylation (Fig. 7). In this system, one palindromic methylation-sensitive ZnF-binding site was placed surrounding the -35 region of the promoter of a GFP reporter gene showing a methylation-dependent repression of reporter gene expression [32]. Next, we aimed to design an artificial epigenetic memory system, in which the methylation-dependent promoter was used to regulate the expression of a CcrM MTase gene. By this, a system with positive feedback was obtained, which was supposed to exist in two stable states, an off-state and an on-state (Fig. 8). The off-state should be maintained by the engineered zinc-finger repressor and the on-state by constant promoter methylation by CcrM, which hinders zinc-finger DNA binding. An enhanced green fluorescent protein gene was used as a reporter in front of the *ccrM* gene and this setup was called reporter-maintenance operon.

However, as mentioned above, we faced the problem that the off-state of the initial system using a dimeric ZnF protein was not stable, indicating insufficient repression. To this end, the methylation-sensitive promoter repressor system was further enhanced by using the tetrameric ZnF protein and introducing the additional palindromic binding sites as described in Subheading 3.2. With this system, we were facing the problem of an instable on-state, which was

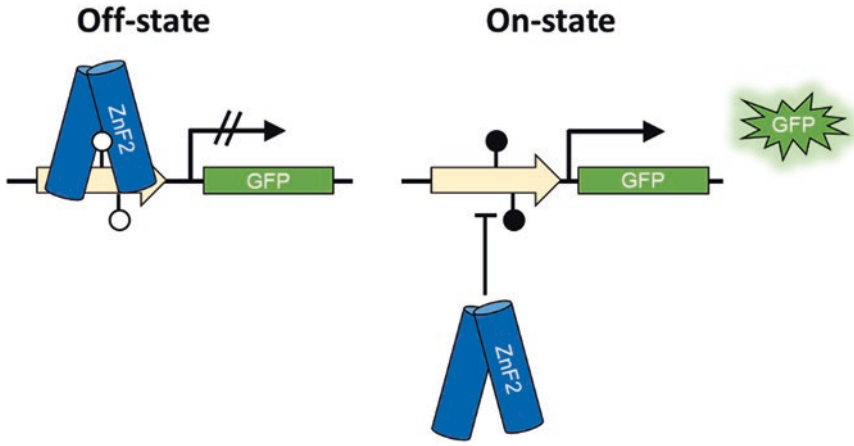


Fig. 7 Scheme of an artificial methylation-sensitive expression system. In the off-state, the synthetic ZnF repressor binds the promoter region of a reporter gene (GFP) and silences gene expression. In the on-state, the promoter region is methylated, repressor binding is hindered, and the *gfp* gene is transcribed. Filled and open lollipops represent methylated and unmethylated 5′-GANTC-3′ sites

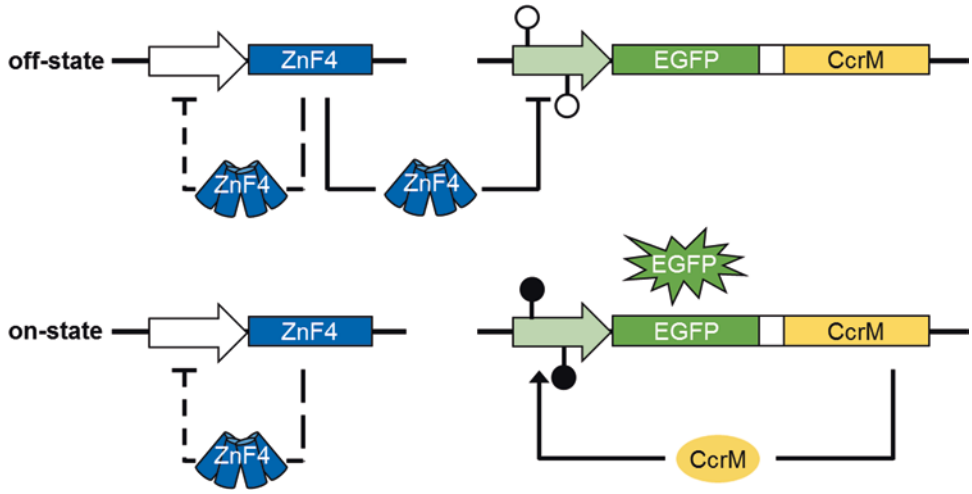


Fig. 8 Schematic circuit design of a synthetic epigenetic memory system with ZnF repressor (ZnF4) and reporter-maintenance operon, consisting of an *egfp* gene and a *ccrM* gene. In the off-state, the ZnF repressor inhibits transcription of the reporter-maintenance operon by binding to the unmethylated promoter region (empty lollipops). Once the system is switched to the on-state, EGFP and CcrM are expressed and binding of the ZnF repressor to the promoter region of the reporter-maintenance operon is prevented by DNA methylation (filled lollipops). CcrM constantly re-methylates the operator ZnF-binding sites and keeps the system in the on-state. The zinc-finger repressor regulates its own expression in a manner not affected by methylation (indicated negative feedback with dashed lines). The image is taken from [32] with permission

due to the accumulation of the ZnF repressor in the cells over several hours and competition of the repressor with CcrM for target-site binding. To overcome this problem, we introduced an additional double-binding site for the ZnF repressor in the promoter region of its own gene in order to maintain a low steady-state level of repressor molecules by autoregulation (Fig. 8). This autoregulatory binding site was designed not to overlap with CcrM sites; hence, it cannot be methylated and repressor binding is independent of CcrM expression. This approach mimics several bacterial repressor systems, which employ autoregulation to maintain a low and stable expression of the repressor and minimize stochastic fluctuations of repressor protein amounts in the cell. The final system illustrated in Fig. 8 showed biphasic behavior with stable on- and off-states.

Next, we showed that the expression of the reporter-maintenance operon could be initiated either by loss of the ZnF protein binding due to thermal instability during a heat shock or by introduction of DNA methylation at the ZnF-binding site by an externally expressed CcrM [32]. The external CcrM expression system could be coupled to different input signals like nutrients (arabinose) or DNA damage (ultraviolet light or cisplatin) inducing a switch from the off- to the on-state. After appearance of the external transient trigger signal, the maintenance CcrM MTase methylates the promoter region of the reporter-maintenance

operon and keeps the system in the on-state even after cessation of the trigger signal. These designed synthetic epigenetic memory circuits were able to sense different transient stimuli and store this information in the form of DNA methylation patterns for many cellular generations in *Escherichia coli* [32]. Further, the epigenetic nature of the developed system could be exploited and it was shown that the system is resettable by addition of regulated protein degradation tags to the methyltransferase [32]. Potential applications and extensions of the system are in the field of bacterial life biosensors.

4 Outlook: Application of ZnF Proteins as Methylation-Specific DNA-Binding Proteins

Designer ZnF proteins were developed as programmable DNA-binding devices [39–43]. They are representing the first example in which a rational design of a DNA-binding protein with predetermined specificity was possible—a true scientific breakthrough. Later CRISPR/Cas systems were discovered in several bacteria [44, 45]. The CRISPR/Cas9 system targets DNA with a short, single-stranded, so-called sgRNA that forms a double-stranded RNA/DNA hybrid structure with one strand of the target site. Since target recognition is based on simple Watson/Crick base pairing, retargeting the CRISPR/Cas9 complex only requires introduction of a new sgRNA sequence, while in ZnF and TAL systems more complicated protein design has to be conducted to change the DNA binding specificity. For this reason, CRISPR/Cas9 is expected to outcompete proteinaceous DNA-targeting systems in many applications in genome and epigenome editing [14, 46–48]. However, the targeting of CRISPR/Cas9 based on Watson/Crick base pairing between the sgRNA and the target DNA is insensitive for DNA methylation, because this does not affect Watson/Crick base pairing. In contrast, the binding of ZnF and TAL proteins to DNA can be designed to be 5mC or 6mA specific, which gives these targeting systems an important advantage over CRISPR/Cas9 systems in special applications. Methylation-specific interaction with DNA can be used for the design of artificial gene expression circuits [32]. It could also be applied in the context of a human cell, where DNA methylation-specific genome or epigenome editing could be triggered, which would allow to address cells in specific physiological states, defined cellular subtypes, or small subpopulations characterized by a specific epigenetic state. Hence, the potential for a combined reading of DNA sequence and DNA methylation may spur the development of a new generation of ZnF and TAL protein applications in genome targeting.

References

1. Allis CD, Jenuwein T (2016) The molecular hallmarks of epigenetic control. Nat Rev Genet 17:487–500
2. Henikoff S, Greally JM (2016) Epigenetics, cellular memory and gene regulation. Curr Biol 26:R644–R648
3. Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of DNA methyltransferases. Chembiochem 3:274–293
4. Schubeler D (2015) Function and information content of DNA methylation. Nature 517:321–326
5. Ambrosi C, Manzo M, Baubec T (2017) Dynamics and context-dependent roles of DNA methylation. J Mol Biol 429:1459–1475
6. O'Brown ZK, Greer EL (2016) N6-methyladenine: a conserved and dynamic DNA mark. Adv Exp Med Biol 945:213–246
7. Bergman Y, Cedar H (2013) DNA methylation dynamics in health and disease. Nat Struct Mol Biol 20:274–281
8. Baylin SB, Jones PA (2011) A decade of exploring the cancer epigenome—biological and translational implications. Nat Rev Cancer 11:726–734
9. Jones PA (2012) Functions of DNA methylation: islands, start sites, gene bodies and beyond. Nat Rev Genet 13:484–492
10. Wion D, Casadesus J (2006) N6-methyladenine: an epigenetic signal for DNA-protein interactions. Nat Rev Microbiol 4:183–192
11. Marinus MG, Casadesus J (2009) Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. FEMS Microbiol Rev 33:488–503
12. Casadesus J, Low DA (2013) Programmed heterogeneity: epigenetic mechanisms in bacteria. J Biol Chem 288:13929–13935
13. Roberts RJ, Vincze T, Posfai J, Macelis D (2015) REBASE-a database for DNA restriction and modification: enzymes, genes and genomes. Nucleic Acids Res 43:D298–D299
14. Barrangou R, Horvath P (2017) A decade of discovery: CRISPR functions and applications. Nat Microbiol 2:17092
15. Gonzalez D, Collier J (2013) DNA methylation by CcrM activates the transcription of two genes required for the division of Caulobacter crescentus. Mol Microbiol 88:203–218
16. Garvie CW, Wolberger C (2001) Recognition of specific DNA sequences. Mol Cell 8:937–946
17. Seeman NC, Rosenberg JM, Rich A (1976) Sequence-specific recognition of double helical nucleic acids by proteins. Proc Natl Acad Sci U S A 73:804–808
18. Travers AA (1989) DNA conformation and protein binding. Annu Rev Biochem 58:427–452
19. Du Q, Luu PL, Stirzaker C, Clark SJ (2015) Methyl-CpG-binding domain proteins: readers of the epigenome. Epigenomics 7:1051–1073
20. Shimbo T, Wade PA (2016) Proteins that read DNA methylation. Adv Exp Med Biol 945:303–320
21. Jeltsch A (2008) Reading and writing DNA methylation. Nat Struct Mol Biol 15:1003–1004
22. Meehan RR, Lewis JD, McKay S, Kleiner EL, Bird AP (1989) Identification of a mammalian protein that binds specifically to DNA containing methylated CpGs. Cell 58:499–507
23. Hashimoto H, Zhang X, Vertino PM, Cheng X (2015) The mechanisms of generation, recognition, and erasure of DNA 5-methylcytosine and thymine oxidations. J Biol Chem 290:20723–20733
24. Prokhortchouk A, Hendrich B, Jorgensen H, Ruzov A, Wilm M, Georgiev G, Bird A, Prokhortchouk E (2001) The p120 catenin partner Kaiso is a DNA methylation-dependent transcriptional repressor. Genes Dev 15:1613–1618
25. Buck-Koehntop BA, Defossez PA (2013) On how mammalian transcription factors recognize methylated DNA. Epigenetics 8:131–137
26. Liu Y, Toh H, Sasaki H, Zhang X, Cheng X (2012) An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. Genes Dev 26:2374–2379
27. Liu Y, Olanrewaju YO, Zheng Y, Hashimoto H, Blumenthal RM, Zhang X, Cheng X (2014) Structural basis for Klf4 recognition of methylated DNA. Nucleic Acids Res 42:4859–4867
28. Yin Y, Morgunova E, Jolma A, Kaasinen E, Sahu B, Khund-Sayeed S, Das PK, Kivioja T, Dave K, Zhong F, Nitta KR, Taipale M, Popov A, Ginno PA, Domcke S, Yan J, Schubeler D, Vinson C, Taipale J (2017) Impact of cytosine methylation on DNA binding specificities of human transcription factors. Science 356:eaaj2239
29. Long HK, Blackledge NP, Klose RJ (2013) ZF-CxxC domain-containing proteins, CpG islands and the chromatin connection. Biochem Soc Trans 41:727–740

30. Isalan M, Choo Y (2000) Engineered zinc finger proteins that respond to DNA modification by HaeIII and HhaI methyltransferase enzymes. J Mol Biol 295:471–477
31. Deng D, Yin P, Yan C, Pan X, Gong X, Qi S, Xie T, Mahfouz M, Zhu JK, Yan N, Shi Y (2012) Recognition of methylated DNA by TAL effectors. Cell Res 22:1502–1504
32. Maier JAH, Mohrle R, Jeltsch A (2017) Design of synthetic epigenetic circuits featuring memory effects and reversible switching based on DNA methylation. Nat Commun 8:15336
33. Sander JD, Zaback P, Joung JK, Voytas DF, Dobbs D (2007) Zinc Finger Targeter (ZiFiT): an engineered zinc finger/target site design tool. Nucleic Acids Res 35:W599–W605
34. Sander JD, Maeder ML, Reyon D, Voytas DF, Joung JK, Dobbs D (2010) ZiFiT (Zinc Finger Targeter): an updated zinc finger engineering tool. Nucleic Acids Res 38:W462–W468
35. Wright DA, Thibodeau-Beganny S, Sander JD, Winfrey RJ, Hirsh AS, Eichtinger M, Fu F, Porteus MH, Dobbs D, Voytas DF, Joung JK (2006) Standardized reagents and protocols for engineering zinc finger nucleases by modular assembly. Nat Protoc 1:1637–1652
36. O'Shea EK, Klemm JD, Kim PS, Alber T (1991) X-ray structure of the GCN4 leucine zipper, a two-stranded, parallel coiled coil. Science 254:539–544
37. Harbury PB, Zhang T, Kim PS, Alber T (1993) A switch between two-, three-, and four-stranded coiled coils in GCN4 leucine zipper mutants. Science 262:1401–1407
38. Oehler S, Eismann ER, Kramer H, Muller-Hill B (1990) The three operators of the lac operon cooperate in repression. EMBO J 9:973–979
39. Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys_2His_2 zinc finger proteins. Annu Rev Biophys Biomol Struct 29:183–212
40. Pabo CO, Peisach E, Grant RA (2001) Design and selection of novel Cys_2His_2 zinc finger proteins. Annu Rev Biochem 70:313–340
41. Segal DJ, Barbas CF 3rd (2001) Custom DNA-binding proteins come of age: polydactyl zinc-finger proteins. Curr Opin Biotechnol 12:632–637
42. Beerli RR, Barbas CF 3rd (2002) Engineering polydactyl zinc-finger transcription factors. Nat Biotechnol 20:135–141
43. Jamieson AC, Miller JC, Pabo CO (2003) Drug discovery with engineered zinc-finger proteins. Nat Rev Drug Discov 2:361–368
44. Deltcheva E, Chylinski K, Sharma CM, Gonzales K, Chao Y, Pirzada ZA, Eckert MR, Vogel J, Charpentier E (2011) CRISPR RNA maturation by trans-encoded small RNA and host factor RNase III. Nature 471:602–607
45. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821
46. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32:347–355
47. Wang H, La Russa M, Qi LS (2016) CRISPR/Cas9 in genome editing and beyond. Annu Rev Biochem 85:227–264
48. Kungulovski G, Jeltsch A (2016) Epigenome editing: state of the art, concepts, and perspectives. Trends Genet 32:101–113
49. Lee S, Doddapaneni K, Hogue A, McGhee L, Meyers S, Wu Z (2010) Solution structure of Gfi-1 zinc domain bound to consensus DNA. J Mol Biol 397:1055–1066
50. Sera T, Uranga C (2002) Rational design of artificial zinc-finger proteins using a nondegenerate recognition code table. Biochemistry 41:7074–7081

Chapter 4

Enhanced Manipulation of Human Mitochondrial DNA Heteroplasmy In Vitro Using Tunable mtZFN Technology

Payam A. Gammage and Michal Minczuk

Abstract

As a platform capable of mtDNA heteroplasmy manipulation, mitochondrially targeted zinc-finger nuclease (mtZFN) technology holds significant potential for the future of mitochondrial genome engineering, in both laboratory and clinic. Recent work highlights the importance of finely controlled mtZFN levels in mitochondria, permitting far greater mtDNA heteroplasmy modification efficiencies than observed in early applications. An initial approach, differential fluorescence-activated cell sorting (dFACS), allowing selection of transfected cells expressing various levels of mtZFN, demonstrated improved heteroplasmy modification. A further, key optimization has been the use of an engineered hammerhead ribozyme as a means for dynamic regulation of mtZFN expression, which has allowed the development of a unique isogenic cellular model of mitochondrial dysfunction arising from mutations in mtDNA, known as mTUNE. Protocols detailing these transformative optimizations are described in this chapter.

Key words Genetic engineering, mtZFN, Mitochondrial disease, mtDNA, Heteroplasmy, Gene therapy, mTUNE

1 Introduction

Human mitochondrial DNA (mtDNA) is a ~16.5 kb, circular, multi-copy molecule encoding structural components of the respiratory chain, ATP synthase, and all RNA components required for translation by mitochondrial ribosomes [1]. The mitochondrial genome is found within the mitochondrial matrix, tightly associated with the inner mitochondrial membrane, packaged into nucleoids by the mitochondrial transcription factor A (TFAM). In human, there are typically ~100–10,000 copies of mtDNA per cell, varying significantly between cell types. As a multi-copy genome, accumulation of mutations in mtDNA can lead to a phenomenon termed heteroplasmy, where multiple mtDNA variants (i.e., wild type and mutant) can coexist. Such heteroplasmic mtDNA mutations account for a substantial proportion of primary mtDNA disease in humans [2]. The hetero-

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_4,

plasmic nature of such pathogenic mtDNA mutations permits a unique approach to gene therapy, relying on the selective recognition and degradation of mutant mtDNA molecules based on mtDNA sequence discrimination and introduction of a double-strand break (DSB). In so doing, mutant mtDNA copy number is specifically diminished, and replication of the spared wild-type mtDNA molecules will result in the cell recovering to initial mtDNA copy number with altered heteroplasmy. The basis for sequence-specific targeting, DNA break introduction, and heteroplasmy shifting of mtDNA has been demonstrated through the use of mitochondrially targeted restriction enzymes [3–8], mitochondrially targeted transcription activator-like effector nucleases (mitoTALENs) [9–12], and mitochondrially targeted zinc-finger nucleases (mtZFN) [13–17].

When manipulating heteroplasmy of mtDNA bearing a single-point mutation, a mtZFN pair is assembled: one mutation-specific monomer that binds mtDNA including the mutation, and a mutation-nonspecific monomer that binds a proximal mtDNA sequence on the opposite strand, contained within both wild-type and mutant genomes (Fig. 1a). Cleavage of mtDNA requires dimerization of the nuclease domains, which only occurs when both mtZFNs are bound to the mutant molecule.

In previous work, we have described the development and application of an effective mtZFN architecture [15, 17]. More recently, we have identified the importance of fine-tuning mtZFN expression levels to avoid off-target effects, likely both at the target site and at other loci in mtDNA, which would otherwise produce undesired, nonspecific depletion of mtDNA copy number [16]. In this chapter, we describe two methods for controlling expression levels of transiently transfected mtZFNs: (1) by means of differential fluorescence-activated cell sorting (dFACS) and (2) by means of an engineered hammerhead ribozyme (HHR) incorporating an inhibitory, tetracycline-binding aptamer, named 3K19 [18] (Fig. 1b, c). In addition to significantly enhanced mtZFN efficiency, the capacity to produce a range of distinct mtZFN expression levels in a transient system allows for the generation of cell lines with variable heteroplasmy from an initial heteroplasmic cell line [16]. Use of this method largely circumvents issues of clonal variation observed for cytoplasmic hybrid (cybrid) cells [19] and analysis of cells bearing variable levels of m.8993T>G generated using these methods, known as mTUNE cells, has allowed identification of previously unknown mechanisms of metabolic rewiring in response to physiological mitochondrial dysfunction [20].

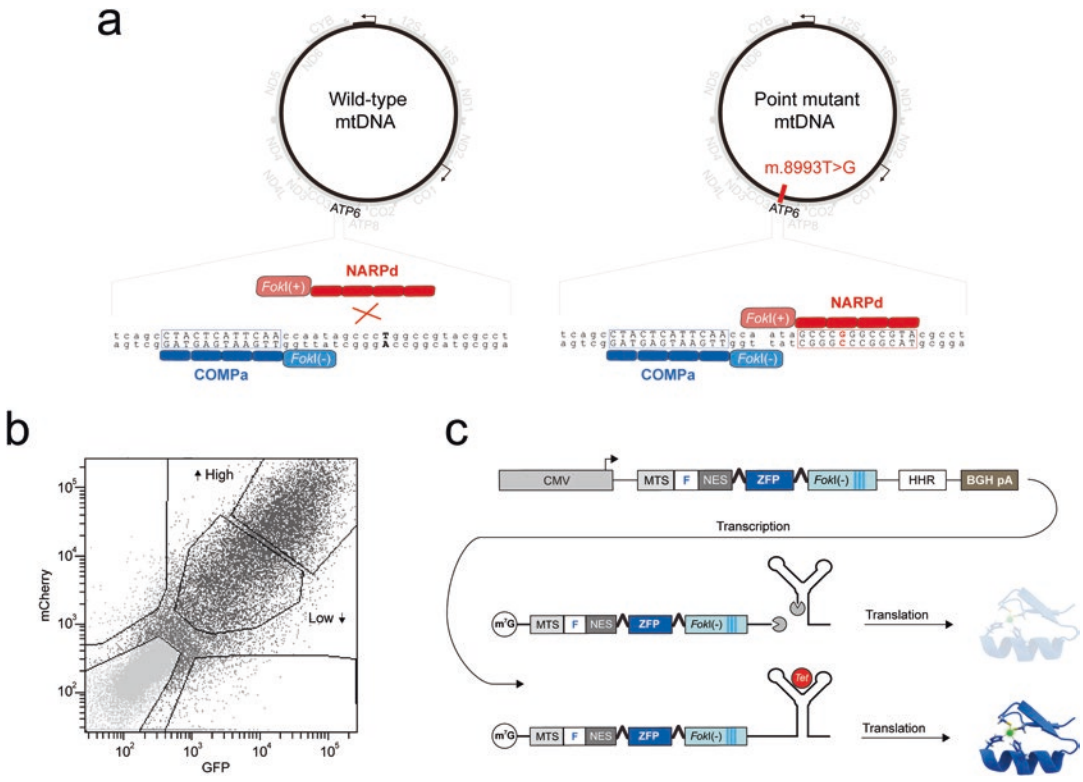


Fig. 1 Fine-tuned mitochondrially targeted ZFNs for manipulation of mtDNA heteroplasmy. (**a**) A general strategy for targeting mtDNA point mutations using mtZFNs, contextualized using the m.8993T>G mutation. The mutation-specific monomer (red) recognizes the mutation site and dimerizes with the companion monomer (blue), which is bound to mtDNA on the opposite strand. Dimerization leads to specific introduction of a DNA double-strand break, followed by degradation of the mutant haplotype only. (**b**) An example of a flow cytometry dot plot, where mCherry/GFP double-positive events have been separated into two distinct populations, based on fluorescence intensity, which should be correlated with mtZFN expression levels. (**c**) Schematic of hammerhead ribozyme (HHR) incorporation into a mtZFN transgene. The HHR is placed downstream of the stop codon, before the polyadenylation signal (BGH pA). Once transcribed, the HHR either can be stabilized by binding of tetracycline or can initiate self-directed cleavage of the transcript, leading to degradation by endogenous RNases. The extent to which these two processes compete determines the resulting protein expression level of transiently expressed mtZFNs. Figure adapted from [16]

It is assumed that, before attempting to enact this protocol, a cybrid cell line bearing a mtDNA mutation of interest and a corresponding library of ZFPs specific to this mutation have been acquired. Published protocols are available for design of ZFPs [21–27] and creation of cybrids [28, 29]. As an example for this protocol we use cells bearing the m.8993T>G point mutation and ZFPs specific to this mutation.

2 Materials

2.1 Molecular Cloning

pcDNA3.1(−)_mCherry_3K19 vector (available from Addgene, 104499).

pTracer CMV/bsd vector (Thermo Fisher Scientific, cat. no. V883-20).

KOD DNA polymerase.

Nuclease-free water.

Restriction endonucleases and buffers.

T4 DNA ligase.

Gel extraction kit.

PCR purification kit.

Miniprep kit.

DNA-loading dye.

UltraPure agarose.

SYBR Safe DNA gel stain.

DNA ladder.

1× TBE electrophoresis buffer: 90 mM Tris, 90 mM boric acid, pH 8, 2 mM EDTA.

DH5α competent cells.

TYE bacterial medium.

Ampicillin.

Midiprep kit.

2.2 Tissue Culture, Transfection, and FACS

Tissue culture plastics:

- 75 cm^2 flask.
- 6-Well plate.
- 15 mL Conical tubes.

5 mL Polystyrene round-bottom tube.

CellTrics cell suspension filter.

1× DMEM, supplemented with 4.5 g/L D-glucose, Glutamax, and 100 mg/L sodium pyruvate.

1× PBS, without $CaCl_2$ or $MgCl_2$.

10× Trypsin-EDTA solution.

Fetal calf serum.

Tetracycline hydrochloride, prepared at 1000× concentration in MilliQ filtered water (*see* **Note 1**).

Opti-MEM, reduced serum media.

Table 1
Antibodies required for western blotting

Epitope	Source (Cat. No.)	Dilution
Primary antibodies		
HA	Roche (1186743100)	1:500
FLAG	Sigma (F3165)	1:2000
β-Actin	Sigma (A2228)	1:100,000
Secondary antibodies		
Goat anti-mouse IgG HRP	Promega (W402B)	1:2000
Goat anti-rat IgG HRP	Santa Cruz (SC2065)	1:1000

Lipofectamine 2000.

FACS-capable flow cytometer.

2.3 Protein Extraction and Western Blotting

BCA protein assay kit.

4× Western sample buffer.

DTT.

Bolt™ 4–12% Bis-Tris polyacrylamide gels.

20× MES SDS running buffer.

iBlot 2™ dry transfer cell.

iBlot 2™ nitrocellulose transfer stack.

Coomassie brilliant blue stain.

Nonfat milk powder.

Antibodies (Table 1).

1× Western lysis buffer: 50 mM Tris-acetate, pH 7.4, 150 mM NaCl, 1 mM EDTA, 1% (w/v) Triton X-100, 1× proteinase inhibitor cocktail.

1× Western incubation buffer: 1× PBS, 5% (w/v) nonfat milk powder.

1× Western washing buffer: 1× PBS, 0.1% (w/v) Triton X-100.

ECL detection reagents.

Empty gel cassettes 1.0 mm.

X-ray film.

XCell SureLock™ mini-cell.

2.4 mtDNA Heteroplasmy Measurements

KOD DNA polymerase.

Nuclease-free water.

Radiolabeled dCTP [α-^{32}P]—3000 Ci/mmol^{-1}, 10 mCi/mL^{-1}, 250 μCi.

*Sma*I restriction endonuclease.

UltraPure agarose.

1× TBE electrophoresis buffer.

DE81 ion-exchange paper.

Gel dryer.

PhosphorImager or X-ray film with developing system.

DNA extraction spin column kit.

2.5 General Laboratory Consumables and Equipment

0.5 mL Tubes.

1.5 mL Tubes.

Standard PCR thermal cycler.

DNA mini gel tanks.

Power pack.

Spectrophotometer.

Tube roller.

Tilt platform.

Benchtop centrifuge.

UV transilluminator.

Gel-imaging system.

2.6 Noncommercial Biological Materials

Human osteosarcoma line 143B, NARP m.8993T>G cybrid [30] (kindly donated by Prof. Eric Schon, Columbia University, NYC).

3 Methods

Carry out all procedures at room temperature unless otherwise stated. All solutions are prepared using ultrapure (deionized, purified) water and analytical grade reagents, and can be stored at room temperature unless otherwise indicated.

3.1 Incorporation of HHR into MtZFN Monomers and Expression Vector Cloning

Using pcDNA3.1(−)_mtZFN(+)^{2G}/pcDNA3.1(−)_mtZFN(−)^{2G} constructs available from Addgene (Nos. 62798 and 62799), the pre-generated library of paired ZFPs must be cloned into their respective +/− backbones, as described previously [17]. The resulting constructs must then be assembled into pcDNA3.1(−)_mCherry_3K19 (Addgene No. 104499), which contains the 3K19 HHR, and then subcloned into appropriate expression vectors by restriction cloning.

1. Using 5′ *Xba*I and 3′ *Kpn*I restriction sites, digest 2 μg of pcDNA3.1(−)_mtZFN(+/−)^{2G} and pcDNA3.1(−)_mCherry_3K19, following the manufacturer's recommendations.
2. Take the entire reaction, resolve on a 0.8% agarose gel, and run at 7 V/cm for 1 h. Visualize DNA using a UV light box, excise appropriate bands using a scalpel, and then purify DNA from the gel using a gel extraction kit.
3. Ligate mtZFN(+/−) fragments into pcDNA3.1(−)_mCherry_3K19 backbone using T4 DNA ligase, following the manufacturer's instructions (*see* **Note 1**). Take 5 μL of the ligation reaction, incubate with DH5α competent cells, transform using the heat-shock method, plate cells on dishes of TYE media supplemented with 50 μg/mL ampicillin, and incubate at 37 °C overnight. Select colonies and amplify in TYE broth supplemented with 50 μg/mL ampicillin at 37 °C overnight in an incubator fitted with shaking platform. After the overnight incubation, isolate plasmid DNA using a miniprep kit, according to the manufacturer's instructions.
4. Repeat *Xba*I-*Kpn*I double digestion and agarose gel electrophoresis from **step 2** using isolated plasmid DNA from the end of **step 3** to confirm the presence of mtZFN(+/−) within the pcDNA3.1(−)_mCherry_3K19 backbone, and then confirm by Sanger sequencing using T7 forward and BGH reverse priming sites located adjacently to the multiple cloning site (MCS) of pcDNA3.1(−).

   Once incorporation of mtZFN(+/−) into pcDNA3.1(−)_mCherry_3K19 is confirmed, the entire MCS of either mtZFN(+)_HHR or mtZFN(−)_HHR be re-cloned into a vector capable of co-expressing mtZFNs and GFP simultaneously in mammalian cells. The vector used in our work is pTracer CMV/Bsd (pTracer). The MCS of pTracer is flanked at the 5′ and 3′ ends by *Pme*I restriction sites, as is the MCS contained within the original pcDNA3.1(−)_mCherry_3K19 vector. There is no apparent benefit to mtZFN(+/−) being co-expressed specifically with mCherry or GFP, though in the following examples mtZFN(+) is cloned into pcmCherry and mtZFN(−) is cloned into pTracer.
5. To re-clone the mtZFN(+)_HHR construct into pTracer, digest 2 μg of pTracer and mtZFN(+)_HHR with 10 U of *Pme*I, according to the manufacturer's instructions, treat briefly with alkaline phosphatase, then resolve the digested fragments on a 0.8% agarose gel run at 7 V/cm for 1 h, and cut gel slices containing relevant bands using a scalpel. Repeat gel extraction, ligation, transformation, digestion screening, and sequencing confirmation as laid out in **steps 2** and **3** (*see* **Note 2**). Once plasmids of correct composition and orientation have been

obtained, produce a concentrated plasmid DNA stock for all plasmids, including empty vectors, using a midiprep kit.

3.2 Assessment of Tunable MtZFN Expression in Heteroplasmic Cells

In order to assess the effects of variable mtZFN expression levels on mtDNA heteroplasmy, a population of cells transiently expressing mtZFN monomers and/or empty vectors must be obtained. Transient expression of mtZFNs in cells allows for permanent mtDNA heteroplasmy modification, without the need for genomic integration of transgenes/selection cassettes and prolonged culture, and can be beneficial for work in cell cultures. Our preferred method to this end is fluorescence-activated cell sorting (FACS), followed by further assessments of mtZFN expression and activity. Throughout the protocols described below, cultured cells are maintained in standard conditions (37 °C, 5% CO_2).

3.2.1 Transfection of mtZFN DNA Constructs

1. On the day prior to transfection, seed cells at a density of 2×10^6 cells per 75 cm^2 flask and leave overnight to attach. On the day of transfection, aspirate the original media, pipette fresh media onto plates, and transfect cells with a total of 25 μg of plasmid DNA (12.5 μg pcmCherry_mtZFN(+) with or without HHR, 12.5 μg pTracer_mtZFN(−) with or without HHR, or 12.5 μg of either empty vector) using Lipofectamine 2000, according to the manufacturer's instructions (*see* **Note 3**).
2. After 6 h, aspirate media containing transfection reagents and replace with fresh media (*see* **Note 4**). If using HHR, this medium should be supplemented with desired concentrations of tetracycline (*see* **Note 5**). With regard to data from [18], we use 25 mM (11 μg/mL) and 250 mM (110 μg/mL).
3. At 24 h post-transfection, wash cells with 1× PBS, trypsinize, and then collect cells in 15 mL Falcon tubes by centrifugation at 300 *g* for 5 min. Resuspend the cell pellet in 750 μL 1× PBS and transfer to 5 mL round-bottomed tube by passing through a CellTrics strainer. It is essential that all samples are kept at 4 °C from this point (*see* **Note 6**).

3.2.2 Fluorescence-Activated Cell Sorting (FACS)

1. Load untransfected cells into the FACS sorter. After selecting monodisperse events, based on forward and side scatter profile, set fluorescence gating using the untransfected and GFP/mCherry-only cells, and then sort experimental samples, collecting double-positive cells (*see* **Note 7**). If attempting to manipulate mtZFN dosage by dFACS, the double-positive population should be sorted in separate groupings, giving "high-expression" and "low-expression" conditions (Fig. 1b),

with downstream analyses performed as described in Subheading 3.2.3.

Samples should be sorted in a sterile environment and collected in sterile tissue culture media in autoclaved 1.5 mL tubes, which are immediately placed on ice once obtained.

3.2.3 Analysis of Tuned mtZFN Activity in Cells

Samples acquired by FACS should be divided into two groups at a ratio of 1:1. These groups are (1) cells to provide a DNA sample, and (2) cells to provide a protein sample, allowing confirmation of mtZFN expression as manipulated by HHR and tetracycline additions. It is expected that at least 5×10^5 cells are collected per sample. If this is not achieved, modify the volumes given in **steps 1** and **2** in Subheading 3.2.3 appropriately.

1. With group (1) cells, intended for DNA analyses, pellet cells in a 1.5 mL tube by centrifugation at 300 *g* for 5 min. Aspirate media and extract total DNA using a DNA extraction spin column kit, according to the manufacturer's recommendation. Elute DNA in a small volume of nuclease-free water, typically ~30–50 μL, and measure DNA concentration by nanodrop.
2. With group (2) cells, intended for protein analyses, pellet an identical number of cells per sample in separate 1.5 mL tubes by centrifugation at 300 *g* for 10 min (*see* **Note 8**). Aspirate media and resuspend pellet in 60 μL western lysis buffer in a 0.5 mL tube. Incubate on ice for 10 min, then add 22.5 μL 4× western sample buffer and 4.5 μL 1 M DTT, mix thoroughly by vortex, and incubate the sample at 95 °C for 5 min. At this point the sample can either be stored for up to a year at −20 °C or analyzed by western blotting, as detailed in Subheading 3.2.5.

3.2.4 Detection of mtZFN Expression

1. To analyze protein extracts for variable mtZFN expression by western blotting, load ~20 μL of extract from group (1), and appropriate protein molecular weight markers, onto a 10-well bolt SDS 4–12% Bis-Tris gel and run at 180 V in 1× MES buffer for ~30 min.
2. When resolved, transfer proteins from the gel to a nitrocellulose membrane using iBlot 2 transfer stack and transfer cell. Once proteins have been electrotransferred, wash the membrane briefly in 1× PBS and then block in 1× PBS with 5% nonfat milk (w/v) (PBS-M) for 1 h at room temperature on a roller. Incubate the gel in Coomassie brilliant blue stain for several hours, and allow to destain in water for several further hours; this allows fast, reliable confirmation of equal protein loading and efficiency of electrotransfer. Image the gel using gel documentation system, and then discard.

3. Once the membrane has been blocked, remove blocking solution, apply primary antibody in 1× PBS-M, and incubate overnight on a roller at 4 °C. On the following day remove the primary antibody solution, wash the membrane three times with 1× PBS with 0.1% Triton X-100 (w/v) (PBS-T) for 10 min on a roller at room temperature, and then incubate with the secondary antibody in 1× PBS-M on a roller for 1 h at room temperature. Remove the secondary antibody solution, wash the membrane three times with 1× PBS-T for 10 min, then develop using ECL reagents, and expose to X-ray film in a darkroom. A list of all primary and secondary antibodies used for western blotting is given in Table 1.

Examples of western blot data obtained using dFACS and HHR in different tetracycline concentrations are presented (Fig. 2a, c).

3.2.5 Assessing the Effects of Tuned mtZFN Treatment on Heteroplasmy

Measurements of mtDNA heteroplasmy can be made by various means. Often the best method for such analyses will change, based on the variant studied and available techniques/resources. Measurements of heteroplasmy, in our hands, typically consist of restriction fragment length polymorphism (RFLP), Southern blotting, ultra-deep amplicon resequencing, and pyrosequencing analyses, though numerous further methods are available. Changes in heteroplasmy of the m.8993T>G mutation, as manipulated by dFACS- or HHR-tuned mtZFN treatments, are measured by PCR amplification of a portion of mtDNA containing the mutation, with incorporation of radiolabeled nucleotides in a final PCR cycle to negate heteroduplex formation. The radiolabeled PCR products are then purified and analyzed by RFLP, based on a unique *Sma*I/*Xma*I restriction site created by the m.8993T>G mutation. This approach to measurement of m.8993T>G has been described in detail previously [17]. Briefly, the procedure is as follows:

1. Prepare 50 μL PCR reactions with 100 ng DNA template obtained from group (2). Prepare three additional reaction volumes: two for 143B wild-type and N100 digest control conditions (*see* **Note 9**), and one further to store on ice. Adjust cycling conditions to repeat the PCR program for a total of 15 cycles.
2. Following completion of the 15-cycle PCR, take the single additional reaction volume, store on ice, add 2 μL dCTP [α-^{32}P] mixing well, then pipette 5 μL of this spiked mixture into each PCR tube, mix thoroughly using the pipette tip, and then place tubes in a thermal cycler for a final PCR cycle. Once

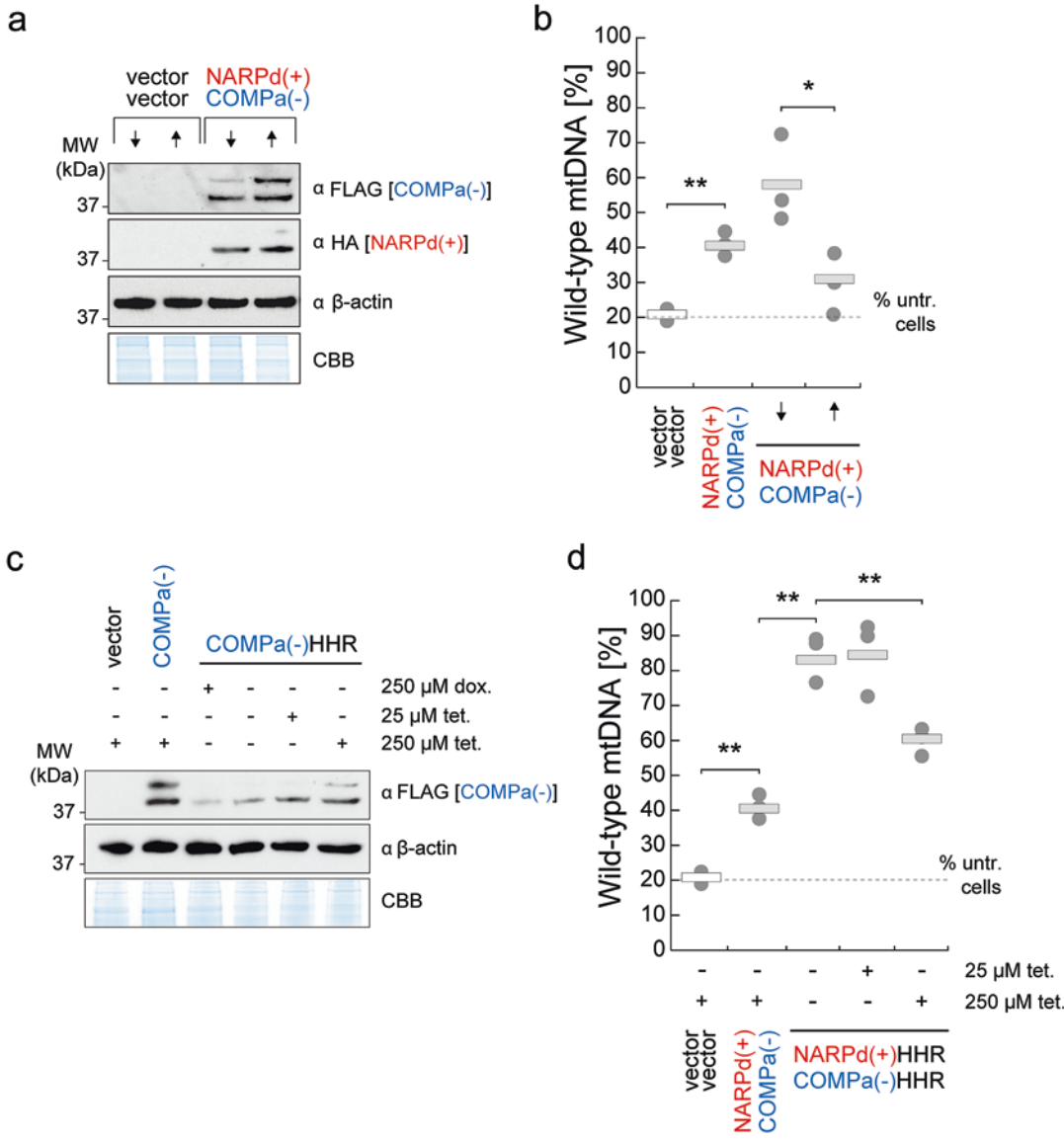


Fig. 2 Comparison of differential FACS and HHR approaches for tunable mtZFN activity on mtDNA heteroplasmy. (**a**) Western blot analysis of mtZFN-expressing cells, and controls, obtained through dFACS at 24 h post-transfection. Arrows indicate "high" and "low" conditions. *HA* hemagglutinin, *FLAG* FLAG epitope. β-Actin and Coomassie brilliant blue (CBB) stain are shown as loading controls. (**b**) Quantification of wild-type heteroplasmy following dFACS treatment with mtZFN. Densitometry of RFLPs from three biological replicates is shown. $*p < 0.05$. $**p < 0.01$. (**c**) Western blot analysis of mtZFN-expressing cells, and controls, obtained by FACS under HHR control at 24 h post-transfection. Tunable, tetracycline-dependent expression of mtZFNs can be achieved. Dox doxycycline. *Tet* tetracycline. *FLAG* FLAG epitope. β-Actin and Coomassie brilliant blue (CBB) stain are shown as loading controls. (**d**) Quantification of wild-type heteroplasmy following HHR-controlled treatment with mtZFN. Densitometry of RFLPs from three biological replicates is shown. $**p < 0.01$. Figures and data adapted from [16]

this is completed, purify PCR products using a PCR purification kit. Elute PCR products in 50 μL nuclease-free water.

3. Prepare 10 μL *Sma*I digest reactions for all samples, including digestion controls. Allow reaction to proceed for 2 h at 25 °C, and then resolve on a 1% agarose gel run at 7 V/cm for 50 min. Dry the gel onto DE81 paper using gel dryer. Expose the dried gel to a phosphor screen overnight and image using the PhosphorImager.

In the example data shown (Fig. 2b, d), 143B cybrid cells bearing m.8993T>G targeted by NARPd(+) and COMPa(−) and lower expression levels of mtZFNs, by either dFACS or HHR methods, lead to more efficient shifts in heteroplasmy. Conversely, higher expression levels of mtZFNs result in less efficient shifts in heteroplasmy, which have also been correlated with copy number depletions in our recent work [16]. Importantly, using the HHR method, control of expression levels can be used to generate a range of heteroplasmies from a single cell line, without clonal selection (Fig. 2d).

The relative mtDNA heteroplasmy modification efficiency of either dFACS- or HHR-tuned mtZFN treatments is dependent on, among other factors, the steady-state copy number of the cells being used and the degree of mtZFN expression exerted (Hoitzing et al., submitted. Pre-print available at https://doi.org/10.1101/145292). It is, therefore, likely that varying results will be achieved across different cellular contexts, and that some applications of mtDNA editing technologies will require further optimization in more challenging circumstances (e.g., cell lines with very low mtDNA copy number). However, the modifications to previous mtZFN protocols [17], detailed in this chapter, should be capable of significantly enhancing mtZFN activity in most heteroplasmic cell lines.

4 Notes

1. T4 DNA ligase is sensitive to guanidinium thiocyanate, contained in buffer QG from the QIAquick gel extraction kit. To limit the impact of this sensitivity on ligation/transformation efficiency, ensure that the total volume of gel-extracted material in ligations is not >20% of the final volume.
2. Digestion with *Pme*I produces blunt ends, and half of the successful ligations will result in insertion of transgenes in the

opposite orientation. Select several clones to ensure a properly ligated product.

3. In order to set gating parameters, untransfected cells and single-fluorescence cells need to be prepared alongside experimental samples transfected with both GFP- and mCherry-expressing plasmids. Without these "single-stained" controls, it is not possible to ensure homogeneity of the sample.
4. Lipofectamine and similar reagents can have cytotoxic effects if left to incubate with cells for extended periods. Although potentially marginal in most cases, this step could be important for some cell types.
5. Tetracyclines are well-known inhibitors of bacterial translation, and have similar activity on mitochondrial translation at relatively low concentrations [31]. As such tetracycline is only ever present for ~18 h in these experiments. Additionally, tetracycline is exquisitely light sensitive, and steps to protect stocks/media containing tetracycline from exposure to light should be taken.
6. If cells are likely to be left for >30 min while other samples are sorted, survival of cells sorted, particularly for continued culture, is greatly enhanced by keeping samples at 4 °C prior to reintroduction to standard tissue culture conditions. If <30-min wait is required, sorting cells directly into a tissue culture dish without storing at 4 °C could be advantageous also.
7. In our FACS experiments, GFP signal is excited by a 488 nm laser and emission detected by photomultiplier tube (PMT) with 530/30 band-pass and 502 long-pass filters. mCherry signal is excited with 561 nm laser and emission detected by PMT 610/20 band-pass and 600 long-pass filters.
8. Use of cell number as a proxy for BCA acquired protein concentration values prior to western blotting can be helpful, especially when handling small samples.
9. These digest controls are essential for production of a reliably interpretable RFLP analysis, and can consist of either cellular DNA extracts or plasmids containing the regions of interest to be used as template for PCR.

Acknowledgments

This work was supported by the Medical Research Council, UK (MC_U105697135).

References

1. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F et al (1981) Nature 290:457–465
2. Gorman GS, Chinnery PF, DiMauro S, Hirano M, Koga Y, McFarland R, Suomalainen A, Thorburn DR, Zeviani M, Turnbull DM (2016) Nat Rev Dis Primers 2:16080
3. Srivastava S, Moraes CT (2001) Hum Mol Genet 10:3093–3099
4. Tanaka M, Borgeld HJ, Zhang J, Muramatsu S, Gong JS, Yoneda M, Maruyama W, Naoi M, Ibi T, Sahashi K et al (2002) J Biomed Sci 9:534–541
5. Bayona-Bafaluy MP, Blits B, Battersby BJ, Shoubridge EA, Moraes CT (2005) Proc Natl Acad Sci U S A 102:14392–14397
6. Bacman SR, Williams SL, Duan D, Moraes CT (2012) Gene Ther 19:1101–1106
7. Bacman SR, Williams SL, Garcia S, Moraes CT (2010) Gene Ther 17:713–720
8. Bacman SR, Williams SL, Hernandez D, Moraes CT (2007) Gene Ther 14:1309–1318
9. Hashimoto M, Bacman SR, Peralta S, Falk MJ, Chomyn A, Chan DC, Williams SL, Moraes CT (2015) Mol Ther 23:1592–1599
10. Bacman S, Williams SL, Pinto M, Peralta S, Moraes CT (2013) Nat Med 19:1111–1113
11. Reddy P, Ocampo A, Suzuki K, Luo J, Bacman SR, Williams SL, Sugawara A, Okamura D, Tsunekawa Y, Wu J et al (2015) Cell 161:459–469
12. Phillips AF, Millet AR, Tigano M, Dubois SM, Crimmins H, Babin L, Charpentier M, Piganeau M, Brunet E, Sfeir A (2017) Mol Cell 65:527–538.e6
13. Minczuk M, Papworth MA, Miller JC, Murphy MP, Klug A (2008) Nucleic Acids Res 36:3926–3938
14. Minczuk M, Kolasinska-Zwierz P, Murphy MP, Papworth MA (2010) Nat Protoc 5:342–356
15. Gammage PA, Rorbach J, Vincent AI, Rebar EJ, Minczuk M (2014) EMBO Mol Med 6:458–466
16. Gammage PA, Gaude E, Van Haute L, Rebelo-Guiomar P, Jackson CB, Rorbach J, Pekalski ML, Robinson AJ, Charpentier M, Concordet JP et al (2016) Nucleic Acids Res 44:7804–7816
17. Gammage PA, Van Haute L, Minczuk M (2016) Methods Mol Biol 1351:145–462
18. Beilstein K, Wittmann A, Grez M, Suess B (2015) ACS Synth Biol 4:526–534
19. Hao H, Morrison LE, Moraes CT (1999) Hum Mol Genet 8:1117–1124
20. Gaude E et al (2018) Mol Cell 69:581–593 PMID:29452638
21. Wright DA, Thibodeau-Beganny S, Sander JD, Winfrey RJ, Hirsh AS, Eichtinger M, Fu F, Porteus MH, Dobbs D, Voytas DF et al (2006) Nat Protoc 1:1637–1652. PMID: 29452638
22. Carroll D, Morton JJ, Beumer KJ, Segal DJ (2006) Nat Protoc 1:1329–1341
23. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA et al (2008) Mol Cell 31:294–301
24. Kim S, Lee MJ, Kim H, Kang M, Kim JS (2011) Nat Methods 8:7
25. Sander JD, Zaback P, Joung JK, Voytas DF, Dobbs D (2007) Nucleic Acids Res 35:W599–W605
26. Fu F, Sander JD, Maeder M, Thibodeau-Beganny S, Joung JK, Dobbs D, Miller L, Voytas DF (2009) Nucleic Acids Res 37:D279–D283
27. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y et al (2011) Nat Methods 8:67–69
28. King MP, Attardi G (1989) Science 246:500–503
29. King MP, Attadi G (1996) Methods Enzymol 264:313–334
30. Manfredi G, Gupta N, Vazquez-Memije ME, Sadlock JE, Spinazzola A, De Vivo DC, Schon EA (1999) J Biol Chem 274:9386–9391
31. Moullan N, Mouchiroud L, Wang X, Ryu D, Williams EG, Mottis A, Jovaisaite V, Frochaux MV, Quiros PM, Deplancke B et al (2015) Cell Rep 10:1681–1691

Chapter 5

Engineering RNA-Binding Proteins by Modular Assembly of RanBP2-Type Zinc Fingers

Simona De Franco, Mitchell R. O'Connell, and Marylène Vandevenne

Abstract

Deciphering the function of the nonprotein-coding portion of genomes represents one of the major challenges that molecular biology is facing today. Numerous classes of RNAs have been discovered over the last past decade and appear to play important regulatory roles in gene expression and disease. The ability to study and manipulate these RNAs relies on the development of programmable RNA-binding molecules such as RNA-binding proteins. Most RNA-binding proteins have modular architectures and combine different RNA-binding domains that provide binding affinity toward a specific RNA sequence and/or structure. Herein, we describe a general strategy to design single-stranded RNA-binding proteins using RanBP2-type zinc-finger (ZF) domains that can recognize a given RNA sequence with high affinity and specificity.

Key words Zinc finger, Protein design, RNA-binding protein, Modular assembly

1 Introduction

RNA-binding proteins (RBP) are involved in all aspects of RNA biology from RNA processing after transcription to RNA transport and localization, and the association and dissociation of these proteins to their target RNA regulate their stability, localization, and function in the cell [1, 2]. Given their important roles in gene regulation, the scientific community has significant interest in investigating the way these proteins interact with RNA and, in particular, what determines their specificity for RNA targets from a highly sequenced and structurally diverse cellular pool of RNA. Most of these proteins use a modular domain arrangement and often combine RNA-binding domain(s) for target recognition with other functional modules that affect the function of the targeted RNA [3].

Furthermore, most RBPs also recognize their target(s) using a modular recognition mode. Most RBPs utilize several RNA-binding domains (RBDs) from the same or different families, which are arranged in numerous ways to create a large repertoire

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_5,

of possible sequence specificities to recognize a large range of different RNA targets. Common RBDs that target single-stranded RNA (ssRNA) include K homology domains (KH domains) [4], RNA recognition motifs (RRMs) [5], Pumilio/FBF repeats (PUFs) [6], pentatricopeptide repeats (PPRs) [7], cold-shock domains (CSDs) [8], and zinc fingers (ZFs) [9].

It is also important to mention that some RRMs and ZFs were also shown to recognize double-stranded RNAs (dsRNAs) [10, 11] together with sterile alpha motifs (SAMs) and Z-RNA-binding domains. In the latter case, the recognition is largely sequence independent [12], whereas interactions with ssRNAs occur, in most cases, in a sequence-specific manner [13].

Over the last past decade, many scientists have attempted to monitor and control RNA functions in living cells using designer RNA-binding proteins (RBPs) that can specifically deliver an effector domain to a given RNA sequence (Fig. 1). These engineered proteins represent extremely valuable tools for research and potentially medical purposes.

One of the first powerful and efficient strategies to target RNA is a system based on the natural interaction of the MS2 bacteriophage coat protein with a RNA stem-loop structure from the phage genome [14]. This strategy consisted of inserting this exogenous RNA stem-loop structure in the targeted RNA [14], so that the corresponding MS2 RNA-binding protein fused to a functional module can specifically recognize the MS2 stem-loop-labeled RNA and deliver the effector domain to the targeted RNA, to observe or alter its function. These RBPs were fused to various effector domains and used to track mRNA localization [15, 16] or to activate the translation of a reporter gene [17]. However, this approach requires the preexistence or insertion of protein-binding sequences in the targeted RNAs, which limits their application to RNAs containing MS2-binding sites.

More recently, very promising data was obtained using PUF proteins to engineer designer sequence-specific RNA-binding proteins. These proteins include eight repeats and recognize with high affinity and specificity a 8-nt-long ssRNA [6]. The success of these proteins for ssRBP engineering is mostly explained because a complete and simple recognition code has been established for these proteins. In this recognition code, each of the four RNA bases is recognized by only two residues exposed on each PUF repeat so that the sequence specificity can be easily customized by changing the identity of these residues [18]. Engineered PUF proteins were combined with various functional domains and successfully used to alter RNA splicing [19], generate sequence-specific RNA endonucleases [20], and modulate the stability and translation of mRNAs in living cells [21].

However, despite this promising work, various complications have been reported for PUF proteins. First, variations in their scaffold allow different binding modes that result in tolerance for alter-

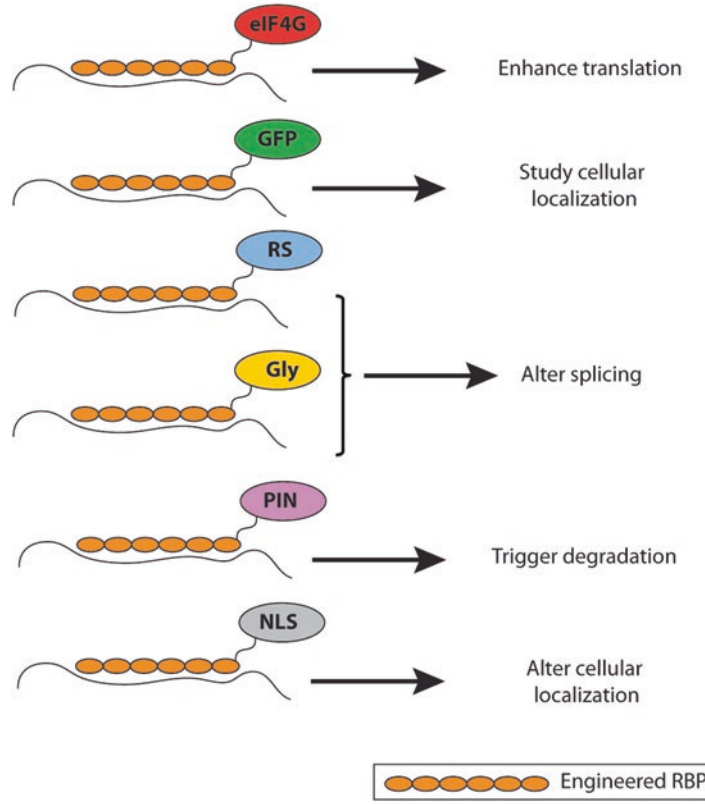


Fig. 1 Schematic representation of the different possible uses of engineered RBPs in combination with various effector domains. The effector domains include (1) eIF4G, an initiator of protein translation; (2) GFP, the green fluorescent protein to study cellular localization; (3) RS or Gly: arginine/serine- or glycine-rich domains that modulate RNA splicing; (4) PIN domain: a nonspecific endonuclease domain that catalyzes RNA degradation; (5) NLS: a nuclear localization signal that promotes transfer to the nucleus

native bases by the different repeats, or the ability to accommodate additional bases that flip away from the protein surface [22–25]. Secondly, the concave shape adopted by PUF proteins limits the number of repeats and therefore the length of the recognized sequence. For these reasons, these proteins exhibit high probability for off-target binding.

More recently, exciting programmable RNA-targeting capabilities have been demonstrated using CRISPR-Cas systems. In 2014, CRISPR-Cas9, an RNA-guided endonuclease, currently widely used for sequence-specific dsDNA cleavage and genome editing [26], was also shown to be able to be redirected to specifically bind and cleave ssRNA [27]. In this system, RNA binding and cleavage by Cas9 require the presence of short DNA sequence (PAMmer: Protospacer Adjacent Motif- containing oligonucleotide) that hybridizes next to the 20 nt recognized ssRNA sequence. This sys-

tem has also been used to track endogenous RNAs in live cells [28]. The main advantage of this system is that it doesn't rely on protein engineering of an RNA-binding surface, and the molecular basis for RNA recognition is much simpler and only requires the presence of a guide RNA that is complementary to the target RNA. However, although this system has proven to be successful to specifically bind or cut a given RNA target from a cellular pool, as well as track RNAs in live cells, it has the inconvenience to require the presence of the PAM either *in cis* or *in trans*, namely presented on a separate DNA oligonucleotide. In addition to PAMmer-activated RNA recognition, a divergent Cas9 from *Francisella novicida* can be used as an alternative/atypical RNA-guided mechanism to target RNA [29, 30]; however, it is unclear how programmable the RNA specificity is and therefore how widely applicable this strategy will be for general RNA targeting. Finally, another CRISPR-Cas protein, Cas13, was recently shown to be a bona fide RNA-guided RNA-targeting enzyme [31, 32]. This protein has found exciting uses for RNA-detection applications [31, 33], but its utility as a programmable RNA-binding protein has yet to be demonstrated.

In this chapter, we focus on the design of RBPs using modular assembly of ZF domains. These small protein domains (≈30 amino acids) are highly diverse family of proteins that have a wide variety of biological functions. These domains fold independently and require coordination of one or more zinc ions to stabilize their 3D structure. ZFs were initially discovered for their dsDNA-binding activity but they can also bind to ssRNA, dsRNA, proteins, or small molecules [34, 35]. We decided to use this protein domain to engineer ssRBPs because (1) it is small and stable; (2) it is robust to mutation; (3) it is modular and recognizes its ligand in a way that is independent from the surrounding protein domains; and finally (4) it has been successfully used to engineer sequence-specific DNA-binding proteins [36]. More specifically, we have used a particular family of ZF that naturally binds ssRNA, the RanBP2-type ZF. This family was first discovered in the human splicing factor ZRANB2. This protein harbors two RanBP2-type ZFs (ZRANB2-ZF-1 and -ZF-2); each of them was shown to bind with micromolar affinity to ssRNA, and specifically recognizes the trinucleotide sequence "GGU" [37]. The crystal structure of the ZRANB2-ZF2:GGU complex has revealed that most of the protein:RNA contacts are side-chain mediated (Fig. 2).

This chapter is divided into two main sections. In the first section, we present the theoretical aspects and a general strategy for modular assembly of individual ZF domains in order to create functional polydactyl ZF proteins. In the second section, we focus on the practical aspects and manipulation of ZF-based proteins by describing the gene design, the recombinant expression, and the purification of a 3-RanBP2-type ZF protein. This protocol can also

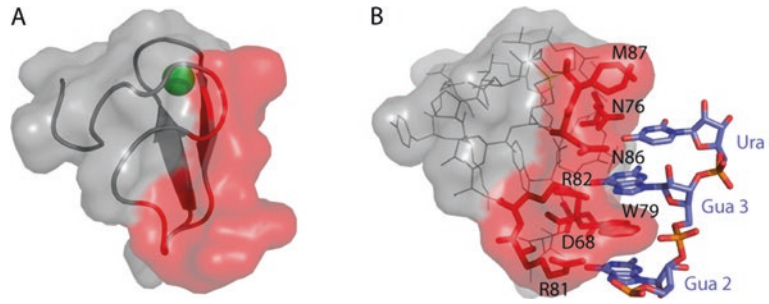


Fig. 2 Representation of the ZRANB2-ZF2 structure alone (**a**) or in complex with its target ssRNA: GGU (**b**) (PDB ID: 3G9Y). (**a**) Structure of the free ZRANB2-ZF2 shown in cartoon. The residues that contact RNA are shown in red. The ZRANB2-ZF2 includes the conserved consensus sequence W-X-C-X_{2-4}-C-X_3-N-X_6-C-X_2-C. Its fold contains two distorted β-hairpins and a single zinc ion (shown in green) that is coordinated by the four cysteines. (**b**) Structure of the ZENAB2-F2:GGU complex. The protein and RNA are shown in grey and blue sticks, respectively. The RNA-binding surface in shown in red. Interactions between the protein and RNA include the W79 side chain that stacks between Gua2 and Gua3. In addition, Gua2, Gua3, and Ura4 make a network of direct or water-mediated hydrogen bonds mostly to the side chains of the protein [37]

be applied and adapted for longer ZF arrays (we've successfully produced arrays with up to six ZFs using this approach).

2 Materials

Prepare all solutions using deionized water (18 MΩ-cm at 25 °C) and analytical grade reagents. Filter all solutions and store them at 4 °C.

2.1 Genetic Constructs

1. ZRANB2 F122 synthetic gene, codon-optimized for *E. coli* expression: The gene was synthetized and cloned into pUC57 by GeneCust (Luxembourg) in between the BamH1 and EcoR1 restriction sites.
2. Expression vector: pMal-c2x (New England Biolabs). Use conventional cloning protocols to insert the F122 gene into the BamH1 and EcoR1 restriction sites present in the pMal-c2x vector.

2.2 Recombinant Expression of ZRANB2 MBP-F122

1. Chemically competent Rosetta™ 2 (DE3) Singles™ competent cells (Novagen).
2. Luria Bertani (LB) media: 10 g Peptone, 5 g yeast extract, 5 g NaCl for 1 L of liquid media prepared with deionized water and sterilized by autoclaving.

3. Luria Bertani (LB)—Agar media: 10 g peptone, 5 g yeast extract, 5 g NaCl, 12 g agar for 1 L of liquid media prepared with deionized water and sterilized by autoclaving.
4. Ampicillin: 100 mg/mL for stock solution prepared with sterile deionized water.
5. Chloramphenicol: 34 mg/mL for stock solution prepared with 100% ethanol.
6. Isopropyl β-D-1-thiogalactopyranoside (IPTG): 1 M for stock solution prepared with sterile deionized water.
7. $ZnCl_2$: 100 mM Stock solution prepared with deionized water and filter sterilized.
8. DURAN® Erlenmeyer flask: Narrow-neck flask capacity 2 L.
9. Incubator shaker with cooling system.

2.3 Purification of ZRANB2 MBP-F122

1. Cell lysis solution: 20 mM Tris–HCl, pH 7.0, 0.5–1 M NaCl (*see* **Note 1**), 1 mM $MgCl_2$, 100 μM $ZnCl_2$, 1 mM D,L-dithiothreitol (DTT) (*see* **Notes 2** and **3**).
2. Complete ethylenediaminetetraacetic acid (EDTA)-free protease inhibitor tablets (Roche).
3. Lysozyme from chicken egg white: 1 mg/mL Prepared in 20 mM Tris–HCl, pH 7.0, 150 mM NaCl.
4. DNase I: 1 mg/mL Prepared in 20 mM Tris–HCl, pH 7.0, 150 mM NaCl.
5. Empty gravity-flow glass column.
6. Amylose Resin High Flow (New England Biolabs).
7. Solution A: 20 mM Tris–HCl, pH 7.0, 150 mM NaCl, 100 μM $ZnCl_2$, 1 mM DTT.
8. Solution B: 20 mM Tris–HCl, pH 7.0, 150 mM NaCl, 100 μM $ZnCl_2$, 1 mM DTT, 5% (v/v) glycerol.
9. Solution C: 20 mM Tris–HCl, pH 7.0, 150 mM NaCl, 100 μM $ZnCl_2$, 1 mM DTT, 50 mM maltose.
10. Thrombin (Thermo Fisher Scientific): 5 μg/μL Prepared in 20 mM Tris–HCl, pH 7.0, 150 mM NaCl.
11. UNO-S (cation-exchange column, Biorad), 1 mL bed volume.
12. Solution D: 20 mM Tris–HCl, pH 7.0, 20 mM NaCl, 100 μM $ZnCl_2$, 1 mM DTT.
13. Solution E: 20 mM Tris–HCl, pH 7.0, 1 M NaCl, 100 μM $ZnCl_2$, 1 mM DTT.
14. NGC™ Medium-Pressure Liquid Chromatography Systems (Biorad) or any medium-pressure liquid chromatography instruments.

3 Methods

3.1 Design Considerations When Engineering ZF-Based Polydactyl ssRBPs

3.1.1 Linker Length

The length of the linker regions used to assemble the ZF domains is critical for the functionality of ssRBPs. In this regard, knowledge of location of the N-terminal (N_{ter}) and C-terminal (C_{ter}) extremities of each ZF domain is important to properly estimate the linker length. If the extremities are close to each other or not exposed on the surface of the domain, modular assembly of these ZFs will require a longer linker (Fig. 3b). In contrast, ZFs that present N_{ter} and C_{ter} extremities located on opposite sides and/or exposed on the protein domain surface can accommodate a shorter linker (Fig. 3a). For example, as mentioned in Subheading 1, ZRANB2 contains two RanBP2-type ZFs (ZF1 and ZF2) that are separated by a ~25 amino acid linker. Both of these RanBP2-type ZFs present N_{ter} and C_{ter} extremities that are located on opposite sides of the ZF domain and are both well exposed on the ZF surface (Fig. 3a).

Based on these considerations, in our previous study we shortened the linker from 25 to 5 amino acids [38] (Fig. 4), and we noticed that this shortened 5-residue linker helped to stabilize the protein against proteolysis without significantly affecting its RNA-binding affinity (relative to the WT protein). This is why we have used successfully this 5-AA linker to generate longer ZF-based protein constructs [38] (*see* **Note 4**).

3.1.2 Linker Amino Acid Composition

In addition to the linker length, the linker composition is also an important parameter to consider when assembling ZFs. The composition of the linker will influence the flexibility/rigidity properties of the resulting multi-domain protein. Glycine- and serine-rich $(GS)_n$ linkers will provide flexibility [39] whereas linkers enriched in alanine $(EAAAK)_n$ offer rigidity [40, 41]. We believe that to reduce any steric hindrance upon RNA binding, it is recommended to use flexible linkers. Another option is to use the natural amino acid composition that is present in the WT protein containing the assembled ZFs. In the case of the RanBP2-type ZFs, we have used both natural linkers (but shortened, see above) [38] and (GSGSG) linkers (publication in preparation). Both strategies led to functional polydactyl ZF proteins (*see* **Note 5**).

3.1.3 Directionality of the ZF:RNA Interaction

Another important parameter to consider when assembling ZFs from different families is the directionality of the ZF:RNA interaction. Indeed, ssRNA recognition by ssRNA-binding domains (including RanBP2 ZFs) is directional. For example, for RanBP2-type ZFs, the interaction between the ZFs and ssRNA takes place using the following orientation: $N_{ter} \rightarrow C_{ter}$ binds $5' \rightarrow 3'$ (Fig. 5a). In contrast, other ssRNA-binding ZFs exhibit the opposite direction when binding to RNA, namely $N_{ter} \rightarrow C_{ter}$ binds $3' \rightarrow 5'$ (Fig. 5b), such as the CCHC-type zinc knuckles (ZnKs) from the nucleocapsid

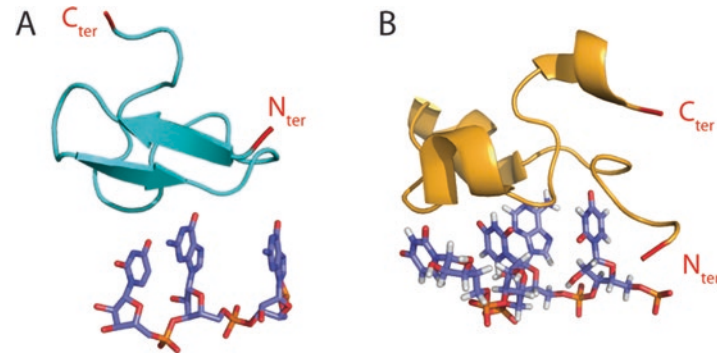


Fig. 3 Representation of the ZRANB2-ZF2 (PDB ID: 3G9Y) (**a**) and the Ts11d-ZF1 (PDB ID: 1RG0) (**b**) structures in complex with their target ssRNA. Their N- and C-terminal extremities are indicated in red. Assembling of ZRANB2 ZFs accommodates short linkers whereas assembling of TIS11d ZFs requires longer linkers to prevent any steric hindrance between the assembled ZFs

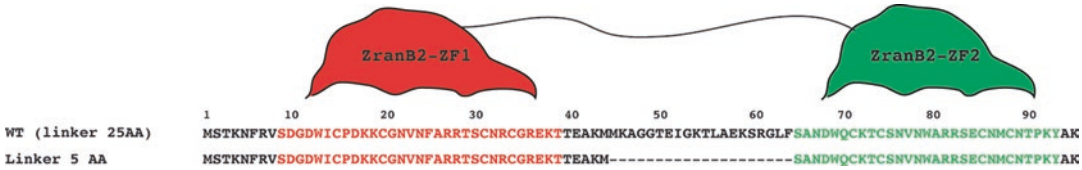


Fig. 4 The inter-domain regions of the two ZranB2-ZFs. ZranB2-ZF1 and ZF2 are represented in red and green, respectively. In the WT protein, the two ZFs are separated by a 25-AA linker region. Previously, we removed 20 residues from this inter-domain region and showed that this new protein construct retained the ability to bind to ssRNA with high affinity. This figure has been adapted from our previous study [38]

protein of the Moloney murine leukemia virus [42] and the ZF from the human posttranscriptional regulator Lin28A [43]. This means that modular assembly of different ZFs to create functional ssRBPs can only be performed if the assembled ZFs recognize RNA using the same directionality.

3.1.4 Construction of a 6-ZF ssRBP

As an illustration of the different aspects that have been mentioned above to create functional ZF-based polydactyl ssRBPs, we have recently generated a functional 6-ZF protein (publication in preparation) using ZFs that belong to different families. This polydactyl 6-ZF protein is represented in Fig. 6 and consists in the assembly of four different members of the RanBP2-type ZF family and two copies of a CCCH-type ZF that belong to the human protein Ts11d/ZFP36L2 (ZF1 and ZF2) [44] that has been shown to be involved in RNA metabolism and definitive hematopoiesis [45]. All six selected ZFs bind to ssRNA using the same orientation ($N_{ter} \rightarrow C_{ter}$ to $5' \rightarrow 3'$). We assembled the RanBP2-type ZFs using a 5-AA linker (GSGSG), except for those followed by the CCCH-type ZFs. In this latter case, four additional residues (GSGSGSGSG) were introduced in order to reduce potential steric hindrances given the

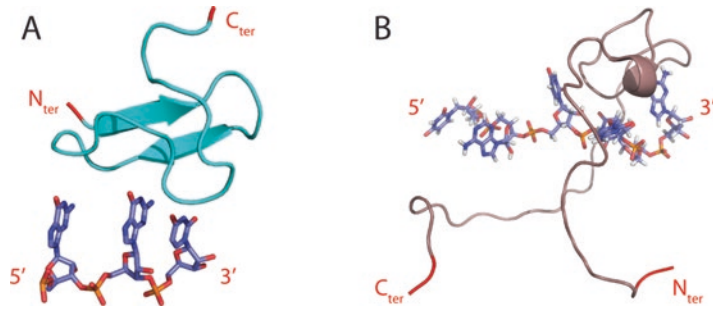


Fig. 5 Representation of the ZranB2-ZF2 (PDB ID: 3G9Y) (**a**) and the Moloney leukemia virus (MLV) nucleocapsid ZnK (PDB ID: 1WWG) (**b**) structures in complex with their target ssRNA. The N_{ter} and C_{ter} extremities are indicated in red on the structures. ZranB2-ZF2 binds to RNA using the following orientation: $N_{ter} \rightarrow C_{ter}$ / 5′ → 3′ whereas the ZnK binds to RNA using the opposite orientation, namely $N_{ter} \rightarrow C_{ter}$ / 3′ → 5′

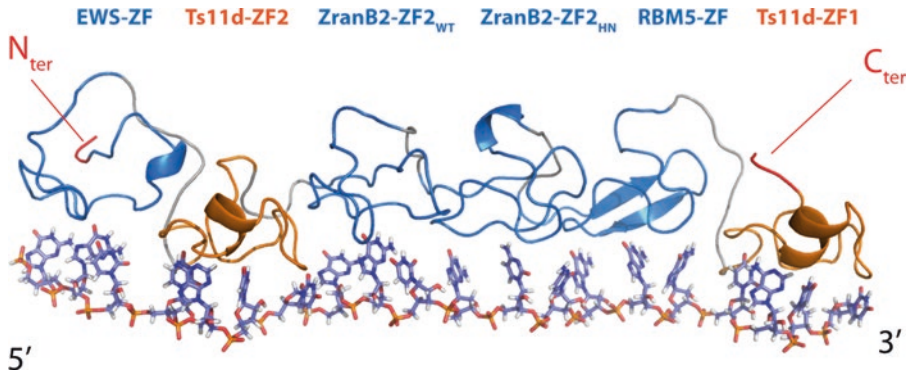


Fig. 6 Representation of the designed 6-ZF chimeric protein in complex with ssRNA. This protein includes (from the N- to the C-terminus) EWS-ZF, Ts11d-ZF2, ZRANB2-ZF2$_{WT}$, ZRANB2-ZF2$_{HN}$, RBM5-ZF, and Ts11d-ZF1. The RanBP2-type and the CCCH-type ZFs are represented in blue and orange, respectively. The linker regions are indicated in grey. All of these ZFs were connected by a 5-AA linker (GSGSG), except for those followed by the Ts11d-ZFs, which were connected by a 9-AA linker region (GSGSGSGSG) to reduce potential steric hindrance. The choice of all ZFs in this array required the selection of ZFs with consistent domain polarity, since ZFs bind to ssRNA in a directional manner (in this case: $N_{ter} \rightarrow C_{ter}$ to 5′ → 3′)

fact that the N_{ter} and C_{ter} extremities of Ts11d-ZFs were located very close to each other on the structure as illustrated in Fig. 3b. This designed 6-ZF protein includes (from the N- to the C-terminus) (1) ZF from the human transcriptional regulator Ewing sarcoma protein, named EWS-ZF (RanBP2-type ZF) [46]; (2) ZF2 from the human protein Ts11d (see above), named Ts11d-ZF2 (CCCH-type ZF); (3) ZRANB2-ZF2$_{WT}$ (RanBP2-type ZF); ZRANB2-ZF2$_{HN}$ a mutated variant of the WT ZRANB2-ZF2 that was engineered for specificity changes (RanBP2-type ZF) [47]; and the ZF domain from the human RNA-binding protein 5, a component of the spliceosome complex A, named RBM5-ZF (RanBP2-type

ZF) [48], ZF1 from the human protein Ts11d (see above), named Tsd11-ZF1 (CCCH-type ZF).

We have recently expressed and purified this 6-ZF construct in fusion with the maltose-binding protein (MBP). Our preliminary data show that the protein is functional and recognizes its target RNA sequence with high affinity (nanomolar range) and relatively high specificity (manuscript in preparation). These data show that it is therefore possible to assemble RNA-binding ZFs from different families to create functional polydactyl ZF proteins.

3.2 Practical Aspects of ZF Domain Modular Assembly

The following part of this chapter is dedicated to the practical aspects of designing and manipulating ZF-based proteins. In particular, we will focus on the design, recombinant expression, and purification of a 3-ZF protein assembled using RanBP2-type ZFs only.

3.2.1 Design of the Gene Encoding the 3-ZF Protein (F122)

Combine the amino sequences of RanBP2-ZF1 (amino acids: 8–41 of ZRANB2: NP_976225) with two copies of RanBP2-ZF2 (amino acids: 65–94 of ZRANB2: NP_976225) using a 5-amino acid linker region (amino acids: 42–46 in between ZF-1 and ZF-2 and amino acids: 95–99 in between the two copies of ZF-2) to connect them together following the scheme presented in Fig. 7.

Order a gene encoding the F122 construct with codon optimization for *E. coli* expression. Clone the gene using well-established cloning protocols in between the BamH1 and EcoR1 restriction sites present in the commercially available vector pMAL-c2X to allow the expression of the F122 construct fused to the C-terminal end of the MBP (*see* **Note 6**). Check the integrity of the gene by DNA sequencing.

3.2.2 Recombinant Expression of MBP-F122

1. Transform *E. coli* Rosetta2 chemically competent cells with 100 ng of the pMal-c2X-F122 vector using the well-established heat-shock protocol and select transformed bacteria overnight at 37 °C on LB agar plate supplemented with 100 μg/mL of ampicillin.
2. Perform a pre-culture by inoculating an isolated bacterial colony into 100 mL of LB media supplemented with 50 μg/mL and 34 μg/mL of ampicillin and chloramphenicol, respectively. Grow the cells overnight at 37 °C with shaking at 200 rpm. Inoculate 4 L of LB (100 μg/mL ampicillin) with 80 mL of pre-cultured cells (maximum 500 mL of culture in a 2 L flask). Grow the cells at 37 °C with constant shaking at 200 rpm until the optical density (OD_{600nm}) reaches ≈0.6–1.0. Induce the expression of MBP-F122 by supplementing the culture with 1 mM IPTG. At the same time add 90 μM of $ZnCl_2$ to obtain

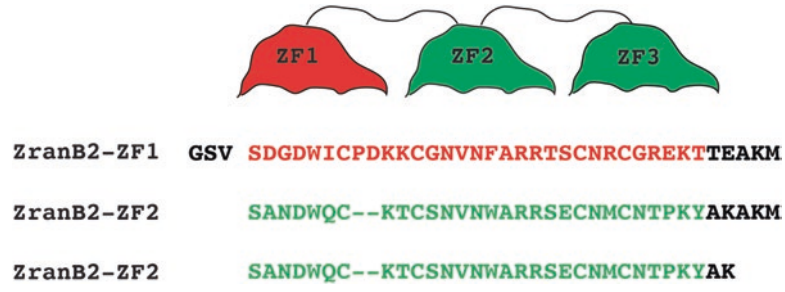


Fig. 7 Schematic representation and amino acid sequence of the 3-ZF protein assembled using RanBP2-type ZFs. This 3-ZF construct consists of one copy of the ZRANB2-ZF1 (N-terminal in red) and two copies from ZRANB2-ZF2 (C-terminal in green). All the ZFs are separated by a five-amino acid linker obtained by shortening the inter-domain region present in WT ZRANB2

reliable overexpression of the protein (*see* **Note 7**). Perform the expression overnight at 18 °C.

3. Harvest the cells by centrifugation at 6500 × *g* for 20 min and discard the supernatant.

3.2.3 Purification of F122-MBP by Affinity Chromatography

1. Resuspend the cell pellets in ~100 mL of lysis buffer (*see* **Note 1**) and add one complete EDTA-free protease inhibitor tablet.
2. Freeze-thaw the cell suspension three times using dry ice and a room-temperature water bath. After the final freeze-thaw, add 200 μL of 1 mg/mL lysozyme and 200 μL of 1 mg/mL DNase I and incubate for 1 h at 4 °C with gentle mixing. Sonicate the cell suspension three times for 60 s (at medium power) while keeping the sample on ice.
3. Centrifuge the cell lysate for 30 min at 24,000 × *g*. Recover and filter the soluble fraction through a 0.22 μm filter. MBP-F122 should be recovered in the soluble fraction.
4. Add 2 mL of amylose fast-flow resin to an empty gravity-flow glass column. Rinse the resin with 50 mL of MilliQ water, and then 50 mL of buffer A. Apply the soluble fraction onto the resin and incubate for 1 h at 4 °C with gentle mixing (*see* **Note 8**).
5. Discard the flow-through and wash the resin with 50 mL of solution B. Then add 10 mL of elution solution (solution C) stepwise, 1 mL at a time, and wait for 10 min before collecting each of the 1 mL elutions.
6. Analyze the eluted fraction using SDS-PAGE.
7. The fractions that contain MBP-F122 are pooled together and used in the next purification steps. The estimation of MBP-F122

concentration is determined by measuring the absorbance at 280 nm.

3.2.4 Purification of F122 on a Cation-Exchange Column

1. To cleave MBP from F122, incubate eluted MBP-F122 with thrombin for 1 h at room temperature or overnight at 4 °C (*see* **Note 9**) in solution D supplemented with 5 mM $CaCl_2$ with gentle mixing. The quantity of thrombin used to cleave the protein fusion is adjusted from the recommendation provided by the manufacturer and varies from 10 to 20 μg of thrombin for 2 mg of MBP-F122.
2. Dialyze the cleaved MBP and F122 mixture using 10 L of solution D for at least 3 h at 4 °C to reduce the NaCl concentration and allow efficient binding of F122 to the ion-exchange column.
3. Purify cleaved F122 from MBP by cation-exchange chromatography (*see* **Note 10**). Equilibrate a UNO-S column (1 mL bed volume) with solution D and load the cleaved MBP-F122 sample at a flow rate of 1 mL/min using a medium-pressure liquid chromatography system (or equivalent). Only F122 (pI: 9.2—positively charged at pH 7) should bind to column, while MBP (pI: 5—negatively charged at pH 7) shouldn't bind to the column. Wash the column with 20 mL of solution D (2 mL/min) and then elute F122 (1 mL/min) using a NaCl gradient starting of solution E (0–100%) over 20 mL (*see* **Note 11**). Eluted fractions are analyzed using SDS-PAGE to check purity of the F122 purification. Pure fractions are pooled together and dialyzed using 10 L of solution A overnight at 4 °C. The concentration is then measured (*see* **Note 12**) and protein aliquots are prepared and stored at −20 °C.
4. Assess the integrity of the F122 protein construct by mass spectrometry and one-dimensional ^{1}H NMR spectroscopy to check for the appropriate length and proper folding of the protein, respectively.
5. The binding activity of purified F122 can be assessed by measuring its affinity for target ssRNA using a range of different methods, e.g., isothermal titration calorimetry (ITC), fluorescence anisotropy (FA), or biolayer interferometry (BLI) (*see* **Note 13**), as reported in our previous studies [38, 47]. Although this chapter doesn't address the binding specificity of assembled ZFs, as well as the different methods to engineer specificity changes (*see* **Note 14**), it is important to mention that the binding specificity can be assessed with the same meth-

ods by measuring the binding affinity of the protein for "control" RNA sequences such as polyA (A_n), polyU (U_n), polyC (C_n), or polyN (N_n).

4 Notes

1. Having a high salt concentration in the lysis buffer (between 0.5 and 1 M NaCl) significantly improves the solubility and stability of our assembled ZF domains. In addition, it also prevents nonspecific binding of nucleic acids present in the cell lysate. This is particularly important during bacterial cell lysis. After lysis, the salt concentration can be decreased during the amylose resin purification step with no significant impact on the stability and solubility of the protein.
2. Maintaining a small concentration of reducing agent and zinc in the purification and storage buffers helps stabilize ZF-containing proteins. The reducing agent prevents oxidation of the cysteines involved in chelating the zinc ion essential for the stability of the protein domain. The presence of a small concentration of zinc also helps each ZF domain maintain its correct fold.
3. Filtering the solutions once all additives have been added ensures the absence of precipitates. This is particularly important when the solution includes zinc chloride. Indeed, despite its high solubility, dilution of highly concentrated solution often leads to the formation of zinc oxychlorides [49]. In addition, zinc can also interact with DTT by coordination to generate an electroneutral ZnL complex (where L is the sulfide diamond of DTT) that has the tendency to precipitate [50]. In the latter case, DTT can be substituted by thiol protectant, Tris (2-carboxyethyl) phosphine (TCEP). In addition to be more stable in solution than DTT, TCEP is a weaker chelator of metal ions [51].
4. The linker length between ZFs can affect the binding specificity of the assembled ZF protein. When using long linkers, the assembled ZF protein can more easily accommodate additional bases in between the sequence motifs that are recognized by the individual ZFs. Thus, reducing the linker length may also improve the binding specificity of the ZF array and help reduce off-target binding.

5. As observed for double-stranded DNA (dsDNA)-binding ZFs [52], positively charged residues can also be added in the linker region in order to favor electrostatic interactions with the negatively charged RNA and help strengthen the overall binding affinity of the $(ZF)_n$:RNA complex.
6. We noticed that single or assembled ZF protein constructs based on RanBP2-type ZFs present low solubility and/or expression yield when produced in *E. coli*. For single- or two-ZF constructs, we obtained the best results using glutathione-S-transferase (GST) fusion proteins expressed from a pGEX-6P expression vector. However for 3-ZF or longer arrays we obtain the best results using MBP.
7. Supplementing the media with zinc upon induction of the protein expression can increase the yield of expressed protein. The presence of zinc most probably helps newly synthetized protein molecules to fold into their native conformation. It is important to note that zinc should not be added before protein expression induction to avoid cell growth inhibition and/or formation of precipitates.
8. We always perform this first purification step of the fusion protein (MBP-F122) on the bench "in batch" using gravity-flow columns. However, you can also use medium-pressure liquid chromatography systems together with high-flow amylose resin.
9. In some cases, nonspecific cleavage and degradation of the fusion protein can occur if incubated with thrombin for too long, or if the quantity of thrombin used is too high. We observed that short cleavage at room temperature limits degradation of the MBP-F122 compared to longer cleavage at 4 °C. Notably, human rhinovirus (HRV) 3C protease can be used rather than thrombin. In this latter case, zinc should be removed from the sample to avoid protease activity inhibition.
10. Cation-exchange chromatography can be substituted for size-exclusion chromatography (such as a Superdex 75) to separate the cleaved F122 from MBP. Benefits for this include the following: First, after thrombin cleavage, salt needs to be removed from the sample by dialysis to allow efficient binding of the protein to the cation-exchange column. This can lead to significant precipitation of the protein. Second, we have observed that if the thrombin cleavage is not complete, the cleaved F122 and the remaining MBP-F122 fusion co-purify on the cation-exchange column, behaving as they have the same isoelectric point (≈9.2). Size-exclusion chromatography can be useful in both cases since you can keep a significant amount of salt in the buffer, as well as efficiently separate F122 (12 kDa) from MBP (42 kDa) and the remaining uncleaved MBP-F122 fusion protein (54 kDa).

11. We do not purify ZF proteins under strictly RNase-free conditions. However, after the purification and concentration process, we add a small amount of RNase inhibitor to the purified protein (100 units of RiboSafe RNase inhibitor from Bioline for 10 mg of purified protein). In addition, we also perform all the binding assays in the presence of RNase inhibitor, which allow us to use any kind of instruments that are not necessarily kept in RNase-free conditions.
12. On average, we obtain medium expression and purification yield for the F122 protein construct, ≈1–2 mg of purified F122 per 1 L of bacterial culture. We observed that for shorter ZF constructs (single ZF or two ZF), the yield increases significantly to 6–8 mg of purified protein for 1 L of bacterial culture expressing a two-ZF protein.
13. We noticed that when performing the binding assays using biolayer interferometry the presence of $ZnCl_2$ affects the amplitudes of the observed signals. In some cases it has also been demonstrated that $ZnCl_2$ affects the RNA-binding activity of proteins [53]. Therefore we recommend not including $ZnCl_2$ in the binding assay solutions.
14. A number of strategies including rational and combinatorial approaches can be used to engineer specificity changes and customize the binding specificity of ssRBPs. In one study, we used a rational approach to engineer specificity changes on the RanBP2-type ZFs. Here we used site-directed mutagenesis based on sequence alignments of various members of the RanBP2-type ZF family to graft particular RNA-binding residues on to the RanBP2-type ZF from the human protein RBM5 [54] with the aim of changing RBM5's RNA-binding specificity. But only little effects on the binding specificity were observed. In a more recent study, we used a combinatorial approach by creating a library of ZRANB2-ZFs where the RNA-binding residues were randomly mutated [47]. Screening of this library by phage display allowed the identification of a ZRANB2-ZF2 variant whose RNA sequence specificity was switched from GGU to GCC. However, the binding affinity of this ZF was much lower compared to the WT ZRANB2-ZF2 and the binding could only take place in the context of a 3-ZF protein where both flanking ZFs were of wild-type sequence. Finally, an alternative strategy is to combine RNA-binding domains of known binding specificities in a desired way to target and bind any chosen RNA sequence. In this context, a SELEX-based (Systematic Evolution of Ligand by EXponentiel enrichment) strategy was recently developed by Burge and coworkers, to identify the binding specificities of numerous ssRBPs [55]. This method, named RNA Bind-n-Seq, consists of coupling a round of SELEX to high-throughput (HT) sequencing for fast and efficient determination of sequence specificities of various ssRBPs.

References

1. Glisovic T, Bachorik JL, Yong J, Dreyfuss G (2008) RNA-binding proteins and post-transcriptional gene regulation. FEBS Lett 582:1977–1986
2. Hogan DJ, Riordan DP, Gerber AP, Herschlag D, Brown PO (2008) Diverse RNA-binding proteins interact with functionally related sets of RNAs, suggesting an extensive regulatory system. PLoS Biol 6:e255
3. Lunde BM, Moore C, Varani G (2007) RNA-binding proteins: modular design for efficient function. Nat Rev Mol Cell Biol 8:479–490
4. Valverde R, Edwards L, Regan L (2008) Structure and function of KH domains. FEBS J 275:2712–2726
5. Clery A, Blatter M, Allain FH (2008) RNA recognition motifs: boring? Not quite. Curr Opin Struct Biol 18:290–298
6. Zamore PD, Williamson JR, Lehmann R (1997) The Pumilio protein binds RNA through a conserved domain that defines a new class of RNA-binding proteins. RNA 3:1421–1433
7. Manna S (2015) An overview of pentatricopeptide repeat proteins and their applications. Biochimie 113:93–99
8. Manival X, Ghisolfi-Nieto L, Joseph G, Bouvet P, Erard M (2001) RNA-binding strategies common to cold-shock domain- and RNA recognition motif-containing proteins. Nucleic Acids Res 29:2223–2233
9. Font J, Mackay JP (2010) Beyond DNA: zinc finger domains as RNA-binding modules. Methods Mol Biol 649:479–491
10. Burge RG, Martinez-Yamout MA, Dyson HJ, Wright PE (2014) Structural characterization of interactions between the double-stranded RNA-binding zinc finger protein JAZ and nucleic acids. Biochemistry 53:1495–1510
11. Parker LM, Fierro-Monti I, Reichman TW, Gunnery S, Mathews MB (2001) Double-stranded RNA-binding proteins and the control of protein synthesis and cell growth. Cold Spring Harb Symp Quant Biol 66:485–497
12. Masliah G, Barraud P, Allain FH (2013) RNA recognition by double-stranded RNA binding domains: a matter of shape and sequence. Cell Mol Life Sci 70:1875–1895
13. Auweter SD, Oberstrass FC, Allain FH (2006) Sequence-specific binding of single-stranded RNA: is there a code for recognition? Nucleic Acids Res 34:4943–4959
14. Keryer-Bibens C, Barreau C, Osborne HB (2008) Tethering of proteins to RNAs by bacteriophage proteins. Biol Cell 100:125–138
15. Shav-Tal Y, Darzacq X, Shenoy SM, Fusco D, Janicki SM, Spector DL, Singer RH (2004) Dynamics of single mRNPs in nuclei of living cells. Science 304:1797–1800
16. Golding I, Paulsson J, Zawilski SM, Cox EC (2005) Real-time kinetics of gene activity in individual bacteria. Cell 123:1025–1036
17. De Gregorio E, Preiss T, Hentze MW (1999) Translation driven by an eIF4G core domain in vivo. EMBO J 18:4865–4874
18. Chen Y, Varani G (2011) Finding the missing code of RNA recognition by PUF proteins. Chem Biol 18:821–823
19. Dong S, Wang Y, Cassidy-Amstutz C, Lu G, Bigler R, Jezyk MR, Li C, Hall TM, Wang Z (2011) Specific and modular binding code for cytosine recognition in Pumilio/FBF (PUF) RNA-binding domains. J Biol Chem 286:26732–26742
20. Choudhury R, Tsai YS, Dominguez D, Wang Y, Wang Z (2012) Engineering RNA endonucleases with customized sequence specificities. Nat Commun 3:1147
21. Cooke A, Prigge A, Opperman L, Wickens M (2011) Targeted translational regulation using the PUF protein family scaffold. Proc Natl Acad Sci U S A 108:15870–15875
22. Wang Y, Opperman L, Wickens M, Hall TM (2009) Structural basis for specific recognition of multiple mRNA targets by a PUF regulatory protein. Proc Natl Acad Sci U S A 106:20186–20191
23. Miller MT, Higgin JJ, Hall TM (2008) Basis of altered RNA-binding specificity by PUF proteins revealed by crystal structures of yeast Puf4p. Nat Struct Mol Biol 15:397–402
24. Gupta YK, Nair DT, Wharton RP, Aggarwal AK (2008) Structures of human Pumilio with noncognate RNAs reveal molecular mechanisms for binding promiscuity. Structure 16:549–557
25. Lu G, Hall TM (2011) Alternate modes of cognate RNA recognition by human PUMILIO proteins. Structure 19:361–367
26. Sander JD, Joung JK (2014) CRISPR-Cas systems for editing, regulating and targeting genomes. Nat Biotechnol 32:347–355
27. O'Connell MR, Oakes BL, Sternberg SH, East-Seletsky A, Kaplan M, Doudna JA (2014) Programmable RNA recognition and cleavage by CRISPR/Cas9. Nature 516:263–266

28. Nelles DA, Fang MY, O'Connell MR, Xu JL, Markmiller SJ, Doudna JA, Yeo GW (2016) Programmable RNA tracking in live cells with CRISPR/Cas9. Cell 165:488–496
29. Sampson TR, Saroj SD, Llewellyn AC, Tzeng YL, Weiss DS (2013) A CRISPR/Cas system mediates bacterial innate immune evasion and virulence. Nature 497:254–257
30. Price AA, Sampson TR, Ratner HK, Grakoui A, Weiss DS (2015) Cas9-mediated targeting of viral RNA in eukaryotic cells. Proc Natl Acad Sci U S A 112:6164–6169
31. East-Seletsky A, O'Connell MR, Knight SC, Burstein D, Cate JH, Tjian R, Doudna JA (2016) Two distinct RNase activities of CRISPR-C2c2 enable guide-RNA processing and RNA detection. Nature 538:270–273
32. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, Regev A, Lander ES, Koonin EV, Zhang F (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353:aaf5573
33. Gootenberg JS, Abudayyeh OO, Lee JW, Essletzbichler P, Dy AJ, Joung J, Verdine V, Donghia N, Daringer NM, Freije CA, Myhrvold C, Bhattacharyya RP, Livny J, Regev A, Koonin EV, Hung DT, Sabeti PC, Collins JJ, Zhang F (2017) Nucleic acid detection with CRISPR-Cas13a/C2c2. Science 356:438–442
34. Klug A (2010) The discovery of zinc fingers and their applications in gene regulation and genome manipulation. Annu Rev Biochem 79:213–231
35. Gamsjaeger R, Liew CK, Loughlin FE, Crossley M, Mackay JP (2007) Sticky fingers: zinc-fingers as protein-recognition motifs. Trends Biochem Sci 32:63–70
36. Carroll D, Morton JJ, Beumer KJ, Segal DJ (2006) Design, construction and in vitro testing of zinc finger nucleases. Nat Protoc 1:1329–1341
37. Loughlin FE, Mansfield RE, Vaz PM, McGrath AP, Setiyaputra S, Gamsjaeger R, Chen ES, Morris BJ, Guss JM, Mackay JP (2009) The zinc fingers of the SR-like protein ZRANB2 are single-stranded RNA-binding domains that recognize 5′ splice site-like sequences. Proc Natl Acad Sci U S A 106:5581–5586
38. O'Connell MR, Vandevenne M, Nguyen CD, Matthews JM, Gamsjaeger R, Segal DJ, Mackay JP (2012) Modular assembly of RanBP2-type zinc finger domains to target single-stranded RNA. Angew Chem Int Ed Engl 51:5371–5375
39. Argos P (1990) An investigation of oligopeptides linking domains in protein tertiary structures and possible candidates for general gene fusion. J Mol Biol 211:943–958
40. Arai R, Ueda H, Kitayama A, Kamiya N, Nagamune T (2001) Design of the linkers which effectively separate domains of a bifunctional fusion protein. Protein Eng 14:529–532
41. Arai R, Wriggers W, Nishikawa Y, Nagamune T, Fujisawa T (2004) Conformations of variably linked chimeric proteins evaluated by synchrotron X-ray small-angle scattering. Proteins 57:829–838
42. Dey A, York D, Smalls-Mantey A, Summers MF (2005) Composition and sequence-dependent binding of RNA to the nucleocapsid protein of Moloney murine leukemia virus. Biochemistry 44:3735–3744
43. Peters DT, Fung HK, Levdikov VM, Irmscher T, Warrander FC, Greive SJ, Kovalevskiy O, Isaacs HV, Coles M, Antson AA (2016) Human Lin28 forms a high-affinity 1:1 complex with the 106~363 cluster miRNA miR-363. Biochemistry 55:5021–5027
44. Morgan BR, Massi F (2010) A computational study of RNA binding and specificity in the tandem zinc finger domain of TIS11d. Protein Sci 19:1222–1234
45. Iwanaga E, Nanri T, Mitsuya H, Asou N (2011) Mutation in the RNA binding protein TIS11D/ZFP36L2 is associated with the pathogenesis of acute leukemia. Int J Oncol 38:25–31
46. Luo Y, Blechingberg J, Fernandes AM, Li S, Fryland T, Borglum AD, Bolund L, Nielsen AL (2015) EWS and FUS bind a subset of transcribed genes encoding proteins enriched in RNA regulatory functions. BMC Genomics 16:929
47. Vandevenne M, O'Connell MR, Helder S, Shepherd NE, Matthews JM, Kwan AH, Segal DJ, Mackay JP (2014) Engineering specificity changes on a RanBP2 zinc finger that binds single-stranded RNA. Angew Chem Int Ed Engl 53:7848–7852
48. Sutherland LC, Rintala-Maki ND, White RD, Morin CD (2005) RNA binding motif (RBM) proteins: a novel family of apoptosis modulators? J Cell Biochem 94:5–24
49. Skou E, Jacobsen T, van der Hoeven W, Atlung S (1977) On the zinc-chloride complex formation. Electrochim Acta 22:169–174
50. Krężel A, Lesniak W, Jezowska-Bojczuk M, Mlynarz P, Brasun J, Kozlowski H, Bal W

(2001) Coordination of heavy metals by dithiothreitol, a commonly used thiol group protectant. J Inorg Biochem 84:77–88

51. Krezel A, Latajka R, Bujacz GD, Bal W (2003) Coordination properties of tris(2-carboxyethyl) phosphine, a newly introduced thiol reductant, and its oxide. Inorg Chem 42:1994–2003
52. Laity JH, Dyson HJ, Wright PE (2000) DNA-induced alpha-helix capping in conserved linker sequences is a determinant of binding affinity in Cys(2)-His(2) zinc fingers. J Mol Biol 295:719–727
53. Worthington MT, Pelo JW, Sachedina MA, Applegate JL, Arseneau KO, Pizarro TT (2002) RNA binding properties of the AU-rich element-binding recombinant Nup475/TIS11/tristetraprolin protein. J Biol Chem 277:48558–48564
54. Nguyen CD, Mansfield RE, Leung W, Vaz PM, Loughlin FE, Grant RP, Mackay JP (2011) Characterization of a family of RanBP2-type zinc fingers that can recognize single-stranded RNA. J Mol Biol 407:273–283
55. Lambert N, Robertson A, Jangi M, McGeary S, Sharp PA, Burge CB (2014) RNA Bind-n-Seq: quantitative assessment of the sequence and structural binding specificity of RNA binding proteins. Mol Cell 54:887–900

Chapter 6

Design of a System for Monitoring Ubiquitination Activities of E2 Enzymes Using Engineered RING Finger Proteins

Kazuhide Miyamoto and Kazuki Saito

Abstract

Ubiquitination is a sequential cascade consisting of ubiquitin-activating (E1), ubiquitin-conjugating (E2), and ubiquitin-ligating (E3) enzymes. It controls numerous processes such as protein degradation, DNA repair, and signal transduction pathways. E2 enzymes are associated with a variety of diseases such as leukemia, breast cancer, lung cancer, and colorectal cancer. To date, the monitoring of E2 activity for cancer diagnosis is challenging due to its intricate cascade reaction. To surmount this hurdle, we have recently developed a novel strategy for monitoring E2 activities. Here, we describe the concise machinery of ubiquitination with artificial RING finger proteins (ARFs) functioning as E3 enzymes. This machinery enables the simplified monitoring of E2 activities. Furthermore, our system combines a signal accumulation ion-sensitive field-effect transistor biosensor with ARFs, allowing for real-time monitoring of the pathological conditions of cancer cells. The present methodology may lead to novel diagnostic techniques for cancers.

Key words Ubiquitination, Cancer diagnosis, E2, E3, RING finger, ARF

1 Introduction

Protein ubiquitination is a posttranslational modification that controls a variety of cellular processes such as protein degradation, DNA repair, and signal transduction [1–3]. Ubiquitination is a cascade reaction consisting of three steps involving ubiquitin (Ub)-activating (E1), Ub-conjugating (E2), and Ub-ligating (E3) enzymes [2, 4]. E3 transfers activated Ub from E2 enzymes to the substrate lysines [5]. There are three major classes of E3 ubiquitination ligases in eukaryotes known as RING E3, Ubox E3, and HECT E3. RING E3 possesses a RING finger and a substrate-binding domain (SBD) responsible for recognizing a particular substrate. RING fingers comprised approximately 50 amino acids and bind two zinc atoms with eight ligands of Cys and/or His in a unique cross-brace arrangement [6, 7]. E2 and E3 activities are associated with various cancers such as leukemia [8], breast cancer [9], lung cancer [10], and colorectal cancer [11]. For example,

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_6,

seven in absentia homolog 1 (SIAH1) belongs to RING E3 and cooperates with UbcH8 of E2. It induces the ubiquitination of the fusion proteins AML1-ETO [acute myeloid leukemia (*RUNX1*)-eight twenty-one (*MTG8*)] and PML-RARα (promyelocytic leukemia-retinoic acid receptor α), which is related to leukemogenesis [12–14].

Ubiquitinated fusion proteins in cells are recruited in the 26S proteasome degradation pathway. Thus, we believe that the detection of E2 activity should enable the monitoring of tumor development and the capturing of pathological conditions of cancers. E2 activities are estimated in terms of the amount of additional Ub that is transferred from E2 conjugates to substrates. However, the capturing of E2 activities for cancer diagnosis is challenging due to the intricate enzymatic cascade reaction of which E2 is a part. Similar to SIAH1, one of the E2s transfers Ub to several different types of substrates via its partner E3. It is difficult to evaluate E2 activity based on the amounts of additional Ub transferred to all these substrates. Furthermore, the intermediate complex of the E2-Ub conjugate and E3 may occur in the Ub-transferring system [15]. To overcome this hurdle, we have recently developed a novel strategy for monitoring E2 activities. In this chapter, we describe the molecular design of artificial E3 enzymes derived from RING and the resulting concise machinery of ubiquitination. This method allowed for the simplified detection of E2 activities, leading to the monitoring of pathological conditions of cancer cells following treatment with the anticancer drug bortezomib. Thus, the present methodology may open up new directions in cancer diagnostic techniques.

2 Materials

Prepare all samples with special grade chemicals and Milli-Q water (18 MΩ·cm). Use HPLC-grade chemicals for peptide purification. Prepare and store all samples at room temperature unless otherwise noted.

2.1 Synthesis and Purification of Peptide

1. Synthesis: Automated peptide synthesizer.
2. Purification: Reverse-phase HPLC equipped with a C18 column.

2.2 Refolding

1. Solution A: 20 mM Tris–HCl, pH 6.8, 50 mM NaCl, 1 mM dithiothreitol (DTT), and 50 μM $ZnCl_2$.
2. Dialysis cassette: Slide-A-Lyzer dialysis cassette.

2.3 Western Blotting

1. Ubiquitination buffer A: 20 mM Tris–HCl, pH 6.8, 5 mM Mg-ATP, 1 mM DTT, 20 U/mL inorganic pyrophosphatase (IPP), 50 μM $ZnCl_2$, 2.5 μM randomly biotinylated Ub, 0.1 μM

human recombinant E1 (His_6-tagged), and 2.5 μg human recombinant E2 (His_6-tagged) (E1, E2, and Ub available from Enzo Life Sciences, Farmingdale, NY, USA).

2. SDS sample buffer: 65.8 mM Tris–HCl, pH 6.8, 26.3% glycerol, 2.1% SDS, and 0.01% bromophenol blue.
3. Streptavidin-horseradish peroxidase solution: Vectastain ABC Elite Kit (Vector Laboratories, Burlingame, CA, USA).
4. Western blotting detection reagent for chemiluminescence: ECL.
5. Luminescent Image Analyzer.

2.4 E2 Activity Monitoring via Biosensor

1. Ubiquitination buffer B: 1.5 mM Tris–HCl, pH 6.8, 375 μM Mg-ATP, 75 μM DTT, 1.5 U/mL IPP, 4.0 μM $ZnCl_2$, 3.0 μM Ub, 9.0 nM human recombinant E1, and human recombinant E2.
2. Culture of NB4 cells: RPMI 1640 medium containing 10% FBS, 100 U/mL penicillin, and 100 mg/mL streptomycin under the conditions of 5% CO_2 at 37 °C.
3. Bortezomib (marketed as Velcade): Funakoshi Co., Ltd.
4. Biosensor: AMIS-101 Micro Bioactive Analyzer (Bio-X Inc., Kyoto, Japan).

3 Methods

3.1 Molecular Design of Artificial RING Fingers (ARFs)

Here we describe the protocol for creating ARFs as artificial E3 enzymes to conveniently measure E2 activities during ubiquitination [16, 17].

1. For engineering ARFs, obtain the amino acid sequences of the E3 RINGs of interest from a public database, e.g., Simple Modular Architecture Research Tool (SMART). As an example, Fig. 1a shows the selection of EL5 and SIAH1. These belong to the RING E3 and compose a RING finger and an SBD.
2. Extract the amino acid sequence of the active α-helical region (DMWLGSHST for EL5; PKLTC for SIAH1) (*see* **Note 1**).
3. When designing ARFs, the PHD finger of Williams–Beuren syndrome transcription factor (WSTF) is needed as the scaffold [18]. Prepare the amino acid sequence of the WSTF PHD finger as shown in Fig. 1b (*see* **Note 2**).
4. Delete the flexible long loop of the PHD finger, and instead transplant the α-helical region as the active site of the RING finger into the structure of the PHD finger (Fig. 1b). This is "α-helical region substitution" method [17] (*see* **Note 3**).

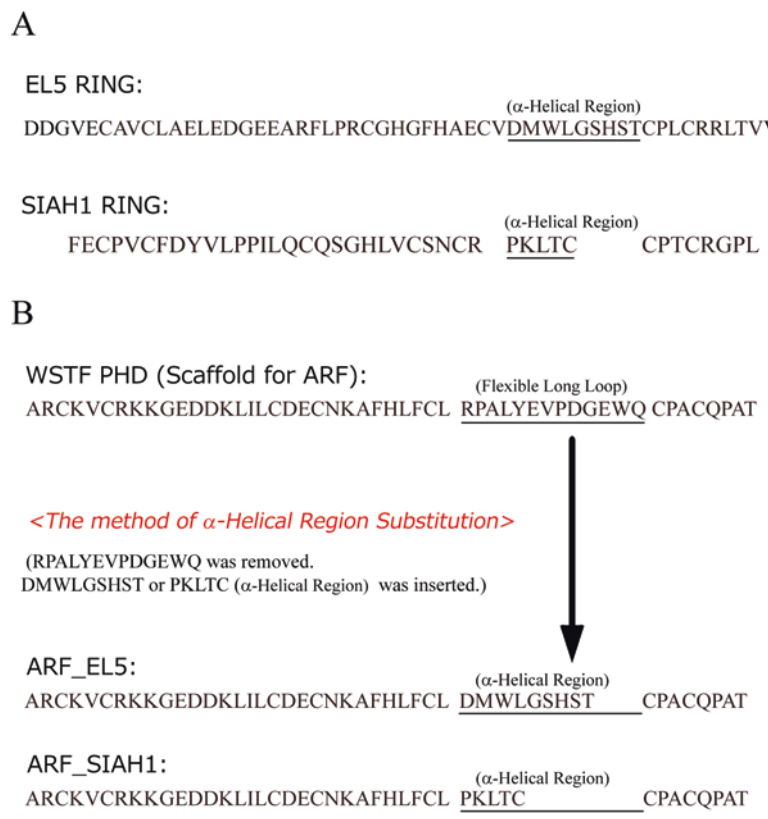


Fig. 1 Schematic for engineering artificial RING fingers (ARFs). (**a**) Amino acid sequences of the RING fingers of EL5 and SIAH1. The α-helical regions for controlling specific E2-binding capabilities are indicated with underlining. (**b**) The method of α-helical region substitution for engineering ARF. The α-helical regions of RINGs are transplanted into the sequence of WSTF PHD. ARF_EL5 and ARF_SIAH1 are designed based on EL5 and SIAH1, respectively. ARFs are chimeric molecules between two different domains

5. The insertion of the only active sites allows for the downsizing of E3s into the small molecule ARFs. For example, SIAH1 functions as a stable homodimer of 564 residues [19]. The molecular size of ARF_SIAH1 (approximately 50 mer) is less than one-tenth that of SIAH1 (Fig. 2).
6. Synthesize ARFs by the standard F-moc solid-phase method on an automated peptide synthesizer based on their amino acid sequences.
7. Purify using HPLC system after cleavage of ARF peptides by the standard method with trifluoroacetic acid.
8. Dissolve the ARF peptide in 1 mL of 8 M guanidine-HCl until it is completely denatured. Dialyze against solution A containing zinc ions overnight at 4 °C using a dialysis cassette,

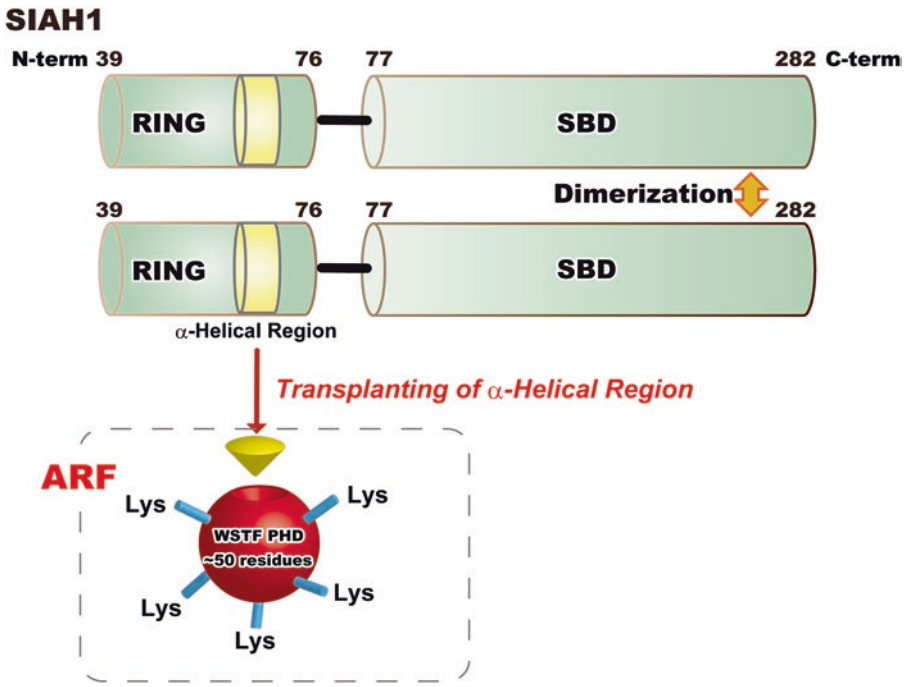


Fig. 2 The grafting design downsizes E3 into a small molecule ARF. SIAH1 has a RING and substrate-binding domain (SBD). It is a stable homodimer of 564 residues. ARF is created by transplanting the α-helical region into an approximately 50 mer peptide. Lys residues of ARF are possible acceptors for the ubiquitin transfer

according to the manufacturer's instructions. This procedure leads to the refolding of the ARF structure.

9. Confirm the formation of the ARF structure by experiments on the chemical modification of Cys residues and by analysis of the circular dichroism spectrum [16, 17] (*see* **Note 4**).

3.2 Unique Ubiquitination Machinery with ARF

ARF ubiquitinates itself without a substrate, although a substrate in traditional ubiquitination is inevitably needed for the Ub transfer. The PHD finger serves as a scaffold for ARF, and its Lys residues are possible acceptors for the Ub transfer (*see* **Note 5**) [20]. Figure 3a shows the novel ubiquitination reaction with ARF. ARF binds to the E2-Ub thioester conjugate and directly transfers activated Ub from E2 to itself. The engineered ARF has specific E2-binding capabilities and also assumes the role of substrate. Thus, ARF is an artificial multifunctional molecule with an auto-ubiquitination reaction. E2 activities can be measured based on ARF reactivity, which is equivalent to the amount of Ub transferred to the ARF.

The procedures to perform in vitro substrate-independent ubiquitination with ARF are as follows:

(Conduct all experiments at room temperature unless otherwise noted).

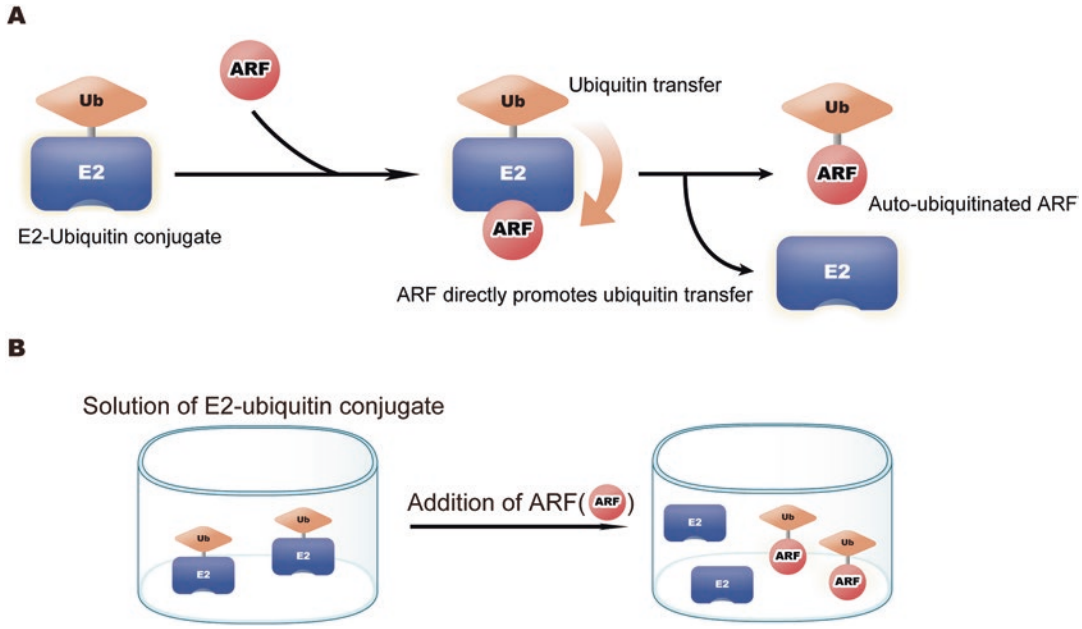


Fig. 3 Novel concise ubiquitination machinery with ARF. (**a**) ARF directly catalyzes ubiquitin (Ub) transfer into itself. ARF is a multifunctional molecule that possesses specific E2-binding and auto-ubiquitination abilities without a substrate. (**b**) Addition of ARF into the solution of E2–Ub conjugates leads to the accumulation of ARF–Ub conjugates and free thiol group (-SH) of E2s

3.2.1 Detection of E2 Activities Using Western Blotting

1. Prepare 45 μL of ubiquitination buffer A, consisting of E1, E2, ATP, and randomly biotinylated Ub. There are approximately 30 E2s in humans [21, 22]. Select the particular E2s of interest.
2. Add 5 μL of ARF (approximately 100 μM), obtained from the above procedure (Subheading 3.1), into 45 μL of ubiquitination buffer A, leading to the accumulation of ARF–Ub conjugates as shown in Fig. 3b. The development of poly- or mono-ubiquitination of ARF depends on the amino acid sequence of the α-helical region that is transplanted by the α-helical region substitution method. For example, ARF_EL5 and ARF_SIAH1 preferentially promote poly-ubiquitination and mono-ubiquitination, respectively. ARF_EL5 and ARF_SIAH1 exhibit similar specific E2-binding capabilities to those of original EL5 and SIAH1, respectively (*see* **Note 6**).
3. Gently mix ubiquitination buffer, and then incubate at 37 °C for 60 min.
4. Add nonreducing SDS sample buffer into the mixture to quench the ARF reaction.
5. Separate using 10–20% SDS-PAGE and transfer to PVDF membrane. Next, for the detection of biotinylated Ub, react the membrane with the streptavidin-horseradish peroxidase

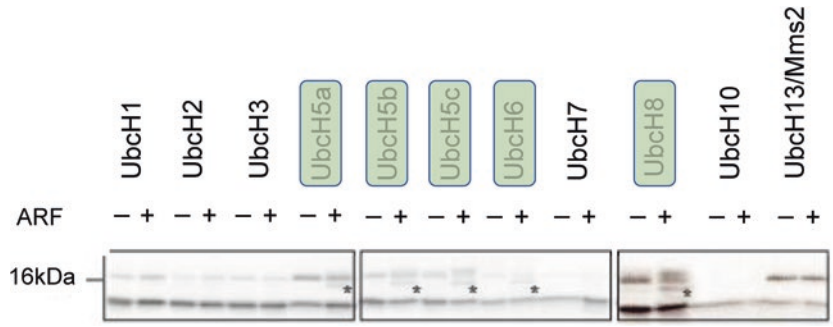


Fig. 4 Substrate-independent ubiquitination reaction with ARF. ARF_SIAH1 was mono-ubiquitinated with UbcH5a, UbcH5b, UbcH5c, UbcH6, or UbcH8, but not with UbcH1, UbcH2, UbcH3, UbcH7, UbcH10, or UbcH13/Mms2. The bands corresponding to the mono-ubiquitination products are identified with stars. Poly-ubiquitination of ARF_SIAH1 was not developed

solution, according to the manufacturer's instructions, and develop by enhanced chemiluminescence of the western blotting detection reagent. Detect the emitted signals on a Luminescent Image Analyzer.

6. As an example, the ubiquitination of ARF_SIAH1 is shown in Fig. 4. ARF_SIAH1 was mono-ubiquitinated in the presence of UbcH5a, UbcH5b, UbcH5c, UbcH6, or UbcH8, but not in the presence of UbcH1, UbcH2, UbcH3, UbcH7, UbcH10, or UbcH13/Mms2. Poly-ubiquitination of ARF_SIAH1 was not promoted if 11 E2s were incubated together. Thus, ARF_SIAH1 has specific E2-binding capabilities and ubiquitinates itself.

3.2.2 E2 Activities Captured Through Polarization Behavior of Electrons in a Semiconductor

A signal accumulation ion-sensitive field-effect transistor biosensor (AMIS sensor) enables the measurement of proton changes in solution [23]. This technique is based on the polarization behavior of electrons in a semiconductor. The system does not require labeling of fluorescent dye molecules, unlike the bioluminescence resonance energy transfer method. The ARF reaction promotes the transfer of Ub from E2–Ub conjugates on the AMIS sensor, yielding ARF–Ub conjugates [16] (Fig. 3b). The aminolytic cleavage of E2–Ub conjugates on the AMIS sensor produces free thiol groups (-SH) of the active-site cysteine of E2s [1], augmenting the free protons in solution. Accordingly, the addition of ARF induces a current that flows through gate channels in the AMIS sensor. This current, which is accumulated and amplified in the sensor, is detected as the AMIS signal, allowing for the high-sensitivity quantitative detection of E2 activities in real time.

1. Prepare ubiquitination buffer B consisting of E1, E2, ATP, and Ub.

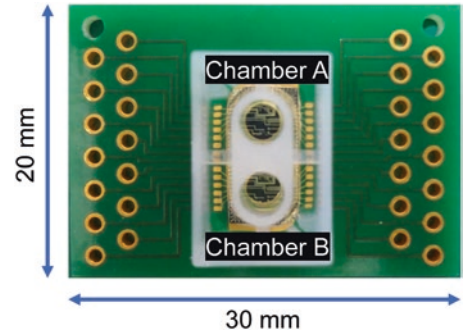


Fig. 5 Signal accumulation ion-sensitive field-effect transistor biosensor (AMIS sensor). The AMIS sensor is a 20 × 30 mm electronic board. It has a two-chambered device (chambers A and B). Each chamber adopts an ultralow volume of 20 μL on the sensor

2. Ubiquitination buffer B (16 μL) was introduced to both chambers A and B of a two-chambered device of the AMIS sensor on an AMIS-101 Analyzer (Fig. 5).
3. Allow the solutions to stand for 10 min to form E2–Ub conjugates.
4. Add 4 μL of solution A into chamber A, and add 4 μL of the 50 μM ARF solution (as obtained from Subheading 3.1) into chamber B. Gently mix the resulting 20 μL mixture in each chamber by pipetting.
5. Acquire the AMIS signals during ubiquitination reactions at 37 °C with an AMIS-101 Analyzer. Set the interval time for collecting real-time signals, typically at 5-s intervals.
6. Obtain the AMIS data for E2 activities by subtracting the AMIS signals of chamber A from those of chamber B (*see* **Note 7**).
7. As an example, Fig. 6 shows the real-time detection of AMIS signals at 37 °C recorded for 600 s at 5-s intervals. ARF_SIAH1 (*see* above Subheading 3.1) was used in the ubiquitination reaction. After the peak of the injection shock, the AMIS signals from E2 UbcH8 activities were detected from approximately 150 s, reaching a plateau at approximately 300 s (*see* **Note 8**). Maintain the device of the AMIS sensor at a constant temperature.

3.2.3 Real-Time Monitoring of E2 Activities Using Cancer Cells with an AMIS Sensor

Bortezomib is a highly selective proteasome inhibitor and induces cell death in various cancers, such as leukemia, multiple myeloma, and breast cancer [24–26]. Treatment with bortezomib reduces cell viability of human acute promyelocytic leukemia-derived NB4 cells, wherein the UbcH8-Ub conjugate leaks out of NB4 cells into the culture supernatant, but the UbcH5 and UbcH6 conjugates do not [8]. The diapedesis of UbcH8-Ub conjugates into the

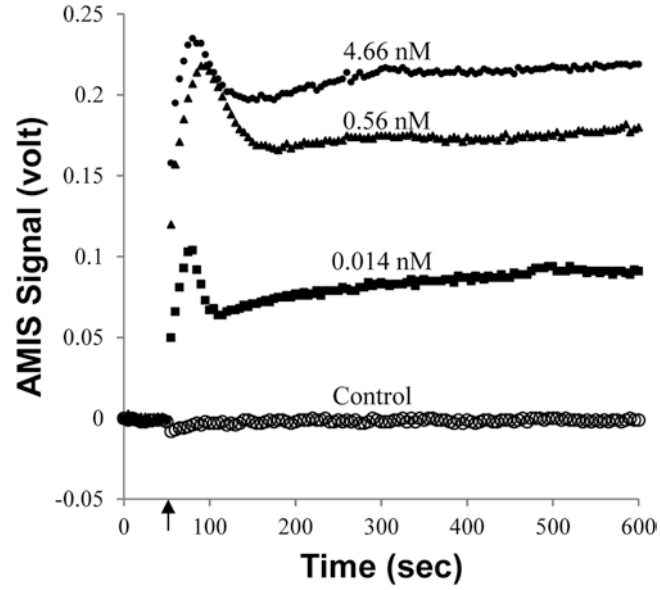


Fig. 6 Real-time monitoring of E2 activities on an AMIS sensor. Addition of ARF_SIAH1 into a solution of E2 UbcH8 causes proton changes, leading to the altering of the electron polarization for the AMIS signals in the sensing layer. The AMIS signals were continuously measured as E2 UbcH8 activities. The representative AMIS signals for 0.014 (filled square), 0.56 (filled triangle), and 4.66 (filled circle) nM E2 UbcH8 were acquired at 5-s intervals, and the AMIS signals in the absence of ubiquitin were used as the control (circle). The arrow indicates the injection time of ARF_SIAH1 on an AMIS sensor

culture supernatant increases in a concentration-dependent manner with drug treatment. The approach described here, combining an AMIS system with the ARF reaction, can be deployed to measure the E2 activity of UbcH8 in crude extracts, such as the culture supernatant of cells [23].

1. Prepare a 1.0 mM stock solution of bortezomib dissolved in DMSO and store at −20 °C.
2. Culture NB4 cells with RPMI 1640 medium [23].
3. After 3 h of growth in FBS-free RPMI-1640 medium, treat NB4 cells (1×10^5 cells/mL) with 0, 10, and 100 nM bortezomib for 48 h at 37 °C.
4. Collect the supernatant of NB4 cells for the following ubiquitination experiments, after centrifugation at $300 \times g$ for 300 s.
5. The supernatant obtained should be diluted fivefold with water (*see* **Note 9**).
6. Inject the diluted samples (16 μL) into a two-chambered device of the AMIS sensor (Fig. 5).

7. Prepare ARF_SIAH1 (*see* above Subheading 3.1) for the ubiquitination reaction.
8. Add 4 μL of solution A into chamber A and 4 μL of the 50 μM ARF solution into chamber B. Gently mix the mixtures (each chamber should contain 20 μL total of its combined solution) by pipetting.
9. Acquire the AMIS signals during the ubiquitination reactions at 37 °C on an AMIS-101 Analyzer.
10. Obtain the AMIS data for the E2 activity of UbcH8 by subtracting the AMIS signals of chamber A from those of chamber B (*see* **Notes 7** and **10**) [23].

Using the Ub machinery with ARF, the rapid and accurate monitoring of E2 activities was conveniently achieved within several minutes. Leukemia is diagnosed by the existence of approximately 2×10^5 cells/mL of blood [27]. With the technique described here, E2 activities leaked out of NB4 cells (cultured NB4 cells, 1×10^5 cells/mL) were successfully detected in the culture supernatant. Thus, the present approach could be applied to the detection of E2 activities in leukemia patients and may prove a convenient method for arranging suitable treatment programs with patients. Different E2 activities are associated with various cancers, such as lung cancer (UbcH10) and breast cancer (UbcH13) [28, 29]. Therefore, the ARF method presented here could be widely applicable to the measuring of E2 activities related to these cancers. We hope that this monitoring system will be used to evaluate responses to treatment with a proteasome inhibitor, as a new companion diagnostic approach.

4 Notes

1. RING fingers have the essential active site of the α-helical region corresponding to the amino acid sequence between the sixth and seventh zinc-binding residues [30–33]. The helical region controls the specific E2–E3 binding capabilities [34].
2. Like the RING finger, the PHD finger adopts a cross-braced zinc-binding arrangement. However, the PHD finger has a randomized flexible long loop rather than the α-helical region. Accordingly, the PHD finger obviously does not have the E2-binding site, and thus it is not associated with the ubiquitination system [35, 36].
3. The substituting region confers specific E2-binding capabilities on the PHD finger. ARF is created as a RING-PHD chimeric molecule. ARF_EL5 and ARF_SIAH1 are designed from EL5 and SIAH1, respectively.

4. The cysteine modification of ARF is performed through treatment with *p*-(hydroxymercuri) benzoic acid, and the released zinc ions are quantified using the metallochromic indicator 4-(2-pyridylazo) resorcinol. The molar ratio of [Zn]/[ARF] should be approximately 2.0. The typical CD spectrum of ARF indicates an ordered secondary structure including the α-helical conformation with double minima occurring at approximately 205 (π-π* transitions) and 225 nm (n-π* transitions) [17].
5. The PHD finger specifically recognizes histone H3 methylated at Lys4 and controls transcriptional regulation in cells [37, 38]. The PHD finger is characterized as a positively charged Lys-rich motif in the nucleus [18].
6. Original EL5 develops poly-ubiquitination on the substrate with UbcH5a but not with UbcH7/8 [31]. Original SIAH1 promotes mono-ubiquitination with UbcH5 or UbcH8 [14, 39].
7. In AMIS measurements, the effects of protons, except for the reaction between ARF and E2s, are completely canceled out by subtracting the AMIS signals of chamber A from those of chamber B.
8. The AMIS signals of E2 activities adopt concentration-dependent behavior and indicate a linear relationship between intensity (volts) and logarithm of concentration of E2s over a wide response range of femtomolar to micromolar concentrations [16].
9. To reduce the viscosity of the sample, it should be diluted with water prior to AMIS measurements.
10. The monitoring system indicated 8.9 pg/100 μL and 30.4 pg/100 μL of UbcH8 activity in the culture supernatant of NB4 cells (1×10^5 cells/mL) following treatment with 10 nM and 100 nM bortezomib, respectively [23].

Acknowledgments

This research was supported by A-STEP from Japan Science and Technology Agency (JST), a Grant-in-Aid for Scientific Research (KAKENHI 26430147), Takeda Science Foundation, Sanyo Chemical Industries Foundation, and Nakatani Foundation.

Author contributions
K.M. designed this study. K.M. and K.S. wrote the main manuscript text and prepared all the figures.

Competing interests
The authors declare no competing financial interests.

References

1. Haas AL, Siepmann TJ (1997) Pathways of ubiquitin conjugation. FASEB J 11:1257–1268
2. Hershko A, Ciechanover A (1998) The ubiquitin system. Annu Rev Biochem 67:425–479
3. Weissman AM (2001) Themes and variations on ubiquitylation. Nat Rev Mol Cell Biol 2:169–178
4. Varshavsky A (1997) The ubiquitin system. Trends Biochem Sci 22:383–387
5. Deshaies RJ, Joazeiro CA (2009) RING domain E3 ubiquitin ligases. Annu Rev Biochem 78:399–434
6. Borden KL (2000) RING domains: master builders of molecular scaffolds? J Mol Biol 295:1103–1112
7. Freemont PS (1993) The RING finger. A novel protein sequence motif related to the zinc finger. Ann N Y Acad Sci 684:174–192
8. Takenokuchi M, Miyamoto K, Saigo K, Taniguchi T (2015) Bortezomib causes ER stress-related death of acute promyelocytic leukemia cells through excessive accumulation of PML–RARA. Anticancer Res 35:3307–3316
9. Ueki T, Park JH, Nishidate T, Kijima K, Hirata K, Nakamura Y, Katagiri T (2009) Ubiquitination and downregulation of BRCA1 by ubiquitin-conjugating enzyme E2T overexpression in human breast cancer cells. Cancer Res 69(22):8752–8760
10. Snoek BC, de Wilt LH, Jansen G, Peters GJ (2013) Role of E3 ubiquitin ligases in lung cancer. World J Clin Oncol 4(3):58–69
11. Chen S, Chen Y, Hu C, Jing H, Cao Y, Liu X (2010) Association of clinicopathological features with UbcH10 expression in colorectal cancer. J Cancer Res Clin Oncol 136:419–426
12. Liani E, Eyal A, Avraham E, Shemer R, Szargel R, Berg D, Bornemann A, Riess O, Ross CA, Rott R, Engelender S (2004) Ubiquitylation of synphilin-1 and alpha-synuclein by SIAH and its presence in cellular inclusions and Lewy bodies imply a role in Parkinson's disease. Proc Natl Acad Sci U S A 101:5500–5505
13. Kramer OH, Muller S, Buchwald M, Reichardt S, Heinzel T (2008) Mechanism for ubiquitylation of the leukemia fusion proteins AML1-ETO and PML-RARalpha. FASEB J 22:1369–1379
14. Lee JT, Wheeler TC, Li L, Chin LS (2008) Ubiquitination of alpha-synuclein by Siah-1 promotes alpha-synuclein aggregation and apoptotic cell death. Hum Mol Genet 17:906–917
15. Brown NG, Watson ER, Weissmann F, Jarvis MA, VanderLinden R, Grace CR, Frye JJ, Qiao R, Dube P, Petzold G, Cho SE, Alsharif O, Bao J, Davidson IF, Zheng JJ, Nourse A, Kurinov I, Peters JM, Stark H, Schulman BA (2014) Mechanism of polyubiquitination by human anaphase-promoting complex: RING repurposing for ubiquitin chain assembly. Mol Cell 56:246–260
16. Miyamoto K (2012) Ubiquitination of an artificial RING finger without a substrate and a tag. J Pept Sci 18:135–139
17. Miyamoto K, Togiya K (2010) The creation of the artificial RING finger from the cross-brace zinc finger by alpha-helical region substitution. Biochem Biophys Res Commun 394:972–975
18. Pascual J, Martinez-Yamout M, Dyson HJ, Wright PE (2000) Structure of the PHD zinc finger from human Williams-Beuren syndrome transcription factor. J Mol Biol 304:723–729
19. Matsuzawa S, Li C, Ni CZ, Takayama S, Reed JC, Ely KR (2003) Structural analysis of Siah1 and its interactions with Siah-interacting protein (SIP). J Biol Chem 278:1837–1840
20. Miyamoto K (2014) Structural model of ubiquitin transfer onto an artificial RING finger as an E3 ligase. Sci Rep 4:6574
21. Markson G, Kiel C, Hyde R, Brown S, Charalabous P, Bremm A, Semple J, Woodsmith J, Duley S, Salehi-Ashtiani K, Vidal M, Komander D, Serrano L, Lehner P, Sanderson CM (2009) Analysis of the human E2 ubiquitin conjugating enzyme protein interaction network. Genome Res 19:1905–1911
22. van Wijk SJ, de Vries SJ, Kemmeren P, Huang A, Boelens R, Bonvin AM, Timmers HT (2009) A comprehensive framework of E2-RING E3 interactions of the human ubiquitin-proteasome system. Mol Syst Biol 5:295
23. Miyamoto K, Sumida M, Yuasa-Sunagawa M, Saito K (2017) Highly sensitive detection of E2 activity in ubiquitination using an artificial RING finger. J Pept Sci 23:222–227
24. Gu Y, Bouwman P, Greco D, Saarela J, Yadav B, Jonkers J, Kuznetsov SG (2014) Suppression of BRCA1 sensitizes cells to proteasome inhibitors. Cell Death Dis 5:e1580
25. Niewerth D, Dingjan I, Cloos J, Jansen G, Kaspers G (2013) Proteasome inhibitors in acute leukemia. Expert Rev Anticancer Ther 13:327–337
26. Ying M, Zhou X, Zhong L, Lin N, Jing H, Luo P, Yang X, Song H, Yang B, He Q (2012) Bortezomib sensitizes human acute myeloid

leukemia cells to all-trans-retinoic acid-induced differentiation by modifying the RARalpha/STAT1 axis. Mol Cancer Ther 12:195–206

27. Finkel R, Clark MA, Cubeddu LX (2009) Lippincott's illustrated reviews: pharmacology, fourth edition pharmacology. Lippincott Williams & Wilkins, Philadelphia
28. van Ree JH, Jeganathan KB, Malureanu L, van Deursen JM (2010) Overexpression of the E2 ubiquitin-conjugating enzyme UbcH10 causes chromosome missegregation and tumor formation. J Cell Biol 188:83–100
29. Wu X, Zhang W, Font-Burgada J, Palmer T, Hamil AS, Biswas SK, Poidinger M, Borcherding N, Xie Q, Ellies LG, Lytle NK, Wu LW, Fox RG, Yang J, Dowdy SF, Reya T, Karin M (2014) Ubiquitin-conjugating enzyme Ubc13 controls breast cancer metastasis through a TAK1-p38 MAP kinase cascade. Proc Natl Acad Sci U S A 111:13870–13875
30. Brzovic PS, Rajagopal P, Hoyt DW, King MC, Klevit RE (2001) Structure of a BRCA1-BARD1 heterodimeric RING-RING complex. Nat Struct Biol 8:833–837
31. Katoh S, Hong C, Tsunoda Y, Murata K, Takai R, Minami E, Yamazaki T, Katoh E (2003) High precision NMR structure and function of the RING-H2 finger domain of EL5, a rice protein whose expression is increased upon exposure to pathogen-derived oligosaccharides. J Biol Chem 278:15341–15348
32. Kostic M, Matt T, Martinez-Yamout MA, Dyson HJ, Wright PE (2006) Solution structure of the Hdm2 C2H2C4 RING, a domain critical for ubiquitination of p53. J Mol Biol 363:433–450
33. Miyamoto K, Uechi A, Saito K (2017) The zinc finger domain of RING finger protein 141 reveals a unique RING fold. Protein Sci 26:1681–1686
34. Katoh S, Tsunoda Y, Murata K, Minami E, Katoh E (2005) Active site residues and amino acid specificity of the ubiquitin carrier protein-binding RING-H2 finger domain. J Biol Chem 280:41015–41024
35. Scheel H, Hofmann K (2003) No evidence for PHD fingers as ubiquitin ligases. Trends Cell Biol 13:285–287; author reply 287–288
36. Aravind L, Iyer LM, Koonin EV (2003) Scores of RINGS but no PHDs in ubiquitin signaling. Cell Cycle 2:123–126
37. Chakravarty S, Zeng L, Zhou MM (2009) Structure and site-specific recognition of histone H3 by the PHD finger of human autoimmune regulator. Structure 17:670–679
38. Pena PV, Davrazou F, Shi X, Walter KL, Verkhusha VV, Gozani O, Zhao R, Kutateladze TG (2006) Molecular mechanism of histone H3K4me3 recognition by plant homeodomain of ING2. Nature 442:100–103
39. Szargel R, Rott R, Eyal A, Haskin J, Shani V, Balan L, Wolosker H, Engelender S (2009) Synphilin-1A inhibits seven in absentia homolog (SIAH) and modulates alpha-synuclein monoubiquitylation and inclusion formation. J Biol Chem 284:11706–11716

Chapter 7

Directed Evolution of Targeted Recombinases for Genome Engineering

Shannon J. Sirk

Abstract

Over the past several years, genome engineering has become an established component of basic research endeavors, and is emerging as a vital element of clinical research applications. Site-specific recombinases are one of the several tools that can facilitate genome modification by catalyzing rearrangements between specific DNA targets. Of particular interest are the small serine recombinases, which are modular in both form and function. This unique structure permits replacement of the native DNA-binding domain with designer targeting modules such as zinc fingers, TALEs, or catalytically inactivated Cas9, enabling modification of investigator-defined genomic loci. Importantly, the catalytic domain of these enzymes also contributes to target specificity, and can be reprogrammed to recognize custom sequences for genomic targeting. Here we describe the steps required to construct, select, and validate hybrid recombinase catalytic domains for targeted genome engineering.

Key words Recombinase, Genome editing, Genetic engineering, Protein engineering, Targeted integration, Site-specific genomic modification, Directed evolution, Rational design

1 Introduction

In order to meet the demand for highly precise genome editing, numerous tools have been developed with increasingly sophisticated capabilities. Among these are site-specific recombinases: enzymes that can perform integration, inversion, or deletion between specific DNA target sites [1]. Unlike targeted nucleases, which cleave DNA and then rely on cellular machinery to repair the resulting double-strand breaks, these enzymes catalyze recombination autonomously. Because recombinases orchestrate DNA cleavage and strand exchange as a concerted process, the genomic outcome can be controlled, unlike nuclease cleavage products which can be subject to random nucleotide insertions or deletions. The resolvase/invertase family of small serine recombinases features a modular design, with structurally and functionally distinct DNA-binding and catalytic

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_7,

domains [2]. Engineering efforts have capitalized on this modularity via the fusion of recombinase catalytic domains with designer DNA-binding domains such as zinc-finger [3–5], TALE [6], or catalytically dead Cas9 proteins, known as dCas9 [7]. Because these DNA-binding platforms are programmable, hybrid recombinases can be engineered to target a broad range of genomic sequences.

An additional layer of recombinase target specificity lies in the catalytic domain, which contributes to recognition through interactions with the minor groove of the target-site DNA [8]. The cooperative specificity of the DNA-binding and catalytic domains ensures that targeted recombinases only act on genomic sites harboring the corresponding recognition sites for both modules. Because recombinases are responsible for regulating many essential cellular functions in their native hosts, their catalytic domains have evolved very strict specificities for their DNA targets [2]. Thus, while the unique modular structure of small serine recombinases enables the straightforward exchange of DNA-binding domains, more sophisticated engineering approaches have been required to alter the stringent specificity of the catalytic domain [9].

To address this challenge, we have identified residues within the recombinase catalytic domain that coordinate target specificity [10–12]. By combining substrate specificity profiling with directed evolution, we generated a panel of engineered catalytic domains capable of recognizing distinct DNA targets [10, 11]. Because these enzymes recombine DNA as dimers, with each monomer recognizing one half of the target sequence, two different engineered recombinase monomers must be paired in order to recognize asymmetric target sites. To prevent nonspecific activity, we also modified the catalytic domain dimerization interface to prevent the formation of homodimer recombinases that can recognize off-target sites [13].

Engineered hybrid recombinases thus represent a promising tool for targeted genomic engineering. Expansion of the current recombinase toolbox to include additional enzymes with complementary targeting specificity will further increase the utility of this approach [12, 14]. To that end, we provide here step-by-step procedures for the creation of recombinase catalytic domains with investigator-defined specificity, using the following approach: (1) randomization of critical DNA-binding residues in the recombinase catalytic domain, (2) selection of active variants using a bacterial genetic screen, and (3) evaluation of the genome-modifying activity of the evolved enzymes in human cells.

2 Materials

2.1 Randomization of Critical DNA-Binding Residues in the Recombinase Catalytic Domain for Construction of Hybrid Recombinase Library

1. DNA sequence of recombinase of interest (*see* **Note 1**).
2. DNA and protein sequence alignment software (*see* **Note 2**).
3. Recombinase of interest gene cassette (*see* **Note 3**).
4. Primers to amplify recombinase library (described in detail in Subheading 3).
5. H1 zinc-finger gene cassette sequence (*see* **Note 4**).
6. H1 zinc-finger gene cassette (*see* **Note 4**).
7. H1 zinc-finger primers ZF Restore Prim1 and pUC18Prim2 [5].
8. Thermal cycler.
9. PCR reagents (*see* **Note 5**).
10. DNA gel electrophoresis equipment, reagents, and supplies.
11. DNA gel purification kit/reagents.
12. PCR purification kit/reagents.
13. 1.5 mL Eppendorf tubes.
14. 0.2 mL Thin-wall PCR tubes.
15. Microcentrifuge.

2.2 Selection of Active Variants Using a Bacterial Genetic Screen

1. Recombinase catalytic domain library (from Subheading 2.1).
2. DNA sequence of genomic target site.
3. Split-beta lactamase reporter plasmid pBLA [15] (*see* **Note 6**).
4. Split-beta lactamase reporter plasmid sequence (*see* **Note 6**).
5. Primers encoding mirrored symmetrical target sites (described in detail in Subheading 3; *see* **Note 7**).
6. Thermal cycler.
7. PCR reagents (*see* **Note 5**).
8. 1.5 mL Eppendorf tubes.
9. 0.2 mL Thin-wall PCR tubes.
10. DNA gel electrophoresis reagents and equipment.
11. DNA gel purification kit/reagents.
12. PCR purification kit/reagents.
13. Restriction enzymes (SacI, XbaI, HindIII, SpeI; *see* **Note 8**).
14. T4 DNA ligase and buffer.
15. Microcentrifuge spin filter, 0.45 μm.
16. Molecular biology grade ethanol, 200 proof.
17. 3 M Sodium acetate, pH 5.2–6.0.
18. Electrocompetent *E. coli* Top10F' cells (*see* **Note 9**).

19. Electroporator.
20. Electroporation cuvettes for bacterial transformation.
21. Super Optimal Broth with Catabolite suppression (SOC): 2% w/v Tryptone, 0.5% w/v yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM $MgCl_2$, 10 mM $MgSO_4$, 20 mM glucose, in water, sterilized. Autoclave before adding glucose, filter sterilize after adding glucose.
22. Super broth (SB): 3.5% w/v Tryptone, 2% w/v yeast extract, 0.5% w/v NaCl, 1 M NaOH, in water, sterilized.
23. Lysogeny broth (LB): 1% w/v Tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, in water, sterilized.
24. 100× Chloramphenicol: 30 mg/mL in ethanol (*see* **Note 10**).
25. 100× Carbenicillin: 100 mg/mL in water, filter sterilized (*see* **Note 11**).
26. LB solid media plates: 1% w/v Tryptone, 0.5% w/v yeast extract, 1% w/v NaCl, 2% agar, in water, sterilized:
    (a) With 30 μg/mL chloramphenicol.
    (b) With 100 μg/mL carbenicillin.
    (c) With 30 μg/mL carbenicillin and 100 μg/mL chloramphenicol.
27. Miniprep kit/reagents.
28. Maxiprep kit/reagents.
29. DNA sequencing services.

2.3 Evaluation in Human Cells via Genomic Integration of Donor Plasmid

1. Left and right evolved recombinase catalytic domain monomers selected in Subheading 2.2 against left and right genomic half sites.
2. Primers to amplify evolved recombinase variants (described in detail in Subheading 3).
3. Gene cassettes and sequences for genomic target-specific DNA-binding domains (*see* **Note 4**).
4. Primers to amplify selected DNA-binding domains.
5. Mammalian expression plasmid pcDNA 3.1 (*see* **Note 12**).
6. Mammalian integration donor plasmid with target-site-flanked gene cassette (*see* **Note 13**).
7. PCR reagents (*see* **Note 4**).
8. Thermal cycler.
9. 1.5 mL Eppendorf tubes.
10. 0.2 mL Thin-wall PCR tubes.
11. DNA gel electrophoresis equipment, reagents, and supplies.
12. Restriction enzymes for subcloning into pcDNA3.1 (*see* **Note 8**).

13. T4 DNA ligase.
14. Electrocompetent *E. coli* Top10F' cells (*see* **Note 9**).
15. Electroporation cuvettes for bacterial transformation.
16. Electroporator.
17. Super Optimal Broth with Catabolite suppression (SOC): *See* Subheading 2.2, **step 21**.
18. Lysogeny broth (LB): *See* Subheading 2.2, **step 23**.
19. 100× Carbenicillin: *See* Subheading 2.2, **step 25**.
20. LB plates with 100 μg/mL carbenicillin: *See* Subheading 2.2, **step 26**.
21. Miniprep kits/reagents.
22. Maxiprep kit/reagents.
23. DNA sequencing services.
24. Biosafety cabinet for mammalian tissue culture.
25. Mammalian target cells (*see* **Note 14**).
26. Growth medium for target cell line (*see* **Note 15**).
27. Humidified CO_2 incubator for mammalian cell culture.
28. Inverted microscope capable of ×10–40 magnification.
29. Hemocytometer and coverslips.
30. 24-Well tissue culture plates, sterile.
31. Genomic isolation kit (*see* **Note 16**).
32. Primers for assessing genomic integration (*see* **Note 17**).

3 Methods

The following procedures outline an approach for creating hybrid recombinases with zinc-finger DNA-binding domains; however, any DNA-binding domain of interest can be substituted (*see* **Note 4**). Prior to initiating the procedures described below, a recombinase candidate and corresponding genomic target site must be selected. Native recombinases require the introduction of hyper-activating mutations in order to function in the absence of accessory proteins. Such mutations have been described for Gin [16], Tn3 [17], γ-δ [18], β [12], Sin [12, 19], Hin [16], Bin [14], and Tn21 [14]. Hybrid recombinase target-site selection is described in detail elsewhere [20].

3.1 Randomization of Critical DNA-Binding Residues in the Recombinase Catalytic Domain for Construction of Hybrid Recombinase Library

1. Select recombinase catalytic domain residues to target for mutagenesis by aligning the gene sequence of the recombinase of interest with the gene sequence of Gin-alpha (the native Gin catalytic domain), and then identifying the residues that correspond to Gin Ile 120, Thr 123, Leu 127, Ile 136, and Gly 137.
2. Design PCR primers to amplify the recombinase gene and introduce mutations to generate a catalytic domain library. In order to encode all of the diversity in the library, the reverse primer will span a large segment of the 3′ end of the gene and will therefore be long (~100 bp; *see* **Note 18**). The design of forward and reverse primers is detailed in the following two steps.
3. Design the forward primer to amplify the recombinase gene beginning at the start codon, with a 5′ extension that includes a SacI cleavage site for cloning into the pBLA split-gene reassembly vector (*see* **Note 6**). Follow standard primer design guidelines.
4. Design the reverse primer to amplify the recombinase gene beginning (in the reverse direction) at the stop codon, with a 3′ extension that includes an overlap of ~15 bp for in-frame fusion with the H1 zinc-finger cassette. The 5′ end of the primer should lie ~10 bp upstream of the first residue that is targeted for randomization (i.e., the residue corresponding to Gin Ile 120). The primer sequence should match the recombinase gene sequence exactly, except for the residues selected for randomization. For these codons, use the IUPAC nucleotide code NNK, where "N" represents all nucleotides and "K" represents G or T (*see* **Note 19**).
5. PCR amplify the recombinase gene using the primers designed above and PCR amplify the H1 zinc-finger cassette using ZF Restore Prim1 and pUC18Prim2 (*see* **Note 4**). Carry out the PCR reactions using 50 ng of template DNA, 0.5 μM of each primer, 0.2 mM of dNTPs, and 1 U of high-fidelity DNA polymerase, with the appropriately diluted PCR buffer and water to 50 μL. Follow the manufacturer's guidelines for denaturation, annealing, and extension; perform 30 cycles.
6. Run the PCR products on an agarose gel, cut out, and gel purify appropriate bands.
7. Perform overlap extension PCR to fuse the recombinase library to the H1 zinc-finger cassette. A total reaction volume of 500 μL is used in order to generate sufficient PCR product for downstream cloning steps (*see* **Note 20**). Assemble the entire reaction, and then separate into 50 μL aliquots in 0.2 mM PCR tubes for thermal cycling. Carry out the PCR reaction using 500 ng of each template DNA, 0.5 μM of Fwd primer

from **step 3** in Subheading 3.1, 0.5 μM of pUC18Prim2 primer, 0.2 mM of dNTPs, and 10 U of high-fidelity DNA polymerase, with the appropriately diluted PCR buffer and water to 500 μL. Follow the manufacturer's guidelines for denaturation, annealing, and extension, keeping in mind that the total template length is equal to the length of the recombinase template plus the length of the H1 zinc-finger template. Perform 20 cycles.

8. Purify the PCR product using a PCR purification kit.
9. The PCR-purified hybrid recombinase library can be stored frozen at −20 °C for several months.

3.2 Selection of Active Variants Using a Bacterial Genetic Screen

1. Construct two pBLA plasmids (*see* **Note 6**) for selection of the left and right recombinase monomers on mirrored symmetrical half sites. The procedures are detailed below in **steps 2–9**.
2. Design primers that will amplify the GFPuv [15] stuffer fragment while adding recombinase target sites at both the 5′ and 3′ ends. The final amplicon should have a XbaI site at the 5′ end and a HindIII site at the 3′ end. The target site for selection of the recombinase monomer recognizing the left half of the genomic target sequence should consist of a mirrored, symmetrical version of the left half of the genomic target site, and the 5′ target site and 3′ target site should be identical to each other. The same procedure should be followed for the right half of the genomic target site, creating a separate pBLA selection vector. Primers will be long. Carry out the PCR reaction as in **step 5** in Subheading 3.1, using 0.4 μM of each primer.
3. In two separate reactions, PCR amplify GFPuv from pBLA for the right and left target sites, using the primers designed above in **step 2**.
4. Run the PCR products on an agarose gel, cut out, and gel purify appropriate bands. The GFPuv gene is approximately 750 bp.
5. Digest both PCR products with XbaI and HindIII. Also digest pBLA vector with SpeI and HindIII.
6. Run the PCR products on an agarose gel, cut out, and gel purify appropriate bands.
7. In separate reactions, ligate the left and right GFPuv inserts into the digested pBLA vector. Use 50–100 ng vector, x ng insert (use a 1:6 molar ratio of vector:insert; calculate based on the size of each piece of DNA), 10 UT4 DNA ligase, 1× buffer, and water up to 10 μL.
8. Transform ligation reactions into *E. coli* cells. Use 2 μL of ligation reaction and 50 μL of competent cells (*see* **Note** 9),

recover in 1 mL SOC media, and plate 200 μL of each on individual LB-chlor plates.

9. To screen clones, pick several single colonies into LB-chlor minicultures, miniprep, and validate positive clones by sequencing analysis.
10. Digest library from Subheading 3.1 and two pBLA vectors from **step 9** with SacI and XbaI. Digest at least 10 μg of each pBLA vector and at least 5 μg of the library.
11. Run the digested products on an agarose gel and purify by freeze-squeeze (*see* **Note 21**). The procedures of freeze-squeeze are detailed in **steps 12–22**.
12. Freeze gel slice at −20 °C for at least 30 min or at –80 °C for at least 10 min.
13. Construct a Parafilm pocket by folding a small piece in half, and then folding and sealing the edges by pressing them together with the conical end of a 1.5 mL Eppendorf tube. Insert frozen gel slice into envelope, fold, and seal shut.
14. Using the flat surface of a closed 1.5 mL Eppendorf tube, extensively crush the frozen gel slice inside the pouch.
15. Load the crushed gel and associated liquid into a micro spin filter unit (0.45 μm).
16. Centrifuge at 6000–7000 rpm in a microcentrifuge for 5 min at 4 °C. The flow-through contains the DNA.
17. Ethanol precipiate DNA from flow-through by first measuring the volume of the flow-through using a micropippetor. Add 3 M sodium acetate equivalent to 1/10 the volume of the flow-through and mix. Add two volumes of ethanol to the flow-through, mix, and place at 4 °C for 2 h.
18. Centrifuge at max speed in a microcentrifuge for 30 min at 4 °C.
19. Decant the supernatant.
20. Centrifuge at max speed for ~1 min.
21. Carefully remove remaining ethanol with micropipette. Allow the DNA pellet to air-dry.
22. Resuspend DNA in 10–50 μL of water.
23. In separate reactions, ligate the digested library into (1) the digested pBLA-right selection vector and (2) the digested pBLA-left selection vector. Use 6 μg vector, x μg insert (use a 1:6 molar ratio of vector:insert; calculate based on the size of each piece of DNA), 10 U T4 DNA ligase, 1× buffer, and water up to the lowest volume necessary to accommodate the reaction.
24. Transform each library into electrocompetent *E. coli* by electroporation. For each library, perform six separate electroporations, using 1 μg of library and 300 μL of cells per sample.

Recover in 3 mL SOC. Plate 200 μL of each recovery culture on an LB-chlor plate for determination of transformation efficiency and library size. As an optional choice, assess library diversity by miniprepping and sequencing 10–20 colonies from these plates. To the remainder of the recovery cultures, add 45 mL SB-chlor to each, and incubate for 6–16 h.

25. Maxiprep cultures (round I libraries).
26. To perform the first round of selection, transform 3 μg of each maxiprepped library by electroporation, as in **step 24**.
27. To determine transformation efficiency (TE), plate the recovery culture from **step 26** on an LB-chlor plate. To ensure that the number of colonies on the plate is not too many to permit counting, it will be necessary to dilute the recovery culture.
28. To determine the percent of active clones in the library ($TE_{chlor/carb}/TE_{chlor}$), plate the recovery culture from **step 26** on an LB-chlor/carb plate. The recovery culture should also be diluted here, as above, but it is recommended to plate a lower dilution, as the number of active clones will be less than the number of total clones.
29. Miniprep and sequence individual active recombinase colonies. Analyze the resulting sequences to determine if the library contains many clones with the same amino acid substitutions. If such a consensus emerges, analyze the recombinase variant(s) for activity using the corresponding pBLA target vector. Digest the recombinase variant with SacI and Xba I and ligate into pBLA vector as described in **steps 10–23**. Transform and plate on LB-chlor and LB-chlor/carb plates as described above and calculate the percent activity of each variant ($TE_{chlor/carb}/TE_{chlor}$). Highly active recombinases demonstrate ~10–50% recombination.
30. To the remainder of recovery cultures from **step 26**, add 45 mL SB-chlor/carb to each, incubate for 6–16 h, and maxiprep to build the round II libraries. To increase the stringency of the selection step, decrease the incubation time.
31. Repeat **steps 26–30** in Subheading 3.2 until a consensus sequence is reached and/or highly active recombinase variants are identified.

3.3 Evaluation in Human Cells

1. Construct a pair of hybrid recombinases to recognize the selected genomic target sequence. The procedures are detailed in **step 2**.
2. Design primers to amplify the evolved catalytic domains from Subheading 3.2, adding restriction sites for cloning into pcDNA3.1 at the 5′ end and an overlap extension for PCR fusion to the selected DNA-binding domains at the 3′ end.
3. Design primers to amplify the selected DNA-binding domains (*see* **Note 4**), adding an overlap extension for PCR fusion to

the evolved recombinase catalytic domains at the 5′ end of the gene and restriction sites for cloning into pcDNA3.1 at the 5′ and the 3′ ends of the gene.

4. In separate reactions, PCR amplify the left and right evolved recombinase catalytic domains generated in Subheading 3.2 and the primers generated above in **step 2**. Carry out the PCR reactions using 50 ng of template DNA, 0.5 μM of each primer, 0.2 mM of dNTPs, and 1 U of high-fidelity DNA polymerase, with the appropriately diluted PCR buffer and water to 50 μL. Follow the manufacturer's guidelines for denaturation, annealing, and extension; perform 30 cycles.
5. At the same time, PCR amplify DNA-binding domains appropriate for genomic target site. Numerous protocols, tools, and online resources exist for designing and assembling custom zinc fingers, TALEs, and catalytically inactivated Cas9 constructs. Carry out the PCR reaction as above in **step 4**.
6. Run the PCR products on an agarose gel, cut out the pieces, and purify using a gel purification kit.
7. Perform fusion PCR as in **step 7** in Subheading 3.1.
8. Purify using a PCR purification kit as described above.
9. Digest fusion PCR products and pcDNA vector with chosen restriction enzymes.
10. Run digested products on agarose gel, cut out appropriate pieces, and purify using a gel purification kit.
11. Ligate, transform, and sequence verify clones as described in Subheading 3.2.
12. Prepare DNA for transfection into mammalian target cells. Select a sequence-verified clone from **step 11** and inoculate a 50 mL maxiprep culture. At the same time, start a maxiprep culture of the selected donor vector. Maxiprep both constructs following the manufacturer's instructions.
13. Seed target mammalian cells in a 24-well plate at 1×10^5 cells/well in cell culture media with appropriate additives (*see* **Note 15**). Maintain at 37 °C, 5% CO_2, in a humidified growth chamber.
14. After 24 h, transfect each well with 10 ng of pcDNA-right-recombinase, 10 ng of pcDNA-left-recombinase, and 80 ng of the selected integration donor plasmid using Lipofectamine 2000, following the manufacturer's instructions.
15. After 72 h, harvest cells from plate and extract genomic DNA.
16. Assess genomic integration by PCR (*see* **Note 17**).

4 Notes

1. Construction of hybrid recombinases has been described in detail elsewhere [20]. In brief, for a given genomic target, select a candidate recombinase with catalytic specificity that closely aligns with the genomic target, particularly at base positions 1–6 (counting from the center of the 20 bp symmetric core) [11].
2. Several online tools are freely available to align and analyze DNA and amino acid sequences (e.g., https://www.ebi.ac.uk/Tools/msa/clustalo/).
3. Active variants are available from Addgene and/or from the publishing laboratories.
4. H1 zinc fingers [5], as well as other DNA-binding domains, and sequences can be obtained from Addgene. For selection, validation, and characterization of the recombinase catalytic domains, it is recommended to use the well-characterized H1 zinc fingers as placeholder DNA-binding domains. H1 zinc fingers reliably bind their target sites, ensuring that the performance of the DNA-binding module is not a variable during recombinase catalytic domain selection and optimization. If constructing hybrid recombinases with DNA-binding domains other than H1 zinc fingers (such as TALEs or catalytically inactivated Cas9), different linker lengths will be required and the H1 zinc-finger target sites described in Subheading 2.2, **step 5**, will need to be modified.
5. It is recommended to use a standard *Taq* polymerase for library construction and a high-fidelity polymerase for all other amplifications.
6. The bacterial genetic screen relies on the concept of substrate-linked protein evolution (SLiPE) [21] to screen for active recombinase variants. Using a split-gene reassembly method, individual recombinase variants are placed into the pBLA plasmid, which has chloramphenicol resistance as well as an ampicillin/carbenicillin resistance marker that is disrupted by an inserted GFPuv gene cassette, and is therefore inactive. The inserted DNA is flanked by the recombinase target sites. When transformed into bacteria, an active recombinase will (1) be expressed from the plasmid, (2) recognize the target sites on that same plasmid, (3) remove the DNA insertion, and (4) recombine the plasmid backbone. Because recombinase-mediated DNA removal results in precise and predictable DNA sequence outcomes, the insertion was designed such that recombination renders the antibiotic resistance cassette functional. Growth in media containing both chloramphenicol and ampicillin/carbenicillin thus leads to survival of bacteria harboring active recombinase variant gene sequences.

7. To avoid selecting for relaxed target specificity, recombinases must be selected as homodimers on symmetric target sites. Because genomic target sites are generally asymmetric, two different recombinase monomers are required, each recognizing one half of the full target sequence. It is therefore necessary to generate symmetrical, mirrored versions of each half of the genomic target and select each monomer independently.
8. The protocols presented here describe cloning procedures using restriction enzyme digestion and DNA ligation. If not specified, any restriction sites within a plasmid's multiple cloning site may be selected by the investigator to best accommodate the cloning approach. Alternative methods such as Gibson assembly may be used.
9. In order to achieve transformation efficiencies high enough to accommodate library construction with tenfold sequence coverage, it is recommended to use electrocompetent, rather than chemically competent, cells. Chemically competent cells may be used for other cloning steps.
10. Chloramphenicol stock solution must be dissolved in 70–100% ethanol.
11. The use of carbenicillin, which is a more stable analog of ampicillin, is recommended to ensure that selective pressure is maintained within the context of the split-gene reassembly approach.
12. This plasmid is available from Invitrogen. The methods presented here were developed using this plasmid, but the procedures can be easily modified to accommodate other mammalian expression vectors.
13. Donor plasmid construction is described in detail elsewhere [11, 13]. Briefly, the donor integration vector must contain an exact copy of the genomic target site somewhere within the plasmid.
14. The methods described here were developed with HEK293 cells, and subsequently expanded to other cell types [22]. When developing genomic targeting protocols for common laboratory cell lines, be aware that many lines have mutated over years of continuous passaging, and therefore published genomic sequences may not match individual lab strains. It is recommended to purchase a new vial of cells for these experiments.
15. Different cell lines require different media formulations. These methods were developed using HEK293 cells, which grow in Dulbecco's modified Eagle's medium (DMEM), supplemented with fetal bovine serum (FBS; 10% v/v). It is also recommended to add antibiotics (e.g., Anti-Anti antibiotic/antimycotic, Gibco) to prevent microbial contamination.

16. These methods were developed using Quick Extract DNA Extraction Solution (Epicentre).
17. To determine if the donor plasmid has integrated, and ensure that the PCR reaction is not amplifying extrachromosomal plasmid DNA, use one primer that lies within the integration donor plasmid (at least 100 bp away from the integration junction) and another primer that lies in the cellular genome, ~100 bp away from the integration junction.
18. While it is generally recommended to perform PCR using primers with similar lengths and melting temperatures, amplification using one long primer and one standard-length primer has worked well, with no need for optimization.
19. The degenerate triplet NNK encodes all 20 amino acids, with only one stop codon possible.
20. It is necessary to produce large quantities of amplicon to ensure that, after downstream processing, there is enough digested and purified product to generate a large-scale ligation reaction. The large-scale reaction is needed to ensure adequate coverage of the theoretical library size (the number of possible different variants based on the randomized positions), to minimize bias during the selection rounds. It is recommended to have ten-fold library coverage, that is to say, that there are ten times as many clones as there are possible different sequences. While a single-PCR reaction is capable of producing such coverage, the practical limitation of the approach lies in the transformation efficiency of the competent cells, requiring the whole procedure to be done at a large scale.
21. Purification of DNA by freeze-squeeze is recommended because the gel fragment may be quite large and would require several gel purification columns. Freeze-squeeze may also result in higher yield and cleaner product.

References

1. Grindley NDF, Whiteson KL, Rice PA (2006) Mechanisms of site-specific recombination. Annu Rev Biochem 75:567–605
2. Smith MCM, Thorpe HM (2002) Diversity in the serine recombinases. Mol Microbiol 44:299–307
3. Akopian A, He J, Boocock MR et al (2003) Chimeric recombinases with designed DNA sequence recognition. Proc Natl Acad Sci U S A 100:8688–8691
4. Gordley RM, Gersbach CA, Barbas CF III (2009) Synthesis of programmable integrases. Proc Natl Acad Sci U S A 106:5053–5058
5. Gordley RM, Smith JD, Gräslund T et al (2007) Evolution of programmable zinc finger-recombinases with activity in human cells. J Mol Biol 367:802–813
6. Mercer AC, Gaj T, Fuller RP et al (2012) Chimeric TALE recombinases with programmable DNA sequence specificity. Nucleic Acids Res 40:11163–11172
7. Chaikind B, Bessen JL, Thompson DB et al (2016) A programmable Cas9-serine recombinase fusion protein that operates on DNA sequences in mammalian cells. Nucleic Acids Res 44:9758–9770
8. Yang W, Steitz TA (1995) Crystal structure of the site-specific recombinase gamma delta

resolvase complexed with a 34 bp cleavage site. Cell 82:193–207

9. Gaj T, Sirk SJ, Barbas CF (2014) Expanding the scope of site-specific recombinases for genetic and metabolic engineering. Biotechnol Bioeng 111:1–15
10. Gaj T, Mercer AC, Gersbach CA et al (2011) Structure-guided reprogramming of serine recombinase DNA sequence specificity. Proc Natl Acad Sci U S A 108:498–503
11. Gaj T, Mercer AC, Sirk SJ et al (2013) A comprehensive approach to zinc-finger recombinase customization enables genomic targeting in human cells. Nucleic Acids Res 41:3937–3946
12. Sirk SJ, Gaj T, Jonsson A et al (2014) Expanding the zinc-finger recombinase repertoire: directed evolution and mutational analysis of serine recombinase specificity determinants. Nucleic Acids Res 42:4755–4766
13. Gaj T, Sirk SJ, Tingle RD et al (2014) Enhancing the specificity of recombinase-mediated genome engineering through dimer interface redesign. J Am Chem Soc 136:5047–5056
14. Wallen MC, Gaj T, Barbas CF (2015) Redesigning recombinase specificity for safe harbor sites in the human genome. PLoS One 10:e0139123
15. Gersbach CA, Gaj T, Gordley RM et al (2010) Directed evolution of recombinase specificity by split gene reassembly. Nucleic Acids Res 38:4198–4206
16. Klippel A, Cloppenborg K, Kahmann R (1988) Isolation and characterization of unusual gin mutants. EMBO J 7:3983–3989
17. Arnold PH, Blake DG, Grindley ND et al (1999) Mutants of Tn3 resolvase which do not require accessory binding sites for recombination activity. EMBO J 18:1407–1414
18. Newman BJ, Grindley ND (1984) Mutants of the gamma delta resolvase: a genetic analysis of the recombination function. Cell 38:463–469
19. Rowland S-J, Boocock MR, McPherson AL et al (2009) Regulatory mutations in Sin recombinase support a structure-based model of the synaptosome. Mol Microbiol 74:282–298
20. Gaj T, Barbas CFIII (2014) Genome engineering with custom recombinases. Methods Enzymol 546:79–91
21. Buchholz F, Stewart AF (2001) Alteration of Cre recombinase site specificity by substrate-linked protein evolution. Nat Biotechnol 19:1047–1052
22. Gersbach CA, Gaj T, Gordley RM et al (2011) Targeted plasmid integration into the human genome by an engineered zinc-finger recombinase. Nucleic Acids Res 39:7868–7878

Chapter 8

Optical Recording of Cellular Zinc Dynamics with Zinc-Finger-Based Biosensors

Dylan H. Fudge, Ray Black, and Yan Qin

Abstract

In addition to serving as an essential structural component, zinc is also involved in intracellular and intercellular signaling pathways to impact a number of cellular functions. Genetically encoded zinc sensors that are specifically targeted to various subcellular compartments (ER, mitochondria, nucleus, plasma membrane, and vesicles) have been proven to provide accurate and sensitive visualization and quantification of zinc. Here we describe the methods to utilize both ratiometric and intensiometric genetically encoded zinc sensors designed based on zinc fingers for imaging and quantification of cellular free, labile zinc concentrations, $[Zn^{2+}]_{free}$. This chapter explains in detail how to quantify $[Zn^{2+}]_{free}$ in live cells as well as how to monitor zinc influx in INS-1 cells stimulated with high glucose.

Key words Fluorescent protein, Zinc fingers, Subcellular targeting, Ratiometric zinc sensor, Intensiometric zinc sensor, Imaging

1 Introduction

Zinc is the second most abundant heavy metal within cells and has been bioinformatically linked to over 2800 proteins, yet our understanding of how the interaction between zinc and these proteins modulates various cellular functions is still in its infancy [1]. Emerging studies have shown that zinc affects certain cellular processes including the alteration of enzymatic function [2], regulation of mitochondrial function [3–5], modulation of cell division [6], and facilitation of apoptosis [6]. The desire to further elucidate the role of zinc in these essential cellular functions has driven the development of fluorescent zinc sensors in the past few decades. Two different approaches for sensor development have resulted in the establishment of small-molecule probes and genetically encoded sensors. While both sensors have their distinct advantages, genetically expressed zinc sensors demonstrate advantages in studying the physiological functions of zinc in living cells due to the ability to control the spatial expression of sensors within specific

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_8,

cellular compartments using various molecular tags. Genetically encoded zinc sensors are designed using various metal-sensing domains for zinc including zinc finger 1 (ZF1) and zinc finger 2 (ZF2) from the Zap1 transcription factor of *S. cerevisiae* for the ZapCY, ZapCmR, and GZnP sensors [7–10]; Atox1 and the fourth domain of ATP7B (WD4) for the eCALWY sensors [11, 12]; and simple metal-coordinating amino acids (cysteine and histidine) introduced at the dimerization interface between the two chromophores in eZinCh sensors [13, 14].

This chapter focuses on zinc sensors designed with zinc fingers and their applications in live-cell imaging. The first two zinc fingers from transcription factor Zap1 (ZF1 and ZF2) have been shown to confer conformational changes in response to zinc. Prior to binding of zinc, the ZF1 and ZF2 domains are unstructured linear strings of protein, while upon binding zinc there is a conformation change in the zinc fingers, resulting in a finger-finger interacting complex [15]. Two types of genetically encoded sensors have been developed using this pair of zinc fingers as the metal-sensing domains: FRET-based (ratiometric) and single fluorescent protein-based (intensiometric) sensors. The FRET (Förster resonance energy transfer) based sensor relies on fluorescent resonance energy transfer between two different fluorescent proteins (FP) to indicate if the sensor is bound or unbound to the metal ion. The two different FP chromophores are specific and unique to each other based on their emission and absorbance wavelengths. The zinc-induced conformational changes in ZF1 and ZF2 can bring the donor and acceptor fluorescent proteins closer to each other (typically within 1–10 nm), resulting in increased FRET (Fig. 1a). FRET sensors do not depend on the sensor concentration or thickness of the sample due to the spectral shift in wavelengths typically making these sensors normalized for quantification of free metal ions. There are limitations to using a FRET-based sensor such as the need for different combinations of two excitation and emission filters for image acquisition, spectral bleed through when using multiple sensors, and smaller dynamic ranges correlating to less sensitivity to the metal ion.

Single fluorescent protein-based zinc sensors utilize one chromophore and detect changes in the metal ion concentration by changes in the sensor fluorescent intensity. The single fluorescent protein-based zinc sensor (GZnP1) utilizes ZF1 and ZF2 attached to the N- and C-termini of circularly permutated green fluorescent protein (cpGFP). When bound to zinc both ZF1 and ZF2 form a finger-finger interaction and undergo a conformational change which stabilizes the cpGFP resulting in an increase in fluorescence [8] (Fig. 1b). The advantages of using intensiometric sensors include the ability to use multiple types of sensors simultaneously without spectral bleed through. Also imaging using one wavelength of light requires only one excitation and emission filter decreasing

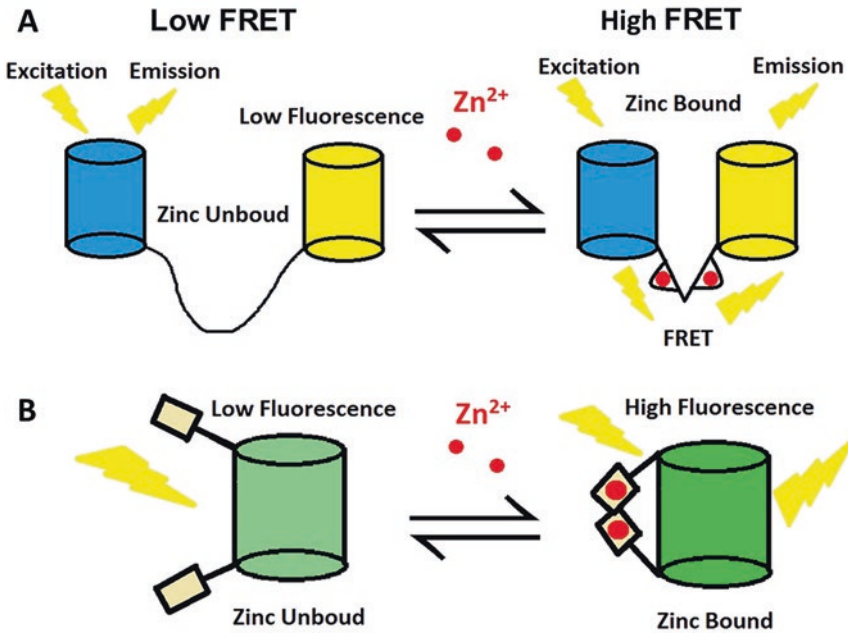


Fig. 1 Sensor design for ratiometric and intensiometric sensors. (**a**) Schematic of a FRET sensor design. The FRET/ratiometric sensor is composed of two separate fluorescent proteins bound via ZF1and ZF2. Upon binding zinc, the conformation changes of ZF1 and ZF2 brought the donor and acceptor fluorescent protein close together. (**b**) Schematic of a single fluorescent protein sensor design. The intensiometric sensor is designed by attaching a pair of zinc fingers to circularly permuted GFP (cpGFP). Zinc-induced conformational change protects the cpGFP fluorophore from solvent attack, resulting in increased fluorescence

the complexity of imaging. Single fluorescent sensors also tend to offer a much higher dynamic range which correlates to more sensitive measurements of the free ions. The purpose of this chapter is to elucidate the application of these two types of sensors (ZapCY, ZapCV and GZnP) in living cells.

2 Materials

2.1 Sensor Plasmid Preparation

1. Mammalian expression plasmid containing the genetically encoded zinc sensors.
2. Competent *E. coli* cells.
3. LB broth.
4. Ampicillin-containing agar plates.
5. Mini-prep DNA purification kit.

2.2 Cell Culture

1. Culture medium for HeLa cells: DMEM and 10% fetal bovine serum.
2. Dulbecco's phosphate-buffered saline (DPBS).

3. Culture medium for INS-1 cells: 43.05 mL of RPMI 1640 media, 5 mL of FBS, 500 μL of 100 mM sodium pyruvate, 500 μL of 1 M HEPES, and 50 μL of 50 mM 2-mercaptoethanol.
4. Culture medium for primary hippocampal neurons: Neural basal and B27.

2.3 Cell Transfection

1. Transfection reagent: 1 mg/mL Polyethylenimine (PEI), pH 7.2.
2. OPTI-MEM.
3. Transfection reagent for neurons: Lipofectamine 3000 transfection reagent.

2.4 Microscopic Imaging

1. Imaging dishes.
2. Phosphate-free HHBSS buffer: 1.26 mM $CaCl_2$, 5.4 mM KCl, 1.1 mM $MgCl_2$, 137 mM NaCl, 16.8 mM D-glucose, 20 mM HEPES, pH 7.4.
3. TPEN (N,N,N′,N′-tetrakis(2-pyridylmethyl)-1,2-ethanediamine) stock solution: 25 mM in DMSO, aliquot and store at −20 °C.
4. Pyrithione (2-Mercaptopyridine *N*-oxide) stock solution: 5 mM in DMSO, aliquot and store at −20 °C.
5. KRB buffer with 3 mM or 35 mM glucose: D-Glucose 0.54 g/L (3 mM) or 6.3 g/L (35 mM), $MgCl_2$ 0.0468 g/L, KCl 0.34 g/L, NaCl 7.0 g/L, sodium phosphate dibasic 0.1 g/L, sodium phosphate monobasic 0.18 g/L.
6. Computer programs: Fiji, Excel, and KaleidaGraph.

3 Methods

3.1 Sensor Plasmid Preparation

1. Order sensor plasmids from Addgene or request from the lab that developed these sensors.
2. Transform the sensor plasmid into competent *E. coli* cells by heat shock.
3. Place transformed *E. coli* cells on agar plates containing ampicillin.
4. Pick single colonies to grow overnight in LB broth.
5. Extract plasmid DNA from *E. coli* cells by miniprep kit according to the manufacturer's protocol.

3.2 HeLa and INS-1 Cell Culture

1. Maintain HeLa cells in DMEM medium with 10% FBS in incubator at 37 °C with 5% CO_2. Maintain INS-1 cells in INS-1 cell media in incubator at 37 °C with 5% CO_2.
2. When the cells reach 80% confluency, pass the cells to maintain healthy cells.

3. Remove the cell media from the cell flask and then wash the cells with DPBS to remove residual medium.
4. Incubate the cells with trypsin (0.05%) at 37 °C for 5 min.
5. Post-trypsin incubation, add 5 mL of pre-warmed cell medium to the cell flask to inhibit trypsin activity. Pipette the cells/cell media up and down ~10 times to thoroughly mix the cells through mechanical force and break apart any cells clumped together.
6. Transfer the dissociated cells to a 15 mL conical tube.
7. Spin down the cells with a cell culture centrifuge at 1000 rpm for 5 min.
8. Remove the supernatant and add fresh cell medium.
9. Dissociate the cell pellet by gentle pipetting.
10. Perform cell count using a hemocytometer.
11. For HeLa cells plate around 5×10^4 cells in each imaging dish with 2 mL of cell media.
12. For INS-1 cells, plate around 5×10^4 cells in each imaging dish precoated with 1 mg/mL of poly-D-lysine (*see* **Note 1**).

3.3 Neuron Preparation

1. A detailed protocol about primary hippocampal neuron culture has been described previously [16].
2. Extract neurons from the dissociated hippocampus of rat fetus and plate cells on 3.5 cm imaging dishes pre-coated with 1 mg/mL poly-D-lysine (*see* **Note 1**).
3. Grow neurons in DMEM + 10% FBS at a density of 360,000 neurons per dish.
4. After 24 h, switch neurons to Neurobasal with 1× B-27 supplement (0.02% v/v).
5. Two days after cell plating, add AraC (5 μM) to hinder the growth of non-neuronal cells.

3.4 HeLa Cell Transfection

1. When the HeLa cells are 40–50% confluent, transfect the sensor into cells.
2. For one transfection reaction, mix 250 μL of OPTI-MEM, 6 μL of PEI transfection reagent, and 1–2 μg of zinc sensor DNA in the same centrifuge tube.
3. Incubate the mixture for 25 min at room temperature.
4. Add the mixture directly to one imaging dish containing HeLa cells, and rock the dish gently side to side.
5. Sensor can be expressed after 12 h.

3.5 INS-1 Cell Transfection

1. Transfection of INS-1 cells is similar to the method used to transfect HeLa cells with slight modification.

2. Mix the following reagents, 250 μL of OPTI-MEM, 6 μL of PEI transfection reagent (*see* **Note 2**), and 1–2 μg of zinc sensor DNA, in one centrifuge tube.
3. Incubate the reaction mixture for 25 min at room temperature.
4. Replace the cell medium with a 9:1 ratio of low-serum INS-1 cell media to normal INS-1 cell media.
5. Add the transfection mixture to the cultured INS-1 cells.
6. After 2–3 h, change the cell media with 2 mL of normal INS-1 cell culture media.
7. Sensor can be expressed 12 h post-transfection.

3.6 Neuron Transfection

1. Mix the following reagents, 250 μL of OPTI-MEM, 2.5 μg of plasmid DNA, and 2.5 μL of p3000 solution, in one centrifuge tube.
2. In a separate centrifuge tube, mix 250 μL of OPTI-MEM and 7.5 μL of Lipofectamine 3000 (*see* **Note 3**).
3. After 5 min, combine the mixture in the above two separate centrifuge tubes and incubate for 15–30 min.
4. From each of the imaging dishes, remove 1 mL of old cell medium, mix with 1 mL of fresh neuron culture medium, then filter with a syringe filter, and store in 37 °C incubator until use.
5. Add the transfection reagent solutions to each imaging dish and incubate the cells in the incubator for 4–6 h.
6. Replace the cell medium containing transfection solution with 2 mL mixed fresh/old medium prepared in **step 4**.
7. Sensor can be expressed 12 h post-transfection.

3.7 Microscopic Imaging of Sensors Targeted to Subcellular Compartments

One of the major advantages to using genetically encoded sensors compared to small-molecule probes is the ability to target the sensor to various subcellular locations. This provides high selectivity for the targeted region [7, 10, 17, 18] as shown in Fig. 2 below. Currently, genetically encoded zinc sensors have been targeted to the cytosol using the nuclear export signal (NES) peptide [10], to the nucleus using the nuclear localization signal (NLS) [10], to the ER using the calreticulin-targeting signal sequence (MLLPVLLLGLLGAAAD) [7], to the mitochondria using the cytochrome c-targeting sequence [8], to the exterior of the plasma membrane using the pDisplay vector [18], to the interior of the plasma membrane using the N-terminal targeting sequence of the protein tyrosine kinase Lck-targeting sequence, and to presynaptic regions using the synaptophysin-targeting sequence and to postsynaptic regions with the PSD95 protein. All the imaging shows high specificity to these locations for the particular targeting

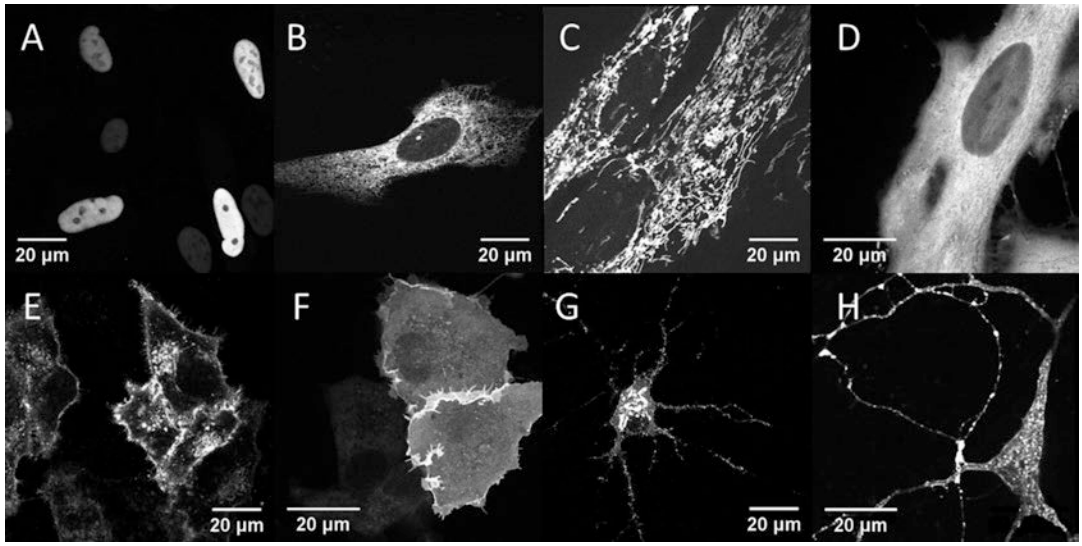


Fig. 2 Subcellular targeting of genetically encoded zinc sensors. (**a**) Nucleus localization of NLS-ZapCY1 in HeLa cell. (**b**) ER localization of ER-ZapCY in HeLa cells. (**c**) Mitochondrial localization of mito-GZnP1 in HeLa cells. (**d**) Cytosol localization in NES-ZapCV2 in HeLa cells. (**e**) Interior localization to plasma membrane using lck-GZnP1 in HeLa cells. (**f**) Exterior localization to plasma membrane using pDisplayGZnP1 in INS-1 cells. (**g**) Postsynaptic localization of PSD95-GZnP1 in primary hippocampal neurons. (**h**) Presynaptic localization of synaphysin-GZnP1 (unpublished) primary hippocampal neurons

sequence (Fig. 2). This genetic specificity essentially allows these zinc sensors to be placed anywhere within the cell for monitoring zinc dynamics in different cellular compartments.

3.8 Quantification of Cytosolic Zinc Concentrations by ZapCV2

The major advantage of FRET-based genetically encoded zinc sensors is that they can provide accurate quantification of labile zinc. Since the sensor concentration is varied in different cells, a calibration must be performed to provide an accurate quantification of cellular zinc concentrations in each cell. The steps of a calibration are outlined below.

1. Two days after transfection, cells transfected with NES-ZapCV2 can be used for imaging experiment.
2. Wash cells in imaging dishes with HBSS buffer three times prior to imaging.
3. Identify the imaging field with healthy cells and acquire time-lapse images every 20 s with a 40X oil objective (NA 1.3).
4. Incubate cells in 2 mL HHBSS buffer for 10 min for the baseline condition.
5. Then add 10 μL of 25 mM TPEN to the imaging dishes to chelate cellular zinc.
6. When sensor signals reach to the minimum, remove the TPEN and wash the cells with HHBSS buffer three times.

7. At last, add a mixture of $ZnCl_2$ and pyrithione to the imaging dish to reach a final concentration of 100 μM $ZnCl_2$ and 5 μM pyrithione.
8. Analyze the data using the computer program Fiji and Excel.
9. Align every image in an image stack to account for drift using the Stackreg plug-in in Fiji.
10. For the quantitative analysis images of NES-ZapCV2 need to be background corrected. Generate a region of interest (ROI) on a blank area of the coverslip and subtract the fluorescence intensity of each channel, e.g., I_{FRET}(sample) − I_{FRET}(background), and I_{CFP}(sample) − I_{CFP}(background). The background-corrected FRET and CFP images are used to calculate the FRET ratio (R, i.e., FRET/CFP).
11. Plot the calibration of NES-ZapCV2 in the KaleidaGraph program (Fig. 3).
12. The baseline FRET ratio (R), minimal FRET ratio achieved with TPEN (R_{min}), and maximal FRET ratio (R_{max}) obtained with zinc saturation are used for cytosolic zinc quantification using the equation $K_d\,[(R - R_{min})/(R_{max} - R)]^{1/n}$ ($K_d = 2.3$ nM, $n = 0.53$) [17].

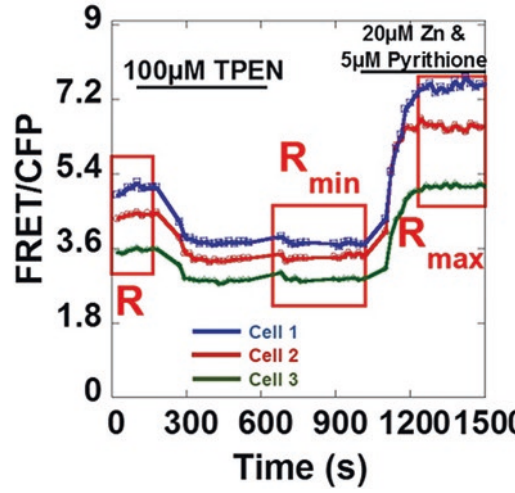


Fig. 3 Calibration and quantification of cytosolic zinc concentrations using the NES-ZapCV2 sensor. Three different cells are shown in the same imaging experiment with different resting FRET ratio (R). Treatment with 100 μM TPEN yielded the minimal FRET ratio (R_{min}), while saturation with 100 μM $ZnCl_2$ and 5 μM pyrithione gives the maximal FRET raito (R_{max}). These were used with the $K_d = 2.3$ nM and $n = 0.53$ to calculate the concentration of labile zinc using the equation $[Zn^{2+}] = K_d\,[(R - R_{min})/(R_{max} - R)]^{1/n}$

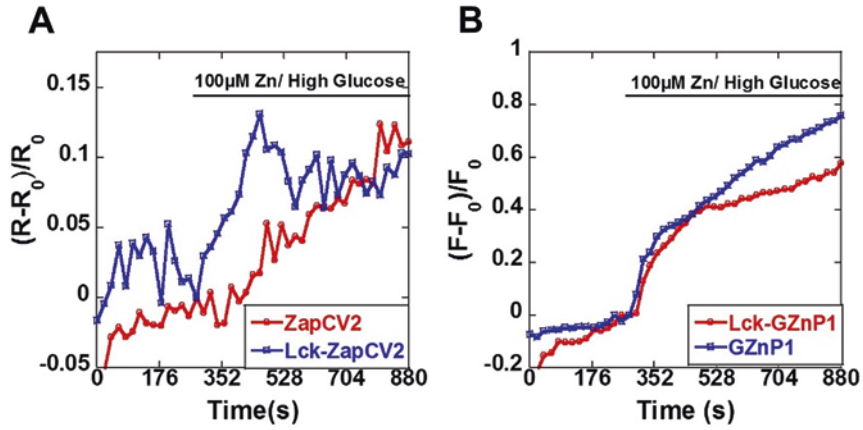


Fig. 4 Monitoring zinc influx in INS-1 cells induced by high glucose. Cells were incubated in 3 mM glucose media to starve the cells of glucose for 5 min. Then 100 μM $ZnCl_2$ was added with 35 mM glucose to induce influx of zinc into the INS-1 cells. Zinc influx was recorded using ZapCV2 (**a**) and GZnP1 (**b**) in cytosols (red line) and cytosolic domains close to plasma membrane (Lck-targeting sequence, blue line)

3.9 Monitor Zinc Flux Using ZapCV2 and GZnP1 Sensor in INS-1 Cells

Here we illustrate an example using ZapCV2 and GZnP1 sensors to study zinc influx in INS-1 cells. INS-1 cells are insulin-secreting pancreatic beta cells and are great model cells when studying any secretory pathway. When glucose is transported into INS-1 cells, the glucose is converted to ATP, and then ATP blocks potassium channels resulting in the depolarization of INS-1 cells. Upon depolarization of the INS-1 cells membrane, voltage-dependent calcium channels open allowing zinc flux through the channels into the cell [19]. When INS-1 cells preincubated with low-glucose medium (3 mM) for 30 min were induced with high glucose (35 mM) and 100 μM $ZnCl_2$, zinc influx was recorded for both ZapCV2 sensor and GZnP1 sensor (Fig. 4). No significant differences were detected between the cytosolic microdomains close to the membrane and the whole cytosolic regions (Fig. 4). However, the GZnP1 sensor displayed a higher sensor response than the ZapCV2 sensor, suggesting that GZnP1 provides advantages in monitoring zinc flux.

4 Notes

1. Coat imaging dishes with poly-D-lysine. Dissolve poly-D-lysine in 0.15 M sodium borate buffer to 10 mg/mL for long-term stock. Before coating, dilute the stock solution to 1 mg/mL in MilliQ water. Coat imaging dishes with 1 mg/mL poly-D-lysine overnight. Then wash three times with water.

2. Our lab has successfully used both Lipofectamine 3000 and homemade PEI transfection reagent to transfect INS-1 cells. Lipofectamine 3000 is more expensive, but it does have a higher transfection rate.
3. This Lipofectamine reaction setup is different than the manufacturer's recommended protocol. The quantities of DNA, Lipofectamine 3000, and P3000 reagent can be changed to optimize transfection efficiency if desired.

References

1. Andreini C, Banci L, Bertini I, Rosato A (2006) Counting the zinc-proteins encoded in the human genome. J Proteome Res 5:196–201
2. Wilson M, Hogstrand C, Maret W (2012) Picomolar concentrations of free zinc(II) ions regulate receptor protein-tyrosine phosphatase beta activity. J Biol Chem 287:9322–9326
3. Dineley KE, Richards LL, Votyakova TV, Reynolds IJ (2005) Zinc causes loss of membrane potential and elevates reactive oxygen species in rat brain mitochondria. Mitochondrion 5:55–65
4. Jiang D, Sullivan PG, Sensi SL, Steward O, Weiss JH (2001) Zn(2+) induces permeability transition pore opening and release of pro-apoptotic peptides from neuronal mitochondria. J Biol Chem 276:47524–47529
5. Faxen K, Salomonsson L, Adelroth P, Brzezinski P (2006) Inhibition of proton pumping by zinc ions during specific reaction steps in cytochrome c oxidase. Biochim Biophys Acta 1757:388–394
6. Prasad AS (1995) Zinc: an overview. Nutrition 11(1 Suppl):93–99
7. Qin Y, Dittmer PJ, Park JG, Jansen KB, Palmer AE (2011) Measuring steady-state and dynamic endoplasmic reticulum and Golgi Zn2+ with genetically encoded sensors. Proc Natl Acad Sci U S A 108:7351–7356
8. Qin Y, Sammond DW, Braselmann E, Carpenter MC, Palmer AE (2016) Development of an optical Zn2+ probe based on a single fluorescent protein. ACS Chem Biol 11:2744–2751
9. Park JG, Qin Y, Galati DF, Palmer AE (2012) New sensors for quantitative measurement of mitochondrial Zn(2+). ACS Chem Biol 7:1636–1640
10. Miranda JG, Weaver AL, Qin Y, Park JG, Stoddard CI, Lin MZ, Palmer AE (2012) New alternately colored FRET sensors for simultaneous monitoring of Zn(2)(+) in multiple cellular locations. PLoS One 7:e49371
11. Vinkenborg JL, Nicolson TJ, Bellomo EA, Koay MS, Rutter GA, Merkx M (2009) Genetically encoded FRET sensors to monitor intracellular Zn2+ homeostasis. Nat Methods 6:737–740
12. Chabosseau P, Tuncay E, Meur G, Bellomo EA, Hessels A, Hughes S, Johnson PR, Bugliani M, Marchetti P, Turan B, Lyon AR, Merkx M, Rutter GA (2014) Mitochondrial and ER-targeted eCALWY probes reveal high levels of free Zn2+. ACS Chem Biol 9:2111–2120
13. Evers TH, Appelhof MA, de Graaf-Heuvelmans PT, Meijer EW, Merkx M (2007) Ratiometric detection of Zn(II) using chelating fluorescent protein chimeras. J Mol Biol 374:411–425
14. Hessels AM, Chabosseau P, Bakker MH, Engelen W, Rutter GA, Taylor KM, Merkx M (2015) eZinCh-2: a versatile, genetically encoded FRET sensor for cytosolic and Intraorganelle Zn(2+) imaging. ACS Chem Biol 10:2126–2134
15. Qiao W, Mooney M, Bird AJ, Winge DR, Eide DJ (2006) Zinc binding to a regulatory zinc-sensing domain monitored in vivo by using FRET. Proc Natl Acad Sci U S A 103:8674–8679
16. Kaech S, Banker G (2006) Culturing hippocampal neurons. Nat Protoc 1:2406–2415
17. Fiedler BL, Van Buskirk S, Carter KP, Qin Y, Carpenter MC, Palmer AE, Jimenez R (2017) Droplet microfluidic flow cytometer for sorting on transient cellular responses of genetically-encoded sensors. Anal Chem 89:711–719
18. Siau A, Huang X, Yam XY, Bob NS, Sun H, Rajapakse JC, Renia L, Preiser PR (2014) Identification of a new export signal in Plasmodium yoelii: identification of a new exportome. Cell Microbiol 16:673–686
19. Gyulkhandanyan AV, Lee SC, Bikopoulos G, Dai F, Wheeler MB (2006) The Zn2+-transporting pathways in pancreatic beta-cells: a role for the L-type voltage-gated Ca2+ channel. J Biol Chem 281:9361–9372

Chapter 9

Delivery of Superoxide Dismutase Using Cys_2-His_2 Zinc-Finger Proteins

Jia Liu, Jiangmei Li, Jie Li, Lianhui Zhu, Shaojie Wang, Xuan Wei, and Peixiang Ma

Abstract

Therapeutic proteins have shown great potential in treating life-threatening diseases, but the hydrophilicity and high molecular weight hamper their passing through the cell membrane. Cell-penetrating peptide (CPP)-assisted protein delivery is a simple and efficacious strategy to promote the cellular uptake of therapeutic proteins. We recently demonstrated that the engineered Cys_2-His_2 zinc-finger domains possess intrinsic cell permeability, which could be leveraged for intracellular protein delivery. Here we applied this method to deliver superoxide dismutase (SOD), a therapeutic protein widely used in preclinical and clinical studies. We present a protocol for the production and delivery of zinc-finger domain-fused SOD. This protocol can be extended for delivering other therapeutic proteins.

Key words Cell-penetrating peptides, Protein delivery, Superoxide dismutase, Zinc-finger proteins

1 Introduction

The selective permeability of cellular membranes separates the cellular components from exogenous molecules. This proves a major challenge for the delivery of therapeutic agents into cells [1]. Various bioactive agents including genes, proteins, and viruses can be efficiently internalized into cells. Among these agents, the direct delivery of functional proteins holds great promise for therapeutic applications because of its safety and efficiency [2, 3]. Protein delivery does not depend on the transcription and translation of imported nucleic acids. Therefore, the delivered proteins can act rapidly and then be degraded by proteasome system, leading to less risk of mutagenesis [4–6]. One major obstacle of protein delivery is the selectivity of cell membrane. Numerous membrane perturbation techniques, such as microinjection and electroporation, have been investigated for accelerating protein delivery. However, these membrane disruption technologies are often associated with low efficiency, high toxicity, penurious bioavailability, and poor

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_9,

specificity [7]. In addition to the physical membrane puncture methods, many biochemical agents were developed to facilitate protein delivery, such as supercharged transduction domains [8, 9], nanoparticles [10], liposomes [11], viruslike particles [12, 13] and polymeric microsphere [14]. In the preclinical or clinical practice, these strategies can be associated with drawbacks such as inefficient cellular uptake [15, 16], poor stability [17], inadvertent cell-type specificity [18], low rate of endosomal escape [19] or toxicity [20]. In the late 1980s, a naturally occurring peptide from TAT trans-activating factor of human immunodeficiency virus (HIV) was found to possess inherent cell-penetrating ability. In the following years, a series of natural peptides with similar cell permeability were identified, which were later recognized as cell-penetrating peptides (CPPs) [21–24]. Based on the features of naturally occurring CPPs, artificial or chimeric CPPs were designed [25–28]. CPPs often have minimal cytotoxicity and can be applied to various cell types for the delivery of a wide range of cargo molecules with different molecular weights [29]. These CPPs can be either genetically fused to or chemically conjugated to the cargo proteins.

Recently, we identified the Cys_2-His_2 zinc-finger proteins (ZFPs) as a novel protein delivery system [30–32]. ZFPs are inherently cell permeable due to the constellation of six positively charged residues on the protein surface [33]. We eliminated the DNA-binding ability of ZFPs by mutating the residues responsible for DNA binding in the α-helices. The engineered zinc-finger proteins (ZFPs) retained cell permeability and can be used as a fusion tag to deliver cargo proteins [31]. The cellular uptake efficacy is tunable by adjusting the number of tandem ZFP domains, which increases the plasticity for different applications [31]. ZFP domains can mediate the efficient intracellular delivery of protein cargos such as green fluorescent protein (GFP) [30] and *Fok* I nuclease [34, 35]. In addition to transformed cell lines, ZFPs can facilitate the delivery of cargo proteins into primary cells and stem cells [35], which is important for the therapeutic applications.

Superoxide dismutase (SOD) family includes a group of well-studied antioxidant enzymes, i.e., SOD1, SOD2, and SOD3 [36]. SODs play a fundamental role in attenuating oxidative stress from cellular reactive oxygen species (ROS) [36, 37]. The disorder of ROS can contribute to the occurrence and progression of a variety of diseases. Preclinical and clinical studies have shown great therapeutic potential of SODs [38, 39]. SODs have been employed for a wide range of medical indications such as ischemia reperfusion injury [40, 41], transplant-induced reperfusion injury [42], inflammation [43, 44], Parkinson's disease [45], cancer [46–49] and acquired immune deficiency syndrome (AIDS) [50, 51]. In this protocol, we provide a detailed protocol for the implementation

of ZFPs on human SOD1 (hSOD1). Key steps to maximize the efficiency of intracellular delivery of ZFP-SOD1 are also highlighted and discussed.

2 Materials

2.1 Construction of Expression Plasmids

1. pET28a plasmid encoding human SOD1 gene optimized for *Escherichia coli* expression (available from commercial gene synthesis service suppliers).
2. ZFP containing plasmid as described in [31].
3. DNA polymerase.
4. Deoxynucleotide mixture, including dATP, dCTP, dGTP, and dTTP.
5. PCR reaction buffers.
6. Sterile water.
7. DNA-staining reagents.
8. Homologous recombination enzymes.
9. DH5α *E. coli* competent cells.
10. Lysogeny broth (LB) medium.
11. Agar, bacteriological grade.
12. Plasmid DNA extraction kit.
13. Gradient thermal cycler for PCR.
14. Agarose gel electrophoresis reagents and equipment.
15. UV transilluminator.
16. Centrifuge.
17. Water bath.

2.2 Protein Expression and Purification

1. Plasmids encoding recombinant ZFP-SOD1 proteins.
2. BL21(DE3) competent *E.coli* cells.
3. Agar, bacteriological grade.
4. 50 mg/mL Kanamycin stock solution
5. 1 M Isopropyl-β-D-1-thiogalactopyranoside (IPTG).
6. Ni-NTA agarose.
7. 1 M Tris–HCl pH 8.0.
8. 4 M Imidazole stock solution.
9. 100 mM $ZnCl_2$ stock solution.
10. 100 mM $MgCl_2$ stock solution.
11. Phenylmethylsulfonyl fluoride (PMSF) (100 mM in ethanol).

12. Lysis buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 100 μM $ZnCl_2$, 1 mM $MgCl_2$, 1 mM PMSF, and 5 mM imidazole.
13. Wash buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 100 μM $ZnCl_2$, 1 mM $MgCl_2$, and 30 mM imidazole.
14. Elution buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 100 μM $ZnCl_2$, 1 mM $MgCl_2$, and 300 mM imidazole.
15. Storage buffer: 50 mM Tris–HCl, pH 8.0, 500 mM NaCl, 100 μM $ZnCl_2$, 1 mM $MgCl_2$, and 10% glycerol.
16. Baffled cell culture flasks.
17. Concentrator.
18. 4–20% Tris-glycine SDS-PAGE.
19. SDS protein-loading dye.
20. BCA protein assay kit.
21. Liquid nitrogen.

2.3 Protein Transduction

1. Purified ZFP-SOD1 proteins.
2. Class II biosafety cabinet.
3. Cell incubator.
4. Bright-field phase-contrast microscope.
5. Dulbecco's modified Eagle's medium (DMEM).
6. Fetal bovine serum (FBS).
7. Penicillin and streptomycin solution.
8. Phosphate-buffered saline (PBS).
9. 9 mM $ZnCl_2$.
10. Triton X-100.
11. 0.05% Trypsin-EDTA solution with phenol red.
12. HeLa cells.
13. Tissue culture flasks.
14. 24-Well flat-bottom tissue culture plates.
15. Centrifuge.
16. SOD assay kit.

3 Methods

3.1 Construction of ZFP-SOD1 Expression Plasmids

1. Miniprep ZFP-encoding plasmid (pET28-ZiF1-EmGFP) [31] and synthesize human SOD genes in pET28a vector (pET28a-hSOD1) using a commercial gene synthesis service (*see* **Note 1**).

2. PCR amplify ZFP genes from the plasmid pET-1F-ZiF with the primers ZFP-Fwd (5′- GCCTGGTGCCGCGCGGCAG| CCCGAAAAAGAAACGCAAAGTGC-3′) and ZFP-SOD1-Rev (5'-GCTTTGGTGGCCATGGATCCACCGGTATGT GTTCTTTGATGG -3′).
3. Prepare PCR mixture to amplify the genes encoding ZFP domain: Use 5 ng of template DNA, 5 μL of 10× polymerase buffer, 1 unit (U) of *Taq* DNA polymerase, 0.2 mM dNTP mixture, and 0.2 μM of each primer in a 50 μL solution. PCR conditions are cycled using the following settings: 95 °C for 5 min; 30 cycles of 95 °C for 30 s, 58 °C for 30 s, and 72 °C for 1 min; and final extension at 72 °C for 10 min. Purify the PCR product by gel extraction and determine DNA concentration using a spectrophotometer measuring Abs260 × 50 ng/μL.
4. Digest 1 μg of pET28a-hSOD1 plasmid with 10 U of each NdeI and BamHI in recommended buffer for 3 h at 37 °C. Visualize DNA by agarose gel electrophoresis using a DNA-staining dye, such as gel red (*see* **Note 2**).
5. Purify the digested plasmid by gel extraction kit and determine DNA concentration by a spectrophotometer measuring Abs260 × 50 ng/μL.
6. Perform homologous recombination reaction (*see* **Note 3**) for constructing ZFP-SOD1 fusion protein (Fig. 1a) as follows: 0.06 pmol ZFP PCR product, 0.03 pmol linearized pET28a-hSOD1 plasmid DNA, 2 μL recombination enzyme such as Exnase II, 4 μL 5× recombination buffer, and deionized water up to 20 μL. Incubate the mixture at 37 °C for 30 min (*see* **Note 4**).

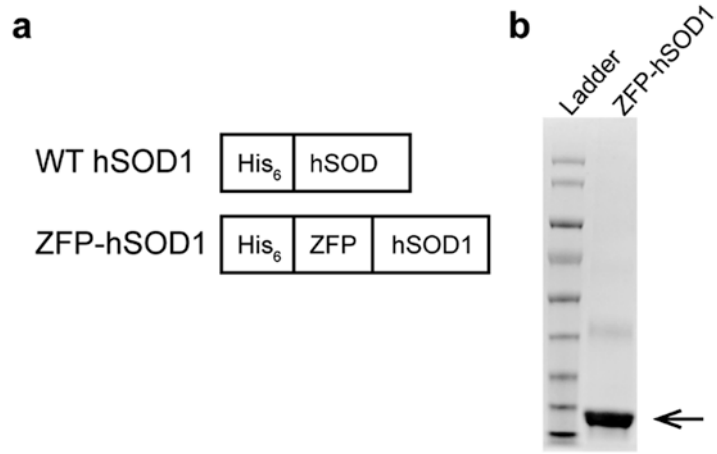


Fig. 1 ZFP-SOD1 protein purification. (**a**) Schematic presentation of ZFP-hSOD1 construct. (**b**) SDS-PAGE of purified ZFP-hSOD1 protein. Arrow indicates the target protein band

7. Thaw 200 μL of chemically competent DH5α *E. coli* cells on ice, mix gently with 20 μL of recombination products, and then incubate on ice for 30 min (*see* **Note 5**).
8. Heat shock the mixture at 42 °C for 45–90 s and recover the cells in 900 μL LB medium for 1 h at 37 °C with shaking.
9. Spread 100 μL of recovery culture on a LB agar plate supplemented with 50 μg/mL kanamycin and incubate overnight at 37 °C.
10. The following day, inoculate a single colony into 5 mL of LB culture containing 50 μg/mL kanamycin and culture overnight at 37 °C.
11. Miniprep pET28a-ZFP-SOD1 plasmid and confirm the construct by DNA sequencing using a primer binding to the T7 promoter (5'-TAATACGACTCACTATAGGG-3′) (*see* **Note 6**).

3.2 Expression and Purification of ZFP-SOD1

1. Thaw 50 μL of chemically competent BL21 *E. coli* cells on ice and mix gently with 100 ng of pET28a-ZFP-SOD1 plasmid. Perform heat-shock transformation as described in Subheading 3.1, **step 8**.
2. The following day, inoculate a single colony into 10 mL of LB medium containing 50 μg/mL kanamycin and grow overnight at 37 °C with shaking.
3. The following day, dilute the 10 mL of overnight culture into 1 L of LB medium supplemented with 50 μg/mL kanamycin. Grow the culture at 37 °C with shaking to an optical density at 600 nm (OD_{600}) of 0.6–0.8 and induce protein expression with 0.5 mM IPTG. After 6 h of expression at 37 °C, harvest cells by centrifugation at 5000 × *g* for 20 min at 4 °C (*see* **Note 7**).
4. Resuspend 1 g cell pellets in 5 mL of lysis buffer. Lyse the cells on ice with cell distributor or sonication (*see* **Note 8**).
5. Centrifuge the cell lysate at 40,000 × *g* for 30 min at 4 °C and transfer the supernatant into a fresh collection tube. For optimum results, perform all the following steps at 4 °C.
6. Flow the supernatant through a column prepacked with 1 mL of equilibrated Ni-NTA agarose. Wash the resin with 20 mL of wash buffer.
7. Elute the protein with 5 mL of elution buffer.
8. Buffer exchange the eluted protein with storage buffer and concentrate the protein to at least 40 μM using a spin concentrator following the manufacturer's instructions.
9. Determine protein concentration by BCA or Bradford assay.
10. Mix 2 μL of purified proteins with 2 μL 2× SDS-PAGE-loading dye, boil at 95 °C for 10 min, and then resolve on 4–20% Tris-glycine SDS-PAGE to assess protein purity (Fig. 1b).

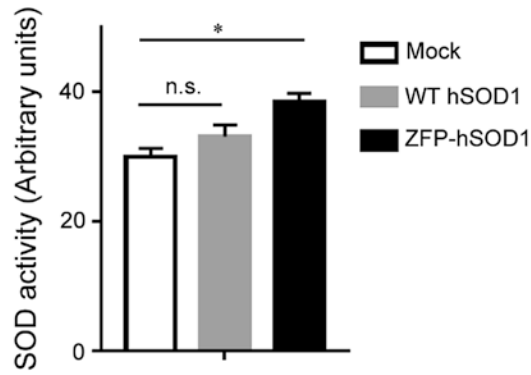


Fig. 2 Internalized SOD1 protein in HeLa cells, as determined by SOD activity. *n.s.* not significant. $*P < 0.05$

11. Aliquot the concentrated protein, flash freeze in liquid nitrogen, and store at −80 °C (*see* **Note 9**).

3.3 Protein Transduction

1. Maintain HeLa cells in DMEM medium supplemented with 10% (v/v) FBS, 100 U/mL penicillin, and 100 U/mL streptomycin at 37 °C in a fully humidified atmosphere with 5% CO_2 (*see* **Note 10**).
2. Pre-coat a 24-well plate with 500 μL of 50 μg/mL of polylysine for 30 to 60 min at 25 °C. Seed HeLa cells onto pre-coated plates at a density of 2×10^5 cells per well.
3. At 24 h after seeding, remove medium from each well and wash with 500 μL of pre-warmed serum-free DMEM.
4. Add to each well 250 μL of SFM containing 2 μM of ZFP-SOD1 proteins and 100 μM $ZnCl_2$. Incubate at 37 °C for 1 h (*see* **Note 11**).
5. Remove media from cells and wash three times with 500 μL of calcium- and magnesium-free PBS supplemented with 0.5 mg/mL of heparin (*see* **Note 12**).
6. Rinse cells with 0.05% trypsin-EDTA, remove trypsin solution, and then incubate at 37 °C for 2 min.
7. Lyse cells using 250 μL of PBS containing 0.1% (v/v) triton X-100.
8. Use SOD assay kit such as SOD Determination Kit (Sigma-Aldrich, St. Louis, MO, USA) to determine internalized SOD proteins (Fig. 2) (*see* **Note 13**).

4 Notes

1. Codon usage is biased in different species. In order to increase the translational efficiency, the genes of interest should be

optimized according to the host species. We optimized the human SOD1 and ZFP genes according to the codon usage of *E. coli* K12.

2. Digestion reaction with single restriction enzyme is also acceptable. However, in order to suppress self-ligation, double digestion and use of phosphatase such as shrimp alkaline phosphatase are recommended.
3. Techniques other than homologous recombination are also acceptable for constructing the fusion protein. However, it is important to avoid the introduction of additional amino acids into the fusion proteins.
4. For the homologous recombination reaction, appropriate incubation time is necessary for high recombination efficiency. Additionally, the reaction should be carried out in a PCR thermo-cycler or other machines with precise temperate control. To achieve optimum results, the recombination reaction should be performed using an insert-to-vector molar ratio of 2:1. Upon completion, the reaction solution can be stored at −20 °C for future experiment.
5. High-competency cells (over 10^8 cfu/μg DNA) are recommended for optimum results. The total volume of the recombination products should not be more than 1/10 of the competent cell volume.
6. A colony PCR screen using the primers ZFP-Fwd and ZFP-SOD1-rev may be performed before DNA sequencing.
7. The experiment can be paused here. The harvested cell pellets could be stored at −20 °C. Induction conditions are highly variable and depend on the stability of the proteins being expressed. We recommend monitoring the OD_{600} every 30 min until an OD_{600} of 0.8 is reached. Different induction conditions should be examined when expressing new recombinant proteins. The variables include but are not limited to the concentration of IPTG, induction temperature, and expression time.
8. Heating during cell lysis may cause the denaturation and degradation of the target protein. Repetitive ice cooling and lower power of sonication are necessary to avoid overheating the cell lysate.
9. Avoid repeated freezing and thawing of the protein solution. ZFP-fused proteins are stable under these conditions for at least 1 month. Inappropriate or over 4-month storage may lead to protein precipitation or degradation.
10. Cells with more than 30 passages are not recommended due to low efficiency of protein transduction.

11. ZFP proteins enter cells primarily through micropinocytosis [31], which is an energy-dependent process. Therefore, cells must be incubated at 37 °C for efficient protein internalization.
12. Heparin is necessary to remove surface-bound protein that may complicate downstream analyses.
13. It is important to have a standard curve using standard SOD protein samples.

Acknowledgments

This work was supported by ShanghaiTech University, the National Natural Science Foundation of China (31500632 to P. Ma), and National Natural Science Foundation of China (31600686 to J. Liu).

References

1. Rehman K, Hamid Akash MS, Akhtar B, Tariq M, Mahmood A, Ibrahim M (2016) Delivery of therapeutic proteins: challenges and strategies. Curr Drug Targets 17:1172–1188
2. van den Berg A, Dowdy SF (2011) Protein transduction domain delivery of therapeutic macromolecules. Curr Opin Biotechnol 22:888–893
3. Lindsay MA (2002) Peptide-mediated cell delivery: application in protein target validation. Curr Opin Pharmacol 2:587–594
4. Luo D, Saltzman WM (2000) Synthetic DNA delivery systems. Nat Biotechnol 18:33–37
5. Guo X, Huang L (2012) Recent advances in nonviral vectors for gene delivery. Acc Chem Res 45:971–979
6. Thomas CE, Ehrhardt A, Kay MA (2003) Progress and problems with the use of viral vectors for gene therapy. Nat Rev Genet 4:346–358
7. Copolovici DM, Langel K, Eriste E, Langel U (2014) Cell-penetrating peptides: design, synthesis, and applications. ACS Nano 8:1972–1994
8. Fuchs SM, Raines RT (2007) Arginine grafting to endow cell permeability. ACS Chem Biol 2:167–170
9. Cronican JJ, Thompson DB, Beier KT, McNaughton BR, Cepko CL, Liu DR (2010) Potent delivery of functional proteins into mammalian cells in vitro and in vivo using a supercharged protein. ACS Chem Biol 5:747–752
10. Panyam J, Labhasetwar V (2003) Biodegradable nanoparticles for drug and gene delivery to cells and tissue. Adv Drug Deliv Rev 55:329–347
11. Zelphati O, Wang Y, Kitada S, Reed JC, Felgner PL, Corbeil J (2001) Intracellular delivery of proteins with a new lipid-mediated delivery system. J Biol Chem 276:35103–35110
12. Kaczmarczyk SJ, Sitaraman K, Young HA, Hughes SH, Chatterjee DK (2011) Protein delivery using engineered virus-like particles. Proc Natl Acad Sci U S A 108:16998–17003
13. Voelkel C, Galla M, Maetzig T, Warlich E, Kuehle J, Zychlinski D, Bode J, Cantz T, Schambach A, Baum C (2010) Protein transduction from retroviral Gag precursors. Proc Natl Acad Sci U S A 107:7805–7810
14. Sinha VR, Trehan A (2003) Biodegradable microspheres for protein delivery. J Control Release 90:261–280
15. Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas CF 3rd (2014) Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One 9:e85755
16. Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res 24:1020–1027
17. Fuchs SM, Raines RT (2005) Polyarginine as a multifunctional fusion tag. Protein Sci 14:1538–1544

18. Mai JC, Shen H, Watkins SC, Cheng T, Robbins PD (2002) Efficiency of protein transduction is cell type-dependent and is enhanced by dextran sulfate. J Biol Chem 277:30208–30218
19. Al-Taei S, Penning NA, Simpson JC, Futaki S, Takeuchi T, Nakase I, Jones AT (2006) Intracellular traffic and fate of protein transduction domains HIV-1 TAT peptide and octaarginine. Implications for their utilization as drug delivery vectors. Bioconjug Chem 17:90–100
20. Jones SW, Christison R, Bundell K, Voyce CJ, Brockbank SM, Newham P, Lindsay MA (2005) Characterisation of cell-penetrating peptide-mediated peptide delivery. Br J Pharmacol 145:1093–1102
21. Frankel AD, Pabo CO (1988) Cellular uptake of the tat protein from human immunodeficiency virus. Cell 55:1189–1193
22. Elliott G, O'Hare P (1997) Intercellular trafficking and protein delivery by a herpesvirus structural protein. Cell 88:223–233
23. Derossi D, Joliot AH, Chassaing G, Prochiantz A (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes. J Biol Chem 269:10444–10450
24. Hamley IW (2017) Small bioactive peptides for biomaterials design and therapeutics. Chem Rev 117:14015–14041
25. Smith BA, Daniels DS, Coplin AE, Jordan GE, McGregor LM, Schepartz A (2008) Minimally cationic cell-permeable miniature proteins via alpha-helical arginine display. J Am Chem Soc 130:2948–2949
26. Daniels DS, Schepartz A (2007) Intrinsically cell-permeable miniature proteins based on a minimal cationic PPII motif. J Am Chem Soc 129:14578–14579
27. Karagiannis ED, Urbanska AM, Sahay G, Pelet JM, Jhunjhunwala S, Langer R, Anderson DG (2013) Rational design of a biomimetic cell penetrating peptide library. ACS Nano 7:8616–8626
28. Gao S, Simon MJ, Hue CD, Morrison B 3rd, Banta S (2011) An unusual cell penetrating peptide identified using a plasmid display-based functional selection platform. ACS Chem Biol 6:484–491
29. Heitz F, Morris MC, Divita G (2009) Twenty years of cell-penetrating peptides: from molecular mechanisms to therapeutics. Br J Pharmacol 157:195–206
30. Gaj T, Liu J (2015) Direct protein delivery to mammalian cells using cell-permeable Cys2-His2 zinc-finger domains. J Vis Exp 97:52814
31. Gaj T, Liu J, Anderson KE, Sirk SJ, Barbas CF 3rd (2014) Protein delivery using Cys2-His2 zinc-finger domains. ACS Chem Biol 9:1662–1667
32. Liu J, Gaj T, Wallen MC, Barbas CF 3rd (2015) Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering. Mol Ther Nucleic Acids 4:e232
33. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 9:805–807
34. Liu J, Shui SL (2016) Delivery methods for site-specific nucleases: achieving the full potential of therapeutic gene editing. J Control Release 244:83–97
35. Liu J, Gaj T, Yang Y, Wang N, Shui S, Kim S, Kanchiswamy CN, Kim JS, Barbas CF 3rd (2015) Efficient delivery of nuclease proteins for genome editing in human stem cells and primary cells. Nat Protoc 10:1842–1859
36. Salvemini D, Riley DP, Cuzzocrea S (2002) SOD mimetics are coming of age. Nat Rev Drug Discov 1:367–374
37. Fridovich I (1995) Superoxide radical and superoxide dismutases. Annu Rev Biochem 64:97–112
38. Maxwell SRJ (1995) Prospects for the use of antioxidant therapies. Drugs 49:345–361
39. McCord JM (1974) Free radicals and inflammation: protection of synovial fluid by superoxide dismutase. Science 185:529–531
40. McCord JM (1986) Superoxide dismutase: rationale for use in reperfusion injury and inflammation. J Free Radic Biol Med 2:307–310
41. Yang G, Chan PH, Chen J, Carlson E, Chen SF, Weinstein P, Epstein CJ, Kamii H (1994) Human copper-zinc superoxide dismutase transgenic mice are highly resistant to reperfusion injury after focal cerebral ischemia. Stroke 25:165–170
42. Land W, Zweler JL (1997) Prevention of reperfusion-induced, free radical-mediated acute endothelial injury by superoxide dismutase as an effective tool to delay/prevent chronic renal allograft failure: a review. Transplant Proc 29:2567–2568
43. Shingu M, Takahashi S, Ito M, Hamamatu N, Suenaga Y, Ichibangase Y, Nobunaga M (1994) Anti-inflammatory effects of recombinant human manganese superoxide dismutase on adjuvant arthritis in rats. Rheumatol Int 14:77–81
44. Droy-Lefaix MT, Drouet Y, Geraud G, Hosford D, Braquet P (1991) Superoxide dis-

mutase (SOD) and the PAF-antagonist (BN 52021) reduce small intestinal damage induced by ischemia-reperfusion. Free Radic Res Commun 12-13(Pt 2):725–735

45. Federica De Lazzari, Alexander J Whitworth, Marco Bisaglia (2017) Superoxide radical dismutation as new therapeutic strategy in Parkinson's disease. Aging Dis http://www.aginganddisease.org/EN/10.14336/AD.2017.1018#1
46. Church SL, Grant JW, Ridnour LA, Oberley LW, Swanson PE, Meltzer PS, Trent JM (1993) Increased manganese superoxide dismutase expression suppresses the malignant phenotype of human melanoma cells. Proc Natl Acad Sci U S A 90:3113–3117
47. Bravard A, Sabatier L, Hoffschir F, Ricoul M, Luccioni C, Dutrillaux B (1992) SOD2: a new type of tumor-suppressor gene? Int J Cancer 51:476–480
48. Safford SE, Oberley TD, Urano M, St Clair DK (1994) Suppression of fibrosarcoma metastasis by elevated expression of manganese superoxide dismutase. Cancer Res 54:4261–4265
49. Yoshizaki N, Mogi Y, Muramatsu H, Koike K, Kogawa K, Niitsu Y (1994) Suppressive effect of recombinant human Cu, Zn-superoxide dismutase on lung metastasis of murine tumor cells. Int J Cancer 57:287–292
50. Mollace V, Nottet HS, Clayette P, Turco MC, Muscoli C, Salvemini D, Perno CF (2001) Oxidative stress and neuroAIDS: triggers, modulators and novel antioxidants. Trends Neurosci 24:411–416
51. Flores SC, Marecki JC, Harper KP, Bose SK, Nelson SK, McCord JM (1993) Tat protein of human immunodeficiency virus type 1 represses expression of manganese superoxide dismutase in HeLa cells. Proc Natl Acad Sci U S A 90:7632–7636

Chapter 10

Genome Editing of MSCs as a Platform for Cell Therapy

Krissanapong Manotham and Supreecha Chattong

Abstract

The rapid growing of genome editing technology leads to an optimistic expectation for the treatment of many incurable diseases. The core of genome editing relies on DNA-repairing processes which occur independently in each cell, thus generating a mix of genetic variation cells. Here, we describe the protocol of using mesenchymal stem cells as a platform to generate high purity of ZFN-edited patients' cell clones which may be useful in conjunction with therapeutic cell conversion and reprograming.

Key words Genome editing, Zinc-finger nuclease (ZFN), Mesenchymal stem cells, Homologous recombination

1 Introduction

The rapid growing of genome editing process leads to an optimistic expectation for the treatment of many incurable diseases [1–5]. Theoretically, gene-editing tools are able to disrupt any normal cellular DNA sequence at the very specific site and generate mutations by an error-prone repairing, and nonhomologous end jointing (NHEJ) that eventually leads to loss of functions of targeted genes [6, 7]. In addition, exogenous nucleotide can be replaced/inserted into the targeted gene via homologous recombination (HR) [8–10]. With these tools, we can correct the pathogenic mutation with the desirable nucleotides or even entirely replace the defective gene with the perfect artificial one [4, 5]. Such therapeutic genome editing can be achieved only if in vivo gene correction can be performed at the very specific cells, easier if we can replace defective cells with the corrected ones. Unfortunately, the core of genome editing relies on DNA-repairing processes which occur independently in each cell, thus creating a large diversified results at the single-cell level.

Mesenchymal stem cells (MSCs) can narrow the therapeutic gap of genome editing technology. MSCs are families of somatic stem cells from various resources, for example, bone marrow, adipose tissue, and dental pulp [11–13]. Those cells are mesoderm

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_10,

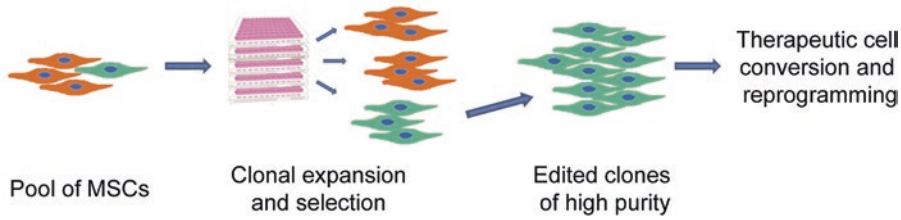


Fig. 1 Schematic presentation of genome editing using mesenchymal cells as a platform

in origin and have very high differentiation potential and remarkable self-renewal activity. It is not difficult to obtain 10^8 cells from a single MSC. Therefore, it is possible to generate patients' corrected clones of high purity with these MSCs. Moreover, the advance on cell programming/reprogramming makes it possible to differentiate the high-purity clones to targeted cell types.

In this current protocol, we provide a simple method to perform ZFN-mediated homologous recombination of 11 nucleotides in MSCs. By using MSCs as a platform we can generate highly homogenous gene-edited clones that may be useful for future research (Fig. 1).

2 Materials

Prepare all tissue culture solutions and equipment under sterile conditions.

2.1 Preparation of Platelet Lysate

1. Expired allogenic platelet stored at −80 °C in blood bank unit.
2. Sterile conical centrifuge tubes (50 mL).
3. Water bath (37 °C).
4. Centrifuge.

2.2 Isolation and Culture of Bone Marrow-Derived MSCs

1. Deionized (DI) water (*see* **Note 1**).
2. Heparinized bone marrow from donor (*see* **Note 2**).
3. Phosphate buffer saline (PBS): Dissolve 140 mM NaCl, 2.7 mM KCl, 1.5 mM KH_2PO_4, and 8.1 mM Na_2HPO_4 in DI water, and adjust to pH 7.4 with 1 N NaOH. Autoclave at 121 °C for 30 min (*see* **Note 3**). Store at room temperature.
4. Ficoll-Paque Plus (GE Healthcare, Piscataway, NJ, USA).
5. 10% Platelet lysate in Dulbecco's modified Eagle's medium (DMEM): Store at 4 °C.
6. Sterile conical centrifuge tubes (50 mL).
7. Cell culture flasks (25 mL).
8. Water bath (37 °C).

9. Centrifuge.
10. Biosafety cabinet class II.
11. CO_2 incubator, set at 37 °C, with 5% CO_2 and 90% humidity.

2.3 Isolation and Culture of Dental Pulp-Derived MSCs

1. DI water.
2. Fresh pulp tissue (*see* **Note 4**).
3. 3 mg/mL Collagenase type I: Dissolve 30 mg of collagenase type I in DI water and then filtrate the solution through a 0.2 μm membrane filter.
4. 10% Platelet lysate in DMEM with 100 units/mL of penicillin and 100 mg/mL of streptomycin.
5. Sterile conical centrifuge tubes (50 mL).
6. Cell culture flasks (25 mL).
7. Water bath (37 °C).
8. Centrifuge.
9. Shaker (37 °C).
10. Biosafety cabinet class II.
11. CO_2 incubator, set at 37 °C, with 5% CO_2 and 90% humidity.

2.4 Characterization of MSCs

1. PBS (*see* **step 3**, Subheading 2.2).
2. 0.25% Trypsin-EDTA.
3. Antibodies against CD34, CD45, HLA-DR, HLA-ABC, CD79a, CD29, CD33, CD44, CD73, CD90, CD11b, and mouse IgG conjugated to fluorescein isothiocyanate (FITC), phycoerythrin (PE), or phycoerythrin-Cy 7 (PE-Cy7).
4. BD FACS Sheath Solution™ (BD Biosciences, San Diego, CA, USA).
5. BD Biosciences FACS Calibur.

2.5 PCR Amplification of DNA Homologous to ~750 bp Upstream and Downstream of the ZFN Cutting Site

1. Template DNA: extract genomic DNA from peripheral blood (concentration 10 ng/μL).
2. Design pairs of primers for amplification of genomic DNA from −750 bp upstream of the left ZFN-binding site to +750 bp downstream of the right ZFN-binding site (Fig. 2) (*see* **Note 5**). Synthesize primers at the scale of 0.05 μmol.
3. 5 U/μL *Taq* DNA polymerase.
4. 10× PCR buffer without Mg^{2+}.
5. 50 mM $MgCl_2$.
6. 10 mM dNTP mix.
7. DI water.
8. 0.2 mL PCR tube.
9. Thermocycler.

Fig. 2 Design of PCR primers for amplification of ZFN-targeted sites. Forward primers anneal to around −750 bp upstream of the ZFN cutting site and the reverse primers anneal to around +750 downstream. Exogenous sequence can be inserted into the PCR products

2.6 Generation of Donor Plasmid for Targeted Integration into ZFN Cutting Site

1. PCR product from Subheading 2.5.
2. TA cloning vector kit (RBC Bioscience, Taipei, Taiwan).
3. SoloPack® Gold competent cells (Agilent Technologies; Santa Clara, CA, USA).
4. SOB medium: Dissolve 20 g of tryptone, 5 g of yeast extract, and 0.5 g of NaCl in 1 L DI water. Autoclave at 121 °C for 30 min. Add 10 mL of filtered sterile 1 M $MgCl_2$ and 10 mL filtered sterile 1 M $MgSO_4$ prior to use.
5. SOC medium: Add 2 mL of filtered sterile 20% (w/v) glucose to a final volume of 100 mL SOB medium (*see* **Note 6**).
6. LB agar containing ampicillin: Dissolve 10 g of tryptone, 5 g of yeast extract, 10 g of NaCl, and 20 g of agar in 1 L DI water, and adjust pH to 7 with 5 N NaOH. Autoclave at 121 °C for 30 min and cool to 55 °C. Add filtered sterile ampicillin to a final concentration of 100 μg/mL and pour 25 mL of medium into each 10 cm petri dish.
7. LB broth containing ampicillin: Dissolve 10 g of tryptone, 5 g of yeast extract, and 10 g of NaCl in 1 L DI water, and adjust pH to 7 with 5 N NaOH. Autoclave at 121 °C for 30 min and cool to 55 °C. Add filtered sterile ampicillin to a final concentration of 100 μg/mL.
8. Blue-white selection agar plates: Spread 100 μL of 10 mM IPTG and 100 μL of 2% X-gal on LB agar containing ampicillin at 30 min before cells are plated (*see* **Note 7**).
9. Water bath (42 °C).
10. Shaker (37 °C).
11. Incubator (37 °C).

2.7 Site-Directed Mutagenesis for Inserting the Universal Stop Codon TAGATA GTTAG Sequence to Donor Plasmid for Targeted Integration into ZFN Cutting Site

1. DNA template (10–100 ng): from Subheading 2.6.
2. Pairs of mutagenic primers containing 10–15 bases of mutagenic sequence in the middle of primers as well as non-mutagenic sequence that anneals to the complementary sequence on the opposite strands of plasmid (*see* **Note 8**).
3. QuikChange® Lightning Site-Directed Mutagenesis kit (Agilent Technologies).

4. DI water.
5. 2-Mercaptoethanol (β-ME).
6. 14 mL Polypropylene round-bottom tubes.
7. NZY+ Broth: Dissolve 10 g of casein hydrolysate, 5 g of yeast extract, and 5 g of NaCl in one liter DI water, and adjust pH to 7.5 with 5 N NaOH. Autoclave at 121 °C for 30 min. Prior to use, add 12.5 mL of filtered sterile 1 M $MgCl_2$, 12.5 mL of 1 M $MgSO_2$, and 20 mL of 20% (w/v) glucose.
8. LB agar containing ampicillin (*see* **step 6**, Subheading 2.6).
9. Blue-white selection agar plates (*see* **step 8**, Subheading 2.6).
10. Thermocycler.
11. Water bath (42 °C).
12. Shaker (37 °C).

2.8 ZFN Targeting CCR5 Gene in MSCs

1. CompoZr® Custom ZFN (human CCR5): mRNA of ZFN pairs specific to human *CRR5* gene.
2. Donor plasmid (*see* Subheading 2.7).
3. 2×10^6 MSCs at log phase (passages 3–5) (*see* **Note 9**).
4. Neon® Transfection system (Invitrogen).
5. Sterile PBS (*see* **step 3**, Subheading 2.2).
6. 10% Platelet lysate in DMEM (*see* **step 5**, Subheading 2.2).
7. 0.25% Trypsin-EDTA.
8. Hemocytometer.
9. Sterile conical centrifuge tubes (50 mL).
10. Cell culture flasks (25 mL).
11. Centrifuge.

2.9 Detection of HDR-Mediated Stop Codon Insertion to MSC Genome

1. ZFN-edited MSCs, harvested at day 5 after treatment (*see* **Note 10**).
2. Genomic DNA extraction kit.
3. Pairs of primers where forward primers anneal to outside of the homologous arm and reverse primers anneal to the targeted insertion region (Fig. 3).
4. *Taq* DNA polymerase (5 U/μL).

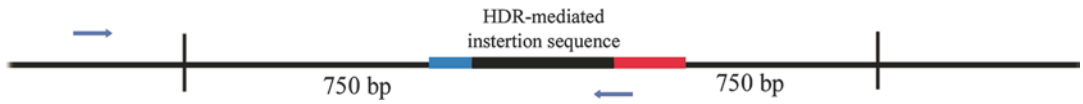


Fig. 3 Design of paired primers for detection of HDR-mediated insertion to MSC genome. Forward primers anneal to outside of the homologous arm and reverse primers anneal to the targeted insertion region

5. 10× PCR buffer without Mg^{2+}.
6. 50 mM $MgCl_2$.
7. 10 mM dNTP mix.
8. DI water.
9. PCR tubes (0.2 mL).
10. Water bath.
11. Spectrophotometer.
12. Thermocycler.

3 Methods

Perform all the tissue culture procedures under sterile conditions.

3.1 Preparation of Platelet Lysate

1. Thaw the frozen platelet at 37 °C in water bath (*see* **Note 11**).
2. After platelet is thawed, centrifuge at 3000 rpm for 10 min at room temperature [14].
3. Remove the supernatant and filtrate through 0.2 μm membrane filter.
4. Aliquot platelet lysate and store at −20 °C.

3.2 Isolation and Culture of Bone Marrow-Derived MSCs

1. Pipette 10 mL of heparinized donor-derived bone marrow into 50 mL sterile conical centrifuge tubes, then add 10 mL PBS, and mix briefly (*see* **Note 12**).
2. Pipette 9 mL Ficoll-Paque plus solution into 50 mL sterile conical centrifuge tubes.
3. Pipette 10 mL of the mixture from **step 1** and carefully overlay on Ficoll-Paque plus solution.
4. Centrifuge at 400 *g* for 30 min at 20 °C.
5. After centrifugation, the mixture is separated to four layers. The upper layer is composed of plasma and platelets, the interphase composed of mononuclear cells, and the upper bottom and the bottom composed of Ficoll-Paque solution and granulocyte erythrocytes, respectively [15].
6. Carefully aspirate all of the interphase layer without contaminating the others. The fluid volume of interphase is typically 1.5–2.0 mL.
7. Transfer the fluid of the interphase to a new 50 mL conical centrifuge tube, then add three volumes of PBS to the tube, and gently mix by pipetting.
8. Centrifuge at 100 *g* for 10 min at 20 °C and remove the supernatant.
9. Resuspend the mononuclear cells in 8 mL PBS by gently pipetting.

10. Centrifuge at 100 *g* for 10 min at 20 °C and remove the supernatant.
11. Resuspend the mononuclear cells in 5 mL of 10% platelet lysate in DMEM and transfer to a 25 mL culture flask.
12. Incubate the cell culture in 5% CO_2 at 37 °C.
13. Change medium every 3 days until the cells reach confluence.

3.3 Isolation and Culturing of Dental Pulp-Derived MSCs

1. Cut the pulp tissue into small pieces with a scalpel blade and then transfer to 15 mL sterile conical centrifuge tubes [13].
2. Add 1 mL of 3 mg/mL collagenase type I to the tissue.
3. Incubate the tube in the shaker at 37 °C with shaking at 220 rpm for at least 1 h.
4. Add 5 mL of 10% platelet lysate in DMEM supplemented with 100 unit/mL of penicillin and 100 mg/mL of streptomycin and then centrifuge at 1200 rpm for 10 min at 4 °C.
5. Remove the supernatant, then resuspend the cell pellet in 5 mL of 10% platelet lysate in DMEM, and transfer to a 25 mL culture flask.
6. Incubate the cell culture in 5% CO_2 at 37 °C.
7. Change medium every 3 days until the cells reach confluence.

3.4 Characterization of MSCs

1. Grow MSCs to 70–90% confluency (passage 3–5) in a 25 mL cell culture flasks.
2. Trypsinize the cells with 0.25% trypsin-EDTA.
3. Wash the cells by adding 5 mL PBS and centrifuge at 200 *g* for 5 min.
4. Remove the supernatant and resuspend the cells in PBS at a concentration of 1×10^7 cells/mL.
5. Dispense 100 μL of the cell suspension from **step 4** to 1.5 mL Eppendorf tubes.
6. Add 5 μL of each antibody as described in **step 3**, Subheading 2.4, to each tube and incubate at room temperature in the dark for 30 min.
7. Wash the cells twice by adding 5 mL PBS and centrifuge at 200 *g* for 5 min.
8. Resuspend the cells in 300 μL of FACS solution and transfer to BD FACS loader tubes.
9. Analyze the cells on a FACS Calibur system with Cell Quest Pro software.

3.5 PCR Amplification of DNA Homologous to ~750 bp Upstream and Downstream of the ZFN Cutting Site

1. Dissolve the primers with sterile DI water to a final concentration of 100 μM as stock solution.
2. Dilute the 100 μM stock primers to 10 μM (working concentration) with sterile DI water.

3. Add the components of the PCR master mix to each tube as follows: 32.1 μL of sterile DI water, 5 μL of 10× PCR buffer without Mg^{2+}, 1.5 μL of 50 mM $MgCl_2$, 1 μL of 10 mM dNTP mix, and 0.4 μL of *Taq* DNA polymerase. The total volume of each reaction is 40 μL. Mix by pipetting and briefly centrifuge the mixture.
4. Add 2.5 μL of both 10 μM forward and reverse primers and then 5 μL of DNA templates to the mixture from **step 3**.
5. Mix all components by pipetting and then briefly centrifuge.
6. Incubate the reactions in a thermocycler under the following conditions: 94 °C for 3 min; 30 cycles of 94 °C for 45 s, 56 °C for 30 s, and 72 °C for 150 s; and finally 72 °C for 10 min. Then keep the reactions at 4° C.
7. Determine the PCR products by running on 1.5% agarose gel.
8. Confirm the PCR products by DNA sequencing.

3.6 Generation of Donor Plasmid for Targeted Integration into ZFN Cutting Site

1. Thaw components of TA cloning vector kit at room temperature (*see* **Note 13**).
2. Briefly centrifuge TA cloning vector and DNA products from Subheading 3.5 to collect the contents at the bottom of the tube.
3. Add the following components to a 0.2 mL tube (*see* **Note 14**): 1 μL of ligation buffer A, 1 μL of ligation buffer B, 2 μL of TA cloning vector, 1.5 μL of PCR product (from Subheading 3.4), 1 μL of T4 DNA ligase, and 3.5 μL of sterile DI water.
4. Mix the reaction solution by pipetting.
5. Incubate the reactions for 30 min at 22 °C.
6. Preheat SOC medium in 42 °C water bath.
7. Thaw 25 μL aliquot of the competent cells on ice.
8. Add 2 μL of the ligated DNA from **step 5** to the competent cells.
9. Gently swirl the tube to mix DNA with the competent cells, and then incubate the tube on ice for 30 min.
10. Heat shock the competent cells by immersing the tube in 42 °C water bath for 60 s (*see* **Note 15**).
11. Immediately put the tube on ice and incubate for 2 min.
12. Add 175 μL of 42 °C SOC medium to the tube.
13. Incubate the tube in the shaker at 37 °C with shaking at 250 rpm for 1 h.
14. Evenly spread 100 μL of the transformed competent cells on blue-white selection agar plates.
15. Incubate the plate upside down at 37 °C incubator overnight.

16. Pick the white colonies and transfer to 200 μL of LB broth containing 100 μg/mL ampicillin in a 1.5 mL tube.
17. Incubate the tube in the shaker at 37 °C with shaking at 250 rpm for at least 1 h until the LB broth becomes turbid.
18. Take out 5 μL of the culture to perform the PCR reactions (*see* **steps 1**–7 in Subheading 3.5).
19. Scale up by transferring the remaining culture from **step 17** to 50 mL of LB broth containing 50 μg/mL ampicillin in a 250 mL Erlenmeyer flask and incubate in the shaker at 37 °C with shaking at 220 rpm overnight (no more than 16 h).
20. Extract the plasmid DNA from the bacteria.
21. Confirm the plasmid by DNA sequencing.

3.7 Site-Directed Mutagenesis of the Donor Plasmid for Insertion of the Universal Stop Codon for Targeted Integration into the ZFN Cutting Site

1. Dilute the mutagenic primers from **step 2**, Subheading 2.7, with sterile DI water to a final concentration of 100 μg/mL.
2. Add the following components to a 0.2 mL PCR tube: 5 μL of the 10× reaction buffer, 2 μL of 50 ng/μL plasmid DNA from Subheading 3.5, 1.25 μL of each 100 μg/mL mutagenic primers 1 and 2, 1 μL of dNTP mix (*see* **Note 16**), 1.5 μL of Quik Solution, and 38 μL of sterile DI water.
3. Add 1 μL of QuikChange lightning enzyme to the tube.
4. Incubate the reactions in a thermocycler under the following condition: 95 °C for 2 min; 18 cycles of 95 °C for 20 s, 60 °C for 10 s, and 68 °C for 90 s; and finally 68 °C for 5 min (*see* **Note 17**).
5. Add 2 μL of *Dpn* I restriction enzyme to the solution from **step 4**. Then, gently and thoroughly mix the solution by pipetting.
6. Briefly centrifuge the mixture to the bottom of the tube and then incubate at 37 °C for 5 min.
7. Thaw the XL-10 Gold ultracompetent cells on ice.
8. Prechill the 14 mL polypropylene round-bottom tube on ice.
9. Preheat NZY$^+$ broth in 42 °C water bath.
10. Take 45 μL of the XL 10 Gold® ultracompetent cells to the 14 mL polypropylene tube.
11. Add 2 μL of β-ME to the competent cells, and then gently swirl the tube.
12. Incubate the tube on ice for 2 min.
13. Transfer 2 μL of *Dpn* I-treated DNA from **step 6** to the tube.
14. Gently swirl the tube to mix DNA with the competent cells, and then incubate the tube on ice for 30 min.

15. Heat shock the competent cells by immersing the tube in a 42 °C water bath for 30 s (*see* **Note 18**).
16. Immediately put the tube on ice and incubate for 2 min.
17. Add 500 μL of preheated NZY^+ broth to the tube.
18. Incubate the tube in the shaker at 37 °C with shaking at 250 rpm for 1 h.
19. Evenly spread 100 μL of the transformed competent cells on blue-white selection agar plates.
20. Incubate the plate upside down at 37 °C overnight.
21. Pick the blue colonies and transfer to 200 μL LB broth containing 100 μg/mL ampicillin in a 1.5 mL tube.
22. Incubate the tube in a shaker at 37 °C with shaking at 250 rpm for at least 1 h until the LB broth becomes turbid.
23. Scale up by transferring the culture from **step 22** to 50 mL LB broth containing 100 μg/mL ampicillin in a 250 mL Erlenmeyer flask and incubate in a shaker at 37 °C with shaking at 220 rpm overnight (no more than 16 h).
24. Extract the plasmid DNA from bacteria.
25. Confirm the sequence of the mutagenic donor plasmid by DNA sequencing.

3.8 ZFN-Mediated Gene Integration in MSCs

1. Plate 1×10^5 cells/mL of MSCs (passages 2–5) into a 25 mL culture flask for 24 h.
2. Trypsinize the cells with 0.25% trypsin-EDTA.
3. Wash the cells by adding 10 mL PBS and centrifuge at 200 *g* for 5 min.
4. Remove the supernatant and count the number of cells.
5. Transfer 2×10^6 cells to 15 mL sterile centrifuge tube and then add PBS to a final volume of 10 mL.
6. Centrifuge at 200 *g* for 5 min and then remove the supernatant.
7. Resuspend the cell pellet in 200 μL of resuspension buffer R and transfer to a 1.5 mL Eppendorf tube (*see* **Note 19**).
8. Add 5 mL of DMEM medium containing 10% platelet lysate into a 25 mL culture flask and then pre-warm at a 37 °C incubator with 5% CO_2 for at least 15 min.
9. Add 3 mL of buffer E2 to the Neon tube in the Neon pipette station.
10. Set the parameters for electroporation: 990 V of pulse voltage, 40 ms of pulse duration, and one cycle of pulse.
11. Transfer one vial of ZFN mRNA and 10 μg of donor plasmid from Subheading 3.7 to a 1.5 mL Eppendorf tube.

12. Add 200 μL of the cell suspension in resuspension buffer R into the tube containing ZFN mRNA and donor plasmid.
13. Gently mix the solution by pipetting.
14. Insert the Neon Tip into the Neon pipette and aspirate the mixture of cells containing ZFN mRNA and donor plasmid into the tip without air bubbles.
15. Insert the Neon pipette with the mixture of cells containing ZFN mRNA and donor plasmid into the Neon pipette station and then press start on the touchscreen to start electroporation.
16. After electroporation, remove the Neon pipettes from the station and then transfer the mixture of cells containing ZFN mRNA and donor plasmid into pre-warmed culture medium in 25 mL culture flasks.
17. Culture the ZFN-edited cells in a 37 °C incubator with 5% CO_2.

3.9 Detection of HDR-Mediated Stop Codon Insertion to MSC Genome

1. At day 5 after transfection, trypsinize the ZFN-edited cells with 2.5% trypsin-EDTA.
2. Count the cells using hemocytometer and then transfer 10^6 cells to 1.5 mL Eppendorf tube.
3. Centrifuge at 2000 rpm for 5 min and then remove the supernatant.
4. Resuspend the cell pellet with 150 μL RBC lysis buffer.
5. Add 200 μL GT buffer to the tube and then vortex for 5 s to mix the components.
6. Incubate the mixture at 70 °C for 10 min. Invert the tube every 3 min during incubation.
7. Preheat RBC elution buffer at 70 °C.
8. After 10-min incubation, add 200 μL of 96% ethanol to the cell lysate and then vortex immediately for 10 s to mix the components.
9. Put the GD column into the 2 mL collection tube and then transfer all the cell lysate to the GD column.
10. Close the cap and then centrifuge at 10,000 *g* for 2 min.
11. Remove the flow-through and then return the GD column to the collection tube.
12. Add 400 μL of W1 Buffer to the GD column and then centrifuge at 10,000 *g* for 30 s.
13. Remove the flow-through and then return the GD column to the collection tube.
14. Add 600 μL of wash buffer to the GD column and then centrifuge at 10,000 *g* for 30 s.

15. Remove the flow-through and then return the GD column to the collection tube.
16. Centrifuge at 15,000 *g* for 3 min to dry the column.
17. Put the dried GD column into a 1.5 mL sterile Eppendorf tube and then add 100 μL of preheated elution buffer to the center of the column matrix.
18. Leave the GD column to stand for 2 min for the absorption of the elution buffer and then centrifuge at 10,000 *g* for 30 s to elute DNA.
19. Determine DNA concentration using spectrophotometer.
20. Dilute the 100 μM stock primers to 10 μM (working concentration) with sterile DI water.
21. Add the following components to an Eppendorf tube on ice to prepare the master mix (*see* **Note 20**): 32.1 μL of sterile DI water, 5 μL of 10× PCR buffer without Mg^{2+}, 1.5 μL of 50 mM $MgCl_2$, 1 μL of 10 mM dNTP mix, and 0.4 μL of *Taq* DNA polymerase. Mix by pipetting and then briefly centrifuge. Dispense 40 μL of the mixture to each PCR tube.
22. Add 2.5 μL of each 10 μM forward and reverse primers and then 5 μL of 10 ng/μL genomic DNA to the mixture in each PCR tube.
23. Mix all components by pipetting and briefly centrifuge.
24. Incubate the reactions in a thermocycler under the following conditions: 94 °C for 3 min; 30 cycles of 94 °C for 45 s, 57 °C for 30 s, and 72 °C for150 s; and finally 72 °C for 10 min. Keep the reaction at 4 °C on hold.
25. Determine the PCR products by running on 1.5% agarose gel.

4 Notes

1. DI water is the type III purified water, which has 0.5 MΩ cm of resistivity, 15 μS/cm of conductivity, and less than 5 ppb of total organics.
2. We use the standard protocol to isolate MSCs from leftover bone marrow aspiration samples. The study protocol is approved by the Ethical Committee of Lerdsin General Hospital.
3. Storage of PBS solution at high temperature may lead to precipitation. Therefore, do not let stand the solution for long time in autoclave machine.
4. Impacted wisdom teeth (third molar) are removed and collected from patients at the Dental Department of Lerdsin General Hospital with the approval of Lerdsin General Hospital's ethic committees. Tooth surfaces are cleaned with

70% ethanol and then cut around the cementoenamel junction by using sterilized dental fissure burs to reveal the pulp chamber. The pulp tissue is gently separated from the crown and roots under sterile conditions. These procedures of stem cell isolation are modified from a previous study [13].

5. To generate donor plasmids for homologous recombination, the homologous regions must be at least 750 bp at each end of the cutting site.
6. SOC medium should be prepared immediately prior to use.
7. Prepare the 2% X-gal and 10 mM IPTG from stock solution using SOC medium before spreading. Do not mix X-gal and IPTG directly without SOC medium because the mixture may precipitate.
8. The length of mutagenic primers should be between 25 and 45 bp and the melting temperature of primers should be more than or equal to 78 °C. The minimum GC content of primers should be 40%. Primers should terminate in one or more C or G base. The 5′ end of primer does not need to be phosphorylated.
9. The MSCs are initially plated at a density of 1×10^4 cells/cm^2. The cells reach log phase at day 4 after plating.
10. At day 1 to day 3 of ZFN mRNA transfection, toxicity on transfected cells may be observed. Cells can proliferate at a regular rate at 5–7 days after transfection. Therefore, we harvest the cells at day 5 for detection of targeted integration and clonal selection.
11. To avoid gel formation in cultured medium with supplemented platelet lysate, we add 500 μL of heparin (100 units/mL) to 50 mL of the thawed platelet.
12. Follow standard procedures and wear lab coats, gloves, and eye protection when working with human tissues.
13. We aliquot 10 μL of TA cloning vectors into each 0.2 mL PCR tube to avoid multiple freezing and thawing that may damage the vector.
14. If there is any precipitation after buffer A in the TA cloning kit is thawed, the solution should be vortexed until the precipitates disappear. If there is any precipitation after buffer B is thawed, the solution should be heated at 60 °C for 5 min or until the precipitation is dissolved.
15. The temperature and duration in heat-shock step are critical for the efficiency of transformation.
16. After thawing dNTP mix for the first time, the solution is aliquoted into 0.2 mL PCR tubes with 3 μL per tube and stored at −20 °C to avoid multiple freeze and thaw.
17. The duration of elongation step in the second segment of cycling depends on the plasmid length. It is recommended to

use 30 s for per 1 kb of plasmid length. In the given example in this protocol, 90 s is used for a plasmid of approximately 3 kb.

18. The temperature and duration in the heat-shock step are critical for the transformation efficiency. The heat shock in this step should be optimized in 14 mL BD Falcon polypropylene round-bottom tube.
19. Do not store the cell suspension in buffer R for more than 15 min as extended incubation may reduce the viability of cells, which leads to low transfection efficiency.
20. We always prepare master mix if multiple reactions are performed to avoid volume loss.

References

1. Santiago Y, Chan E, Liu PQ, Orlando S, Zhang L, Urnov FD et al (2008) Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci 105:5809–5814
2. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O et al (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 7:808–816
3. Zou J, Mali P, Huang X, Dowey SN, Cheng L (2011) Site-specific gene correction of a point mutation in human iPS cells derived from an adult patient with sickle cell disease. Blood 118:4599–4608
4. Manotham K, Chattong S, Setpakdee A (2015) Generation of CCR5-defective CD34 cells from ZFN-driven stop codon-integrated mesenchymal stem cell clones. J Biomed Sci 26:22–25
5. Chattong S, Ruangwattanasuk O, Yindeedej W, Setpakdee A, Manotham K (2017) CD34+ cells from dental pulp stem cells with a ZFN-mediated and homology-driven-repair-mediated locus-specific knock-in of an artificial β-globin gene. Gene Ther 24:425–432
6. Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300:764
7. Carroll D (2008) Progress and prospects: zinc-finger nucleases as gene therapy agents. Gene Ther 15:1463–1468
8. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300:763
9. Moehle EA, Rock JM, Lee YL, Jouvenot Y, DeKelver RC, Gregory PD et al (2007) Targeted gene addition into a specified location in the human genome using designed zinc finger nucleases. Proc Natl Acad Sci U S A 104:3055–3060
10. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M et al (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31:294–301
11. Castro-Malaspina H, Gay RE, Resnick G, Kapoor N, Meyers P, Chiarieri D et al (1980) Characterization of human bone marrow fibroblast colony-forming cells (CFU-F) and their progeny. Blood 56:289–301
12. Boquest AC, Collas P (2012) Obtaining freshly isolated and cultured mesenchymal stem cells from human adipose tissue. Methods Mol Biol 879:269–278
13. Gronthos S, Mankani M, Brahim J, Robey PG, Shi S (2000) Postnatal human dental pulp stem cells (DPSCs) *in vitro* and *invivo*. Proc Natl Acad Sci U S A 97:13625–13630
14. Capelli C, Domenghini M, Borleri G, Bellavita P, Poma R, Carobbio A et al (2007) Human platelet lysate allows expansion and clinical grade production of mesenchymal stromal cells from small samples of bone marrow aspirates or marrow filter washouts. Bone Marrow Transplant 40:785–791
15. Smith JR, Pochampally R, Perry A, Hsu SC, Prockop DJ (2001) Isolation of a highly clonogenic and multipotential subfraction of adult stem cells from bone marrow stroma. Stem Cells 22:823–831

Part II

Methods for the Evaluation and Prevention of ZFN Cytotoxicity

Chapter 11

Integrated Multimodal Evaluation of Genotoxicity in ZFN-Modified Primary Human Cells

Jaichandran Sivalingam, Dimitar Kenanov, Wai Har Ng, Sze Sing Lee, Toan Thang Phan, Sebastian Maurer-Stroh, and Oi Lian Kon

Abstract

Iatrogenic adverse events in clinical trials of retroviral vector-mediated gene-corrected cells have prioritized the urgent need for more comprehensive and stringent assessment of potentially genotoxic off-target alterations and the biosafety of cells intended for therapeutic applications. Genome editing tools such as zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced palindromic repeats (CRISPR)-Cas9 nuclease systems are being investigated as safer and efficient alternatives for site-directed genome modification. Using site-specific integration into the AAVS1 locus of primary human cells as an example, we present an integrated approach to multimodal investigation of off-target alterations and an evaluation of potential genotoxicity induced by ZFN-mediated integration of a therapeutic transgene.

Key words Zinc finger nuclease, AAVS1 site-specific, Off-target, Genotoxicity, Transgene integration, Cell and gene therapy, Primary human cells, Whole-genome sequencing, RNA-seq, bioinformatics

1 Introduction

The use of gene-modified cells to correct or mitigate genetic disorders has been investigated for decades in the field of gene and cell therapy [1]. Although conventional gene therapy strategies which use retroviral vectors to integrate cDNA permanently into cellular genomes have proved efficacious in clinical trials of life-threatening monogenic disorders, these have been complicated and set back by serious iatrogenic adverse complications [2].

Cases of insertional oncogenesis due to the quasi-random nature of gene integration mediated by retroviral vectors have driven vigorous efforts toward safer strategies. In this regard, ex vivo-modified cells have advantages over in vivo gene modifications as modified cells can be assessed for the desired functional phenotype and potentially genotoxic off-target alterations prior to

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_11,

clinical treatment of patients. An important biosafety precaution is to target DNA integration into genomic safe harbors, specific sites which enable persistent transgene expression without dysregulating the expression of neighboring and other endogenous genes, especially tumor suppressors and oncogenes [3]. Site-specific modification of preselected chromosomal safe harbors requires induction of highly accurate double-strand breaks in genomic DNA whose repair by error-prone nonhomologous end joining or homology-directed repair achieves the desired modification. Homology-directed repair in the presence of exogenously supplied homologous donor DNA can correct a mutant gene or integrate a therapeutic transgene. Such targeted modifications are currently made mainly with zinc finger nucleases (ZFNs), transcription activator-like effector nucleases (TALENs) and clustered regularly interspaced palindromic repeats, CRISPR/Cas nucleases [4].

ZFNs are synthetic protein chimeras composed of DNA-binding Cys_2-His_2 zinc fingers in tandem array fused with the catalytic domain of *Fok*I endonuclease and have been investigated widely for their ability to induce accurate site-specific genome modifications. As *Fok*I nuclease is active only as a dimer, ZFNs function as monomer pairs in inverted orientation. Each monomer typically consists of 4–6 fingers (each finger formed by 30 amino acids), recognizes, and binds to a 12 to 18 bp target half-site on opposite DNA strands. The entire recognition sequence of two half-sites with an intervening 5 to 7 bp spacer is long enough to target a unique site in the human genome. Several modifications to the *Fok*I catalytic domain have improved the specificity and activity of ZFNs by obligate heterodimerization [5], enhanced cleavage activity [6], and reduced toxicity and off-target activity [7, 8]. To date, ZFNs are more advanced toward clinical application than TALENs and CRISPR-Cas9, there being a current total of eleven completed and ongoing ZFN clinical trials (www.clinicaltrials.gov; accessed on 4 July 2018).

TALENs resemble ZFNs except for the DNA-binding domains which are tandem arrays of transcription activator-like (TAL) effectors. TALENs are as efficient as ZFNs but are more difficult to assemble and deliver, being much larger than ZFNs. More recently, CRISPR-Cas systems are set to transform genome editing by their simplicity, efficiency, versatility, and affordability. Basic CRISPR-Cas9 editing is performed by a nonspecific endonuclease, Cas-9, guided to the intended genomic site by a single guide RNA which supplies a targeting sequence of 17–20 nucleotides at its 5′end. Local duplex DNA strand separation and heterodimerization of the guide RNA with its DNA complement are followed by cleavage of both DNA strands at the target site. It is known that the basic monomeric CRISPR-Cas9 system is prone to frequent off-target cleavage because the short homology sequence for RNA-guided recognition lacks high specificity in large genomes.

Regardless of which genome editing tool is utilized, all are known to generate unintended off-target modifications which may have genotoxic or oncogenic effects. Accurately identifying these collateral events with high sensitivity is essential in the context of clinical applications because even low-frequency (<0.1%) off-target alterations may compromise biosafety. The most common off-target alterations are small indels; others are mutations, large deletions and insertions, and gross chromosomal rearrangements (inversions, translocations, and dicentric chromosomes). The frequency and nature of off-target alterations in any specific situation are unpredictable because cell type, genomic target site, intracellular DNA repair capacity, cellular transcriptional activity, delivery method, temporal conditions of nuclease activity, and other variables all influence the accuracy of genome editing. Moreover, as no single technique can detect all types of unintended modifications and their likely adverse functional consequences, a combination of different detection methods should be employed. Off-target detection methods may focus only on likely off-target sites predicted in silico or known from previous experience [9]. However, bioinformatic algorithms alone are not a sufficiently secure approach to identify off-target sites comprehensively and should be combined with unbiased methods to identify different types of off-target events. Unbiased methods tend to be technically demanding in execution and data analysis. Such methods currently are systemic evolution of ligands by exponential enrichment (SELEX), ChIP-seq, capture and sequencing of integrase-deficient lentivirus (IDLV) or adeno-associated virus integrations, array comparative genomic hybridization (aCGH), direct in situ breaks labeling, enrichment on streptavidin and next-generation sequencing (BLESS), Digenome-seq, GUIDE-seq, double-strand break sequencing, linear amplification-mediated high-throughput genome-wide translocation sequencing (LAM-HTGTS), whole-exome sequencing, and whole-genome sequencing [9, 10]. As not all off-target genomic alterations are deleterious to cell functions, it is helpful to complement genomic methods with unbiased transcriptome analysis and appropriate tumorigenic assays.

In this chapter we describe a suite of genomic methods combined with RNA-seq to assess potential genotoxicity and biosafety of primary human cells after ZFN-mediated integration of a therapeutic transgene into the human AAVS1 locus (chromosome 19q13.3).

2 Materials

2.1 Polymerase Chain Reaction

1. High-fidelity DNA polymerase.
2. DNA polymerase for long read PCR and GC-rich templates.
3. dNTP mix.

4. Microfuge tubes.
5. DNase-free water.
6. Gradient thermal cycler.
7. Agarose gel electrophoresis reagents and equipment.
8. DNA stain.
9. UV transilluminator to image and document DNA in agarose gels.

2.2 Plasmid DNA Cloning

1. Wild-type FokI cleavage domain cloned in pST1374 (Addgene; http://www.addgene.org).
2. Promoter trap vector for AAVS1 integration: pAAVS1 SA-2A-puro-pA donor (Addgene).
3. Restriction enzymes and buffers.
4. Shrimp alkaline phosphatase.
5. T4 DNA ligase and ligation reaction buffer.
6. Site-directed mutagenesis kit.
7. TA cloning vector.
8. Blunt PCR cloning vector.
9. DH5α *E. coli* competent cells.
10. LB agar plates, antibiotics, reagents and equipment for heat-shock transformation, and bacterial culture.

2.3 Plasmid DNA Extraction

1. Plasmid DNA extraction kit.
2. Isopropanol.
3. 70% Ethanol.
4. DNase-free water.
5. Polypropylene tubes.
6. Centrifuge.

2.4 Cell Culture

1. Class II biosafety cabinet.
2. Incubator (37 °C, 5% CO_2).
3. Water bath (37 °C).
4. Bright-field phase-contrast microscope.
5. PTTe3 cell culture medium for primary human umbilical cord-lining epithelial cells: Medium 171, 50 ng/mL insulin-like growth factor 1, 50 ng/mL platelet-derived growth factor BB, 5 μg/mL transforming growth factor β1, 5 ng/mL insulin.
6. Phosphate-buffered saline.
7. Cell dissociation enzyme (trypsin or similar) and reagents.
8. Tissue culture flasks.

9. 96-Well flat-bottom tissue culture plates.
10. Centrifuge.
11. 0.4% Trypan blue solution.
12. 0.1% Crystal violet.
13. Neubauer hemocytometer chamber.
14. Tetrazolium salt and reagents to assay cell proliferation.

2.5 Electroporation

1. Tissue culture flasks.
2. Cell culture medium.
3. Cell dissociation enzyme (trypsin or similar) and reagents.
4. Centrifuge.
5. Plasmids: ZFN expression plasmid, donor DNA plasmid, and a reporter plasmid, e.g., pEGFP.
6. Lonza 4D-Nucleofector™ system (Lonza, Basel, Switzerland).
7. Lonza primary cell Nucleofector™ P1 solution for cord-lining epithelial cells.
8. Lonza 4D-Nucleofector™ cuvettes.
9. Inverted fluorescence microscope with a filter for FITC (excitation 425 nm; emission 500 nm).
10. Flow cytometer configured to detect FITC emission.
11. Permeabilization solution: 0.5% Triton-100, 2% bovine serum albumin in PBS.
12. Fluorochrome-conjugated anti-phosphohistone H2AX (Ser139) antibody.

2.6 Genomic DNA Extraction and Whole-Genome Amplification

1. Genomic DNA extraction kit and manufacturer's instructions.
2. phi29 DNA polymerase and reaction buffer.
3. DNase-free water.

2.7 Heteroduplex Mismatch Cleavage Assay

1. Genomic DNA.
2. High-fidelity DNA polymerase and reaction buffer.
3. dNTP mix.
4. Thermal cycler.
5. Agarose gel electrophoresis equipment and reagents.
6. Gel extraction kit to purify DNA from agarose gel.
7. Spectrophotometer.
8. Purified PCR amplicons.
9. Mismatch cleaving endonuclease (e.g., Cel 1, T7 endonuclease 1, resolvase) and manufacturer's protocol.
10. Thermal cycler.

11. 30% Acrylamide/Bis solution (29:1) from commercial suppliers.
12. *N*,*N*,*N*′,*N*′-tetramethylethylenediamine.
13. Ammonium persulfate.
14. TBE buffer: 89 mM Tris, pH 8.3, 89 mM boric acid, 2 mM EDTA.
15. Gel casting and electrophoresis equipment.
16. DNA stain.
17. UV transilluminator.

2.8 Total RNA Extraction

1. RNase-decontaminating reagent.
2. Phenol-guanidinium isothiocyanate reagent.
3. Selective RNA-adsorbing reagents to purify crude RNA.
4. RQ1 RNase-free DNase I and reaction buffer.
5. DNase-RNase-free water.
6. Bioanalyzer (Agilent).
7. 1% Agarose gel in TAE buffer.
8. TAE buffer: 40 mM Tris–acetate, pH 8.2–8.4, 1 mM EDTA.
9. RNA loading buffer: 25% Ficoll 400, 0.25% bromophenol blue in RNase-free water.
10. UV transilluminator.

2.9 RT-PCR

1. Nuclease-free thin-wall tubes, 0.2 mL.
2. cDNA synthesis kit.
3. Thermal cycler.
4. 96-Well clear PCR plates or 8-strip PCR tubes with caps.
5. Gene-specific primers.
6. qPCR master mix.
7. Real-time PCR detection system.

2.10 Droplet Digital PCR

1. Droplet Digital PCR kit and manufacturer's protocol.
2. Droplet generator and droplet reader.
3. Thermal cycler.

3 Methods

3.1 Culture of Primary Cord-Lining Epithelial Cells

1. Cannulate a fresh human umbilical cord and flush with sterile phosphate-buffered saline (PBS) containing 5 IU/mL heparin to remove blood (*see* **Note 1**).
2. Cut the cord into 2 cm segments, wash and disinfect with 70% ethanol, and wash again with sterile PBS-containing antibiotics

(50 IU/mL penicillin, 50 μg/mL streptomycin, 250 μg/mL amphotericin B, and 50 μg/mL gentamicin).

3. Dissect the amniotic membrane free from other cord contents using sterile technique, cut into 0.5 cm squares, and place the cut squares in a cell culture dish containing 5 mL of PTTe3 culture medium.
4. Culture the explants at 37 °C and 5% CO_2; change the medium every third day.
5. Trypsinize outgrowing cord-lining epithelial cells (CLECs) and seed them at a density of 5000 cells/cm^2 to expand the culture.
6. Split the culture (1:10) at 70% confluence in fresh culture medium. Renew the culture medium every 3–4 days.
7. If CLECs are not used immediately, cryopreserve in culture medium containing 10% DMSO (5 × 10^6 cells/mL).

3.2 Assembly of Plasmid Constructs

3.2.1 AAVS1 ZFN Expression Plasmid

1. Mutagenize DNA coding the wild-type catalytic domain of *Fok*I in pST1374; for obligate heterodimerization (OH) [5] use site-directed mutagenesis and oligonucleotides listed in Table 1 (*see* **Note 2**).
2. Mutagenize OH ZFN from **step 1** to enhance cleavage activity and reduce toxicity [6, 8] with the following amino acid substitutions: S418P, K441E, and H537R (right *Fok*I monomer) and S418P, K441E, and N496D (left *Fok*I monomer) using site-directed mutagenesis and oligonucleotides in Table 1 (*see* **Note 3**).
3. Synthesize DNA to encode a pair of zinc finger peptides targeted at the AAVS1 locus, 5′CCACCCCACAGTGGGG CCACTAGGGACAGGATTG 3′ (ZFN-binding sites underscored; *see* **Note 4**). Ligate the DNA sequence encoding each zinc finger peptide in frame to DNA encoding its corresponding mutagenized *Fok*I monomer. Clone each zinc finger peptide and its *Fok*I monomer in a separate plasmid.
4. Combine both right and left ZFN AAVS1-*Fok*I expression cassettes into a single plasmid (Fig. 1).

3.2.2 Promoter Trap Vector for AAVS1 Site-Specific Integration of Donor DNA

1. Linearize pAAVS1 SA-2A-puromycin-pA at the multiple cloning site between the homology arms (*see* **Note 5**). If necessary, polish and dephosphorylate the cut ends.
2. Ligate DNA expressing the therapeutic gene of interest between the cut ends to obtain the donor DNA vector.

3.3 Electroporation and Assessment of Transfection Efficiency

1. Trypsinize adherent cells (*see* **Note 6**), mix with an equal volume of fresh medium in 15 mL polystyrene tubes, and centrifuge at 500 × *g* for 3 min. Rinse the cell pellet once with PBS, centrifuge again, and resuspend the cells in PBS.

Table 1
Oligonucleotides for mutagenizing FokI endonuclease, PCR primers for Surveyor™ assay, top 10 predicted off-target sites, and 19q13 gene transcripts

Purpose	Forward (5′ → 3′)	Reverse (5′ → 3′)	Amplicon size (bp)	Annealing temp. (°C)	Section
Oligonucleotides for mutagenizing FokI endonuclease					
E490K mutation	gca acg ata tgt caa aga aaa tca aac acg	cgt gtt tga ttt tct ttg aca tat cgt tgc	Not applicable	55	3.2.1
I538K mutation	cac gat taa atc ata aga cta att gta atg gag c	gct cca tta caa tta gtc tta tga ttt aat cgt g	Not applicable	55	3.2.1
Q468E mutation	cca agc aga tga aat gga acg ata tgt cga ag	ctt cga cat atc gtt cca ttt cat ctg ctt gg	Not applicable	55	3.2.1
I499L mutation	cac gaa aca aac atc tca acc cta atg aat gg	cca ttc att agg gtt gag atg ttt gtt tcg tg	Not applicable	55	3.2.1
S418P mutation	att gaa att gcc aga aat ccc act cag gat aga att ctt	aag aat tct atc ctg agt ggg att tct ggc aat ttc aat	Not applicable	55	3.2.1
K441E mutation	gtt tat gga tat aga ggt gaa cat ttg ggt gga tca agg	cct tga tcc acc caa atg ttc acc tct ata tcc ata aac	Not applicable	55	3.2.1
H537R mutation	cag ctt aca cga tta aat cgt aag act aat tgt aat gga	tcc att aca att agt ctt acg att taa tcg tgt aag ctg	Not applicable	55	3.2.1
N496D mutation	gaa gaa aat caa aca cga gac aaa cat ctc aac cct aat	att agg gtt gag atg ttt gtc tcg tgt ttg att ttc ttc	Not applicable	55	3.2.1
Primers for Surveyor™ assay					
Amplify AAVS1 site	ttcgggtcacctctcactcc	ggctccatcgtaagcaaacc	469	60	3.5.2
Primers for integration junction PCR					
Amplify AAVS1 integration site	aagaagcgcaccacctccaggttctc	atgacctcatgctcttggccctcgta	976	60	3.6.1
Primers to amplify top 10 predicted off-target sites of AAVS1 ZFN					
OT1	tttaagaactgtaacctattttccaaagtgtttg	cag cct ggc caa cat ggt gaa ac	389	62	3.5.3
OT2	aaggtgtaagtggagccacaaggct	tgtggtccttgctggatcaggaa	308	60	3.5.3

OT3	ttggaaataagacccatttgttgatgaga	ctggctcattccaacgtccatgt	389	58	3.5.3
OT4	gacttggtggttggcagaatacacc	gggtaaggtcagatagggctgtaagactc	601	60	3.5.3
OT5	ggaacaaggcacctggctcc	ccattcccgggagaaatctc	353	58	3.5.3
OT6	tgagtttgggcctgaggtcatc	ggcttggaaacacccaggtg	320	60	3.5.3
OT7	ctttgagtttagcagcttccaggaacc	gttttatcttcataaggtagtgggcagatgg	631	60	3.5.3
OT8	ggtcctcaccccatcttcatc	aaagagagggctggtgaggc	375	60	3.5.3
OT9	gttgcgagagtccctactgg	agcctgaagttgagcctgtc	348	60	3.5.3
OT10	cacagagttcaggggatcgt	gccactttgtattgggtggt	650	60	3.5.3
Primers to verify transcriptome data by RT-PCR					
PPP1R12C exons 4 - 6	agaggaattgctccttcatgacac	caacaggctcagtacttcctcatc	322	60	3.6.3
LILRB4	ggacattggcccagagacag	gtcttcatcgtgtgggctct	197	60	3.6.3
ISOC2	tgacggagcagtacccacaag	ggataagaacgggggtccaagatg	196	60	3.6.3
PPP6R1	gcgctacaagtaccccagtg	tccgaagaaaggacacgagc	215	60	3.6.3
NAT14	gctcccgaaaccttgtcgaa	ccttcacgccggccttc	175	60	3.6.3
ZNF579	aaggaggcgaggcatggat	gtaggggaaacggaagaggc	184	60	3.6.3
FIZ1	cagagggaggtagagagccc	ctgtgcttgaaacccttgcc	235	60	3.6.3
RDH13	ggtccccttcccagctgaa	atgatgttgcctcctctcctg	278	60	3.6.3
DUSP1	ggatacgaagcgttttcggc	ggccaccctgatcgtagagt	151	60	3.6.3
DUSP6	acctggaaggtggcttcagta	accatccgagtctgttgcac	197	60	3.6.3
CDC6	agaagggccccatgattgtg	tgcagcattgtccagaacct	296	60	3.6.3
DUSP16	ctgcttgcaggtgggtttg	gcacacgcaggaaatgagac	271	60	3.6.3
GAPDH exons 6–7	gcctcctgcaccaccaact	cgcctgcttcaccaccttc	348	60	3.6.3

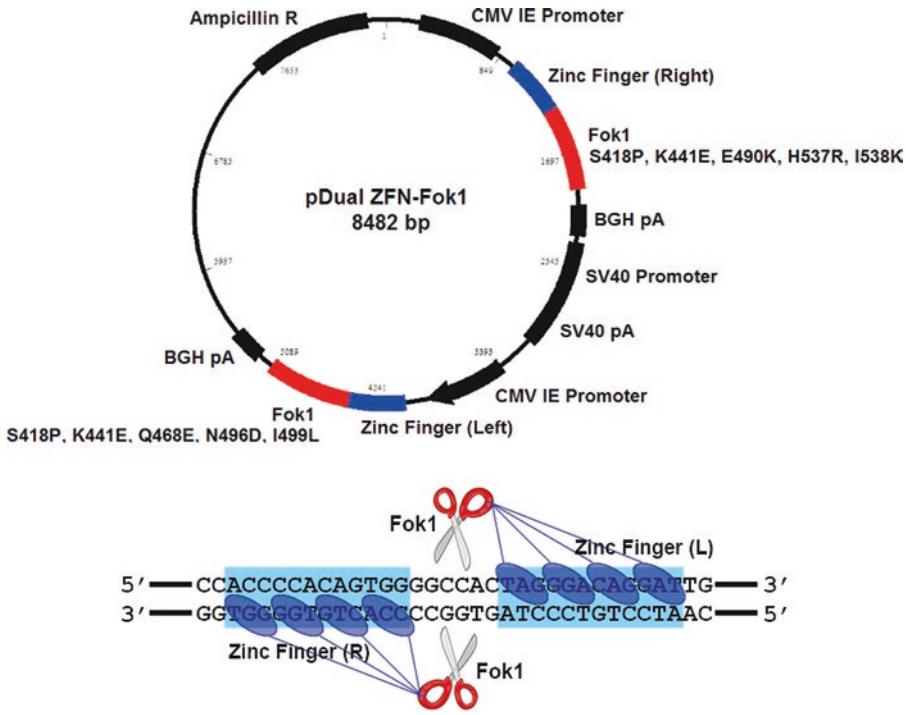


Fig. 1 Diagram of a single plasmid construct expressing right and left zinc finger DNA-binding peptides and *Fok*I monomers modified for obligate heterodimerization

2. Remove 10 μL of the cell suspension and determine cell concentration and viability in a hemocytometer with trypan blue.
3. Transfer 2–10 × 10^6 cells into a separate polystyrene tube, centrifuge, and remove the supernatant (*see* **Note 7**).
4. Electroporate primary cells using an Amaxa® 4D Nucleofector™, the appropriate proprietary electroporation solution and cuvette.
5. Resuspend the cell pellet from Subheading 3.3, **step 3**, in 100 μL of the appropriate Nucleofector™ solution containing 5 μg ZFN plasmid DNA alone or with 20 μg donor DNA. Transfer the mixture to a Nucleofector™ cuvette and electroporate using the recommended setting (*see* **Note 8**).
6. Immediately add 100 μL of pre-warmed (37 °C) cell culture medium to the electroporated contents of the cuvette and let stand for 5 min at room temperature. Then aspirate the entire contents of the cuvette with a sterile plastic Pasteur pipette and transfer to a T-75 tissue culture flask containing 10 mL of pre-warmed cell culture medium.
7. Culture the cells in a humidified 5% CO_2 incubator at 37 °C.
8. If a GFP reporter gene has been co-transfected, GFP-positive cells can be observed by fluorescence microscopy 8–12 h post-electroporation. Transfection efficiency is assessed by flow cytometry 24 h post-electroporation.

3.4 Cell Viability, Proliferation, and In Vitro Colony-Forming Assays

Assess cell viability by trypan blue exclusion and cell proliferation by the tetrazolium salt reduction assay.

1. Detach adherent cells by trypsin treatment, wash once with PBS, and resuspend in 1 mL PBS.
2. For trypan blue exclusion cell counts, load 10 μL of an equal mixture of cell suspension and 0.4% trypan blue solution into a hemocytometer by capillary action.
3. Count the number of unstained live cells and blue stained dead cells in triplicate aliquots. % Cell viability = (Number of live cells)/(Number of live + dead cells) × 100 (*see* **Note 9**).
4. Perform the tetrazolium salt reduction assay for cell proliferation using a commercial kit and the supplier's protocol.
5. Perform in vitro colony-forming assay by seeding 100 cells into each of triplicate wells of a 6-well plate. Wash the cells after 14 days in culture, fix in 10% formalin for 15 min at room temperature, rinse with PBS, and stain with 1% (w/v) crystal violet for 30 min at room temperature. Remove excess crystal violet with three PBS rinses and air-dry. Capture images of stained colonies with an inverted microscope, and score the number and size of colonies.

3.5 Assessing Site-Specific ZFN Activity

Assess the frequency of ZFN-induced double-stranded DNA breaks by flow cytometry detection of phosphorylated histone H2AX 2 days after transfection.

3.5.1 Phosphorylated Histone H2AX as an Indicator of ZFN Toxicity

1. Fix trypsinized cells with 3.7% formaldehyde in PBS for 10 min and 90% methanol at −20 °C for 2 h. Permeabilize fixed cells for 10 min and incubate with fluorochrome-conjugated anti-phosphohistone H2AX (Ser139) for 1 h at 25 °C.
2. Wash cells twice with PBS, resuspend in 500 μL PBS, filter through 40 μm nylon mesh into flow cytometry tubes, and detect positive cells using an appropriate laser for the fluorochrome. Analyze untreated cells as negative controls and cells treated with 10 μM etoposide for 1 h as positive controls for double-stranded DNA breaks.
3. Use flow cytometry analytical software (e.g., FlowJo) to determine the percentage of phosphoH2AX-positive cells (*see* **Note 10**).

3.5.2 Mismatch Cleavage Assay to Evaluate Site-Specific Genomic Cleavage

In the absence of donor DNA, ZFN-mediated site-specific cleavage of genomic DNA is repaired mainly by error-prone nonhomologous end joining (NHEJ) which generates indels at cleavage sites. Cel-I, Cel-II, and T7 endonuclease 1 are DNA endonucleases that specifically cleave mismatched heteroduplexes. Cleavage frequency at target site PCR amplicons can be quantitatively evaluated by a mismatch cleavage assay performed according to the

manufacturer's protocol. The extent of NHEJ repair in the locus amplicons reflects the frequency of site-specific genomic cleavage.

1. Extract genomic DNA from ZFN-treated cells and dissolve in nuclease-free water. Quantify DNA concentration by A_{260} spectrophotometry.
2. Amplify the AAVS1 locus region from 200 ng genomic DNA with the PCR primer pairs in Table 1 and a high-fidelity DNA polymerase.
3. Confirm specificity of amplification by visualizing a single 469 bp fragment on DNA gel electrophoresis.
4. Denature approximately 200 ng of gel-purified AAVS1 locus amplicon at 95 °C for 5 min and reanneal by gradually cooling to room temperature. Set up DNA digestion with the endonuclease according to the supplier's protocol and resolve the reaction products by 10% polyacrylamide gel electrophoresis. Load a DNA ladder in an adjacent well as size markers.
5. Post-stain the gel with a DNA dye for 30 min at room temperature protected from light in an orbital shaker (50 rpm). Visualize and capture the image of DNA bands on a UV transilluminator. Quantify band intensities using densitometric analytical software (e.g., Quantity One®, Bio-Rad; ImageJ, imagej.nih.gov). The unmodified AAVS1 genomic DNA locus is 469 bp; site-specific genome cleavage and repair and mismatch cleavage generate 287 and 182 bp fragments (*see* **Note 11**).

3.5.3 Deep Sequencing of Top 10 Predicted Off-Target Sites

One approach to assessing off-target activity is by deep sequencing the most likely in silico-predicted off-targets for indels [11] (*see* **Note 12**).

1. Amplify the AAVS1 locus and the top 10 predicted off-target sites (OT1–OT10) from genomic DNA extracted from unmodified and AAVS1 ZFN-treated cells using a high-fidelity DNA polymerase and PCR primers listed in Table 1.
2. Resolve and gel purify each amplicon by agarose gel electrophoresis. Quantify DNA concentration by A_{260} spectrophotometry.
3. Spike a commercially synthesized DNA fragment identical to the AAVS1 locus sequence except for a 5 bp deletion between the ZFN-binding half-sites into the wild-type AAVS1 locus amplicon at mass ratios of 1:10, 1:100, 1:500, and 1:1000 to determine the sensitivity of indel detection by deep sequencing (*see* **Note 13**).
4. Subject equal amounts of PCR amplicons to targeted next-generation paired-end sequencing. This is best done in a specialized sequencing facility where library construction is performed

according to the manufacturer's protocol. Mean depth of sequencing should be ≥500× to detect low-frequency indels.

5. Indels that occur within or 10 bp on either side of a predicted off-target site are scored as off-target events. The normalized frequency of indels is the indel frequency divided by the total number of aligned reads at that position.

3.6 Assessing Site-Specific Integration of Donor Vector

1. Electroporate cells with 5 μg ZFN expression plasmid and 20 μg donor DNA plasmid.
2. Commence selection for puromycin resistance 4 days after electroporation. Stably integrated cells are selected after 7 days in 0.5 μg/mL puromycin (*see* **Note 14**).
3. Extract genomic DNA from approximately 2×10^6 stably integrated cells (hereafter designated PURO cells).
4. Determine the presence of site-specific integration by integration junction PCR (*see* **Note 15**).
5. Resolve and purify the amplified products on agarose gel electrophoresis. Sequence each amplicon to confirm the identity of right and left integration junctions.

3.6.1 Identifying Integration of Donor Vector at AAVS1 Locus

3.6.2 Quantify Copy Number of Site-Specific Integrations by Droplet Digital PCR

The number of AAVS1 site-specific integrations versus the total number of integrations anywhere in the genome quantified by droplet digital PCR (ddPCR) gives an objective measure of the specificity of ZFN cleavage. Set up at least five different reactions: one for each integration junction, one to quantify total copy number of the integrated transgene (both site specific and random), and two reactions to quantify genomic loci neighboring but outside the target site to serve as the control copy number of unmodified genome (*see* **Note 16**).

1. Mix 50 ng genomic DNA from stably integrated cells (PURO cells), 1 μM of each PCR primer, 0.25 μM BHQ1-FAM hydrolysis probe, and ddPCR supermix in a final reaction volume of 20 μL (*see* **Note 17**).
2. Set up an identical reaction mixture with 50 ng genomic DNA from unmodified control cells (WT cells).
3. Transfer the mixture to a droplet generator to generate 20,000 droplets.
4. Perform 40 amplification cycles (annealing temperature 61 °C and extension time 1 min per cycle) using a thermal cycler.
5. Score droplets for fluorescent PCR products on droplet reader and analyze data using the system's software.
6. Primers and probes for two control loci in 19q13.42 close to but outside the target site in *PPP1R12C* intron 1:

Control locus 1	Forward primer: cctgccttaaacccagccag
	Reverse primer: atgacctcatgctcttggccctcgta
	Probe: aaccaccccagcagatactct
Control locus 2	Forward primer: tcccctcccagaaagacctgc
	Reverse primer: tcccctcccagaaagacctgc
	Probe: tacctaacgcactcctgggtga

Each locus is quantified in quadruplicate reactions.

3.6.3 RT-qPCR for Effects of Stable Transgene Integration on PPP1R12C and Neighboring Genes

1. Extract and purify total RNA from approximately 2×10^6 stably integrated cells and 2×10^6 control cells (electroporated without ZFN and donor DNA, and of the same number of population doublings). Quantify the amounts of RNA by A_{260} spectrophotometry.
2. Perform first-strand cDNA synthesis on 500 ng total RNA using a cDNA synthesis kit according to the manufacturer's protocol.
3. Perform reverse transcriptase-quantitative PCR (RT-qPCR) on equal amounts of 1:10 diluted cDNA samples and gene-specific primers (Table 1) using a real-time detection system to determine changes in expression levels of *PPP1R12C* and neighboring genes within a 1 Mb interval.
4. Design intron-spanning exonic primers to amplify transcripts of *PPP1R12C* transcript (exons 4–6) (*see* **Note 18**), *LILRB4*, *ISOC2*, *PPP6R1*, *NAT14*, *ZNF579*, *FIZ1*, *RDH13*, and *GAPDH* (the latter is a housekeeping gene for data normalization). Potential interacting partners of *PPP1R12C* predicted by Gene Network Central™ and Spring 9.05 may also be assayed by RT-qPCR.
5. Calculate *GAPDH*-normalized transcript levels and the fold change in stably integrated cells relative to control cells using the $\Delta\Delta C_T$ method [12] (*see* **Note 19**).

3.7 RNA-Seq to Evaluate ZFN-Induced Whole-Transcriptome Changes

1. Extract and purify total RNA from approximately 2×10^6 stably integrated and 2×10^6 control cells. Quantify the amount of RNA using a fluorimeter and assess RNA integrity using Bioanalyzer (*see* **Note 20**).
2. Use 2 μg RNA (RIN > 9) for library preparation following the manufacturer's standard protocol (e.g., Illumina TruSeq® RNA Sample Prep kit; CA, USA).
3. Purify poly-A mRNA on oligo-dT magnetic beads, fragment mRNA (150–250 bp), and convert into first-strand cDNA using reverse transcriptase and random primers.

4. Synthesize second-strand cDNA synthesis with DNA polymerase I and RNaseH.
5. Blunt-end cDNA fragments and add a single A nucleotide for ligation to indexed adapters with complementary T-overhangs.
6. Purify the indexed products and enrich with PCR to create the final cDNA library.
7. Check indexed libraries for size and purity by Bioanalyzer, and quantify dsDNA using a sensitive fluorescence assay.
8. Normalize libraries to 10 nM by real-time PCR and pool equal volumes.
9. Denature pooled libraries and dilute to 20 pM for clustering on the cBot before loading on HiSeq 2000 (Illumina) to generate paired-end reads of 76 bp.
10. Map sequence reads to the reference human genome using TopHat (http://ccb.jhu.edu/software/tophat/index.shtml).
11. Calculate differential expression using Cufflinks (http://cole-trapnell-lab.github.io/cufflinks/) (*see* **Note 21**).
12. Annotate significantly altered transcripts and pathways using DAVID (Database for Annotation, Visualization and Integrated Discovery) 2.1 Functional Annotation Tool (http://david.abcc.ncifcrf.gov) [13].
13. Cross-reference altered transcripts to an aggregated compilation of oncogenes and tumor-suppressor genes (http://www.bushmanlab.org/links/genelists).
14. Identify predicted and known interacting partners of PPP1R12C from Gene Network Central™ (http://gncpro.sabiosciences.com/gncpro/gncpro.php) and String 9.0 (http://string-db.org/) (*see* **Note 22**).

3.8 Whole-Genome Sequencing

1. This is best performed in a specialized core facility.
2. Prepare libraries with 500 bp inserts from 5 μg randomly fragmented genomic DNA from control and stably integrated cells for sequencing on HiSeq 2000 (Illumina) (*see* **Note 23**).
3. Aim for ≥60× sequencing depth.
4. Remove adapter sequences, duplicate reads (marked using Picard tools; https://broadinstitute.github.io/picard/), low-quality reads (more than half the bases in a read having a quality ≤5), and reads with >10% unknown bases from the raw data.
5. Align paired-end clean reads (90 bases) from both samples to the reference human genome using Burrows-Wheeler Aligner (http://bio-bwa.sourceforge.net/) [14] and store as BAM format files.

3.8.1 Bioinformatic Data Analyses

Detection of On-Target Indels

1. BLAST all reads of the stably integrated sample (PURO) against a 1 kb sequence containing the intended ZFN cleavage site with e-value cutoff 0.00001. Using a stringent cutoff minimizes false alignments.
2. Extract the reads which align to a 200 bp window centered on the ZFN cleavage site within of the 1 kb sequence.
3. Create an index of the 1 kb sequence with the command:
   smalt index -k 11 -s 1 <INDEX_NAME> <1kb-seq.fasta>.
4. Map the extracted reads to the reference human genome using SMALT (https://www.sanger.ac.uk/science/tools/smalt-0) to generate a sequence alignment file. Command line used:
   smalt map -f samsoft <INDEX_NAME>_R1_R2 | samtools view -bS - > map.bam (*see* **Note 24**).
5. Create a pileup file with samtools mpileup:
   samtools mpileup -f 1kb-seq,fasta map.bam > PILEUP
6. Use VarScan v2.3.6 (http://dkoboldt.github.io/varscan/) to detect indels using command "pileup2indel" [15]:
   java -jar VarScan.v2.3.6.jar pileup2indel PILEUP > INDELS

Detection of Off-Target Indels

1. Map WT (from control cells) and PURO (from stably integrated cells) reads to the reference human genome using BWA which results in two BAM files. This is done by BWA and SAMtools:
   bwa mem <index> <R1> <R2> | samtools view -bS - > BAM.
2. Create pileup files from the BAM files produced in **step 1** using SAMtools (http://www.htslib.org/) [16]:
   samtools mpileup -f <Ref-genome.fasta> BAM > PILEUP.
3. Submit the resulting pileup files for somatic variant calling by VarScan using the command "somatic" taking control data (WT) as "normal" and stably integrated data (PURO) as "tumor" (java –jar VarScan.jar somatic <pileup file of WT sample> <pileup file of PURO sample> output).
4. Classify somatic variants identified by VarScan as high-confidence (.hc) or low-confidence (.lc) using the command "processSomatic" (java –jar VarScan.jar processSomatic output.indel).
5. Further analyze high-confidence somatic variants to generate the final list of candidate indels for experimental validation. This is done as follows:
   (a) Create a BED file for the high-confidence indels in the form shown below. Create with Perl script:
   CHR POS1 POS2 INDEL-POS
   where
   CHR = chromosome

INDEL-POS = position of the indel
POS1 = INDEL_POS – 50
POS2 = INDEL_POS + 50

(b) Get sequences from the reference human genome (hg19) in FASTA format based on the BED file created in **step 5(a)**.
Program used is BEDTools:
bedtools getfasta -fi <FASTA-IN> -bed BED.file -fo <FASTA-OUT>.

(c) Check the fasta file for LOW_COMPLEXITY (repeats). Program used: dustmasker (from BLAST v.2.6.0).

(d) Parse the output of dustmasker with Perl script. The output looks like the example below:
chr1 coor1
chr1 coor2
etc.,
one line per INDEL.

(e) Extract sequences which are not in LOW-COMPLEXITY regions:
grep -A 1 -f <COOR-FILE> <FILE-TO-EXTRACT-FROM.fasta>
FILE-TO-EXTRACT-FROM.fasta is the file from Subheading "Detection of Off-Target Indels", **step 5(b)**) COOR-FILE is the file resulting from Subheading "Detection of Off-Target Indels", **step 5(d)**).

(f) Identify putative off-target sites in the reference human genome in silico using ZFN-Site (http://ccg.vital-it.ch/tagger/targetsearch.html) [17]. Submit sequences of the right (5′ ACCCCACAGTGG 3′) and left (5′ ATCCTGTCCCTA 3′) ZFN half-sites. Allow a spacer of 5 or 6 bp between half-sites and set the maximum number of mismatches per half-site to 2.

(g) We next adopted another method to identify possible off-target effects by scanning the reads for hotspot genomic regions in which multiple overlapping high-confidence somatic indel variant map. This can be done using a simple Perl script that reads positions of indels.

(h) Validate the final list of high-confidence somatic variants by PCR-Sanger sequencing.

Identification of AAVS1 ZFN-Mediated Transgene Insertion

1. Generate a 1 kb genome sequence centered on the AAVS1 target site and digitally insert the donor DNA sequence at the cleavage site (intron 1, *PPP1R12C*).
2. Align all sequence reads from stably integrated cells to the modified sequence from Subheading "Identification of AAVS1

ZFN-Mediated Transgene Insertion," **step 1**, and to the reference human genome.

3. This identifies donor DNA vector-specific sequences that are unique to the integrated transgene.
4. The absence of donor DNA vector-specific sequences in genomic DNA of control cells and their presence only in genomic DNA of stably integrated cells confirms transgene integration.
5. Identify off-target integrations by reads from stably integrated DNA that are combinations of human and vector sequences (hereafter called mixed reads) (*see* **Note 25**).
6. Map all PURO reads (from stably integrated DNA) to AAVS1 using BWA.

   bwa mem <AAVS1> R1 R2 | samtools view -bS -F 4 - > mapped-reads.bam

   -F 4 stores only reads that map to AAVS1 into the bam file. This saves space and speeds up further processing.
7. Extract the reads that map to AAVS1.
8. Map the extracted reads to hg19.
9. Only reads that map to AAVS1 are considered to be on-target ZFN-mediated insertions.

Detection of Structural Variants

1. Pre-process raw reads with SAMtools and Picard (https://broadinstitute.github.io/picard/) as SVDetect [18] and BreakDancer [19] require only abnormally mapped reads.
2. The command line for SAMtools is:

   samtools view -b -h -F 10 -q 22 input.bam > output.bam. This command retains in output.bam only reads which are abnormally mapped (mapped to different chromosomes and on different strands) and have a minimum mapping quality of 22.
3. Discard duplicate reads in the resulting bam file using the following command in Picard tools:

   java -jar /opt/picard-1.111/MarkDuplicates.jar INPUT=output.bam OUTPUT=output.nodup.bam REMOVE_DUPLICATES=true ASSUME_SORTED=true METRICS_FILE=metrics.output.txt.
4. Submit the resulting bam files to SVDetect (version r0.08; http://svdetect.sourceforge.net) and BreakDancer (version 1.4.4; http://breakdancer.sourceforge.net/) for calling structural variants (SV).
5. Configure SVDetect and BreakDancer to detect rearrangements with two or more supporting read pairs using eight times the standard deviation as threshold.

6. After the first step of SVDetect, filter the resulting ".links" file containing all the called SVs for "imperfect duplicates" [20] with an in-house Perl script. Remove links supported solely by clusters of imperfect duplicates. Retain links which have only some imperfect duplicates after removing the supporting imperfect duplicates.
7. Filter the files resulting from **step 6** using SVDetect's own filtering process.
8. Compare the filtered SVs called for both control and stably integrated cells. Turn off the option for comparing only the same SV type.
9. It is not necessary to filter and remove imperfect duplicates for BreakDancer if the anchoring region is set to 3 and the default value in BreakDancer is 7.
10. Subject results of BreakDancer and SVDetect to another filter to identify overlaps with repetitive DNA and low-mappability regions [20]. Remove SVs that are only supported by reads which overlap in any of these regions.
11. Filter using an in-house Perl script and BEDtools (version 2.17.0; http://bedtools.readthedocs.org/en/latest/) [21]. BED files of these regions needed for filtering are extracted as described [20] from the UCSC genome browser (http://genome.ucsc.edu/).
12. In the final step, inspect outputs from BreakDancer and SVDetect for genomic loci that have clusters of reads consistent with inter- or intrachromosomal SVs specific to stably integrated cells. This is done with an in-house Perl script and different degrees of freedom (DF) of clustering. Set the maximum DF to 3; that is, genomic coordinates that differ by up to 999 bases are still considered as evidence of clustering.

3.9 Validation of Indels and Structural Variants Identified by Whole-Genome Sequencing

1. Amplify genomic DNAs from control and stably integrated cells with phi29 polymerase.
2. Validate high-confidence indels by PCR-Sanger sequencing.
3. Validate unbalanced structural variants identified, by SVDetect and BreakDancer, by quantitative PCR of genomic breakpoint regions with β-actin amplification as internal control in each reaction.
4. Run PCR on a real-time detection system for 50 amplification cycles in triplicate reactions. Confirm product specificity with melt curves and determine the threshold cycle (C_T).
5. Normalize the mean C_T value of each test locus to its own actin C_T value.

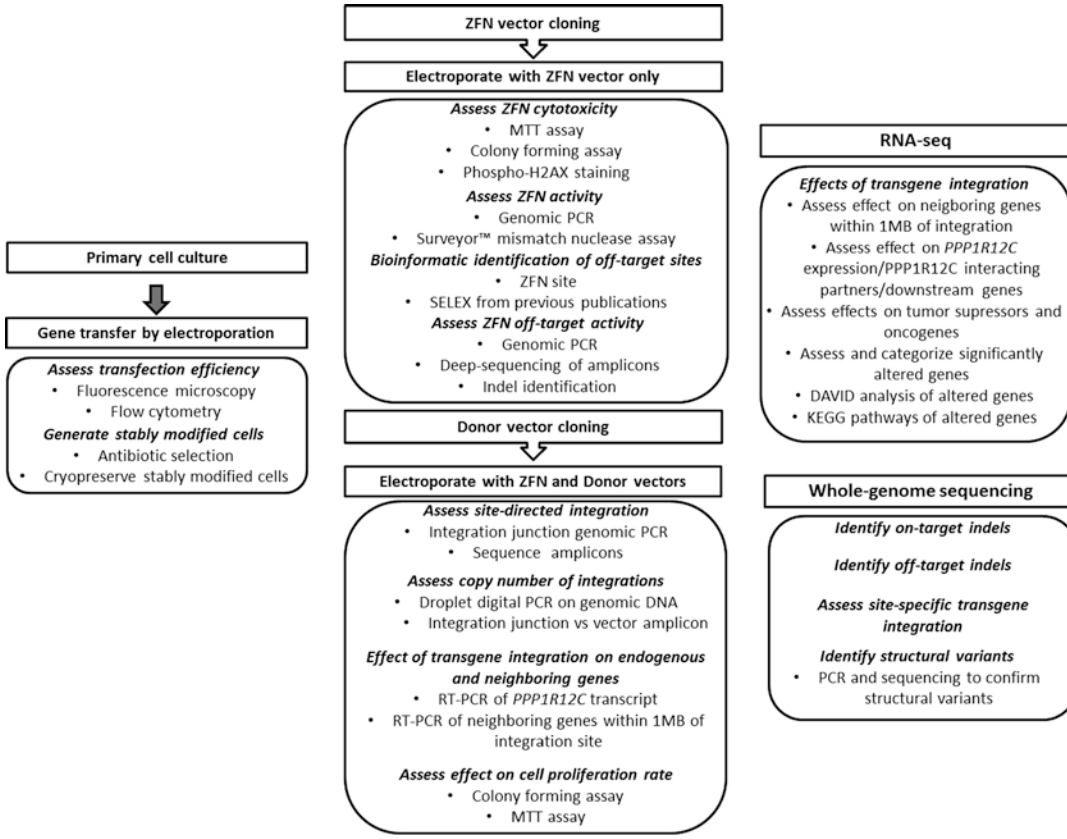


Fig. 2 Schematic of the overall multimodal workflow for assessment of genotoxicity in primary human cells modified with ZFN-mediated AAVS1 site-directed transgene integration

6. Express results as the ratio of normalized C_T value of stably integrated genomic DNA to normalized C_T value of control genomic DNA at each putative breakpoint locus (*see* **Note 26**).
7. Confirm structural variants by PCR-Sanger sequencing.

Fig. 2 is an overview of the multimodal workflow described in this chapter.

4 Notes

1. Cord-lining epithelial cells express pluripotency genes [22]. Other primary human cell types such as fibroblasts and bone marrow- and adipose tissue-derived stromal cells may also be modified by AAVS1 ZFN [23]. Cell culture and electroporation conditions must be optimized for different cell types.
2. Obligate heterodimeric (OH) ZFN has two amino acid changes in the *Fok*I monomer fused to the right AAVS1 homol-

ogy arm (E490K and I538K) and the monomer fused to the left homology arm (Q468E and I499L).

3. **Steps 1** and **2**, when completed, introduce the following amino substitutions: right *Fok*I monomer (S418P, K441E, E490K, H537R, and I538K) and left *Fok*I monomer (S418P, K441E, Q468E, N496D, and I499L).
4. The human AAVS1 locus is within intron 1 of *PPP1R12C* on 19q13.42. AAVS1 ZFN can be purchased from Sigma Aldrich or designed by investigators. If the latter, consider codon optimization and consult the references 24–27 for detailed protocols for designing, constructing, and testing ZFNs.
5. Using a promoterless puromycin resistance gene to select antibiotic-resistant cells from the *PPP1R12C* endogenous promoter significantly improves selection of cells with AAVS1 site-directed integration. In contrast, integration of donor vector with an exogenous promoter-driven antibiotic resistance gene results in a high frequency of random integrations.
6. Cells should be in a proliferative state and preferably 50–70% confluent at the time of electroporation.
7. Remove as much supernatant as possible before resuspending in Nucleofection solution. Residual buffer or media can negatively affect Nucleofection.
8. Optimized Nucleofector solutions and programs for nucleofection of different cell types can be found here (http://bio.lonza.com/resources/product-instructions/protocols/). Use Nucleofector solution P1 for CLECs.
9. For cell viability and cell proliferation assays, include non-electroporated cells and cells electroporated with donor DNA only as controls to compare against cells electroporated with both donor DNA and ZFN.
10. Phosphorylated histone H2AX is an established marker for double-stranded DNA breaks and serves as an indicator of genotoxicity. Conditions which result in low H2AX staining should be favored.
11. The percentage or ratio of cleaved bands to the uncleaved PCR amplicon gives an estimate of the proportion of cells modified by NHEJ repair and thus, by inference, the efficiency of ZFN site-specific genomic cleavage.
12. ZFN-Site (http://ccg.vital-it.ch/tagger/targetsearch.html) and PROGNOS (http://bao.rice.edu/Research/BioinformaticTools/prognos.html) are programs that predict potential ZFN off-target sites.
13. It is important to determine the detection sensitivity of deep sequencing by spiking known indel controls. The lower detection limit should be identified and used as the threshold to

positively identify indels present only in ZFN-modified cells and absent in unmodified cells.

14. These selection conditions for puromycin resistance apply to cord-lining epithelial cells. Conditions should be optimized for other types of primary cells.
15. Integration junction PCR is performed with one primer specific for the integrated transgene and the second primer specific for the AAVS1 locus outside the homology arm. Junctional amplification of antibiotic selected cells should yield a clearly positive band at each junction. A weak signal may indicate poor enrichment of cells with site-specific transgene integration. It is also important to optimize the primers and PCR reaction conditions (annealing temperature, GC melt reagents, robust DNA polymerase) to improve integration junction PCR.
16. Our protocol describes droplet digital PCR performed on the Bio-Rad system. It is essential to read Bio-Rad's user manual carefully before embarking on experiments. Similarly, read the manufacturer's user manual if using another ddPCR system. Droplet digital PCR will quantify comparable concentrations of donor transgene/vector, and left and right integration junction amplicons if the majority of integrations are on-target. Significantly different copy numbers (donor transgene/vector concentration >> integration junction concentrations) indicate significant off-target integrations. *See* reference 23 for an example of ddPCR data for this application.
17. Hydrolysis probes for ddPCR can be custom synthesized by several companies, e.g., Sigma-Aldrich, IDT, Thermo Fisher.
18. It is important to measure the expression of endogenous genes around the site of transgene integration to determine what effect, if any, ZFN-mediated integration has on neighboring gene expression. Monoallelic integration into intron 1 of *PPP1R12C* will be reflected by approximately 50% reduction in *PPP1R12C* transcripts. *PPP1R12C* disruption is well tolerated by cells.
19. If there are significant transcriptional changes of neighboring genes located close to the ZFN integration site, functional assays should be performed to evaluate potentially oncogenic alterations of cellular phenotype.
20. If RNA-seq is outsourced to a specialized core facility, it is important to be clear which steps are performed in the investigator's laboratory and which by the core facility. It is also important to understand the quantity and quality of RNA for RNA-seq, and how the data will be returned.
21. Although a greater than twofold difference in transcripts levels between two groups is commonly used to designate significantly altered expression, this is an arbitrary threshold. Investigators decide on the threshold to use for their data. The

unpaired t-test or Mann-Whitney test is used for Gaussian or non-Gaussian transcriptome data, respectively. Significance *p*-values must be corrected for multiple testing. A false discovery rate of 0.05 is commonly accepted.

22. Genes with significantly altered transcription should be evaluated for likely adverse functional consequences by mapping to gene networks and pathways. It is especially important to look for oncogene activation, tumor suppressor silencing, and dysregulation of oncogenic signaling pathways.
23. Whole-genome sequencing is performed on bulk populations of cells as it is not possible to develop single clonal populations from untransformed primary human cells. It is helpful to work with bioinformatics specialists in analyzing whole-genome sequence data.
24. Reads were extracted with a Perl script. This can be done alternatively using a program called seqtk (https://github.com/lh3/seqtk); the command to use is:

    seqtk subseq <R1.fastq> <NAMELIST.file> > NEW_R1.fastq
    seqtk subseq <R2.fastq> <NAMELIST.file> > NEW_R2.fastq
25. Mixed reads are reads which map to both the human genome and the digitally created insert (Subheading "Identification of AAVS1 ZFN-Mediated Transgene Insertion," **step 1**).
26. The ratio of genome copy number at a putative chromosomal translocation breakpoint (ZFN-modified genome: control unmodified genome) should be significantly different from 1 in a bona fide translocation because translocations are almost always unbalanced. Ratios which are not significantly different from unity indicate absence of unbalanced translocations or other chromosomal rearrangements. *See* reference 23 for an example of this analysis.

Acknowledgments

Supported by the National Medical Research Council (grant CIRG/1326/2012), National Cancer Centre and CellResearch Corporation (all in Singapore). D.K. and S.M.S. were supported by the Agency for Science, Technology and Research (A*STAR), Singapore.

References

1. Collins M, Thrasher A (2015) Gene therapy: progress and predictions. Proc Biol Sci 282:20143003
2. Kaufmann KB, Büning H, Galy A et al (2013) Gene therapy on the move. EMBO Mol Med 5:1642–1661
3. Papapetrou EP, Schambach A (2016) Gene insertion into genomic safe harbors for human gene therapy. Mol Ther 24:678–684
4. Carroll D (2014) Genome engineering with targetable nucleases. Annu Rev Biochem 83:409–439

5. Miller JC, Holmes MC, Wang J et al (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25:778–785
6. Doyon Y, Vo TD, Mendel MC et al (2011) Enhancing zinc-finger-nuclease activity with improved obligate heterodimeric architectures. Nat Methods 8:74–79
7. Szczepek M, Brondani V, Buchel J et al (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. Nat Biotechnol 25:786–793
8. Guo J, Gaj T, Barbas CF 3rd (2010) Directed evolution of an enhanced and highly efficient FokI cleavage domain for zinc finger nucleases. J Mol Biol 400:96–107
9. Hendel A, Fine EJ, Bao G et al (2015) Quantifying on- and off-target genome editing. Trends Biotechnol 33:132–140
10. Zischewski J, Fischer R, Bortesi L (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. Biotechnol Adv 35:95–104
11. Hockemeyer D, Soldner F, Beard C et al (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27:851–857
12. Livak KJ, Schmittgen TD (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2(-\Delta\Delta C_T)$ method. Methods 25:402–408
13. Huang da W, Sherman BT, Lempicki RA (2009) Bioinformatics enrichment tools: paths toward the comprehensive functional analysis of large gene lists. Nucleic Acids Res 37:1–13
14. Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows-Wheeler transform. Bioinformatics 26:589–595
15. Koboldt DC, Zhang Q, Larson DE et al (2012) VarScan 2: somatic mutation and copy number alteration discovery in cancer by exome sequencing. Genome Res 22:568–576
16. Li H, Handsaker B, Wysoker A et al (2009) The sequence alignment/map format and SAMtools. Bioinformatics 25:2078–2079
17. Cradick TJ, Ambrosini G, Iseli C et al (2011) ZFN-site searches genomes for zinc finger nuclease target sites and off-target sites. BMC Bioinformatics 12:152
18. Zeitouni B, Boeva V, Janoueix-Lerosey I et al (2010) SVDetect: a tool to identify genomic structural variations from paired-end and mate-pair sequencing data. Bioinformatics 26:1895–1896
19. Chen K, Wallis JW, McLellan MD et al (2009) BreakDancer: an algorithm for high-resolution mapping of genomic structural variation. Nat Methods 6:677–681
20. Mijuskovic M, Brown SM, Tang Z et al (2012) A streamlined method for detecting structural variants in cancer genomes by short read paired-end sequencing. PLoS One 7:e48314
21. Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842
22. Sivalingam J, Krishnan S, Ng WH et al (2010) Biosafety assessment of site-directed transgene integration in human umbilical cord-lining cells. Mol Ther 18:1346–1356
23. Sivalingam J, Kenanov D, Han H et al (2016) Multidimensional genome-wide analyses show accurate FVIII integration by ZFN in primary human cells. Mol Ther 24:607–619
24. Carroll D, Morton JJ, Beumer KJ et al (2006) Design, construction and *in vitro* testing of zinc finger nucleases. Nat Protoc 1: 1329–1341
25. Porteus M (2008) Design and testing of zinc finger nucleases for use in mammalian cells. Methods Mol Biol 435:47–61
26. Sander JD, Maeder ML, Joung JK (2011) Engineering designer nucleases with customized cleavage specificities. Curr Protoc Mol Biol Chapter 12:Unit12.13 1–19
27. Fujii W, Kano K, Sugiura K et al (2013) Repeatable construction method for engineered zinc finger nuclease based on overlap extension PCR and TA-cloning. PLoS One 8:e59801

Chapter 12

Selection and Characterization of DNA Aptamers Against FokI Nuclease Domain

Maui Nishio, Ayana Yamagishi, Kaori Tsukakoshi, Yoshio Kato, Chikashi Nakamura, and Kazunori Ikebukuro

Abstract

Genome editing with site-specific nucleases (SSNs) may be effective for gene therapy, as SSNs can modify target genes. However, the main limitation of genome editing for clinical use is off-target effects by excess amounts of SSNs within cells. Therefore, a controlled delivery system for SSNs is necessary. Previously we have reported on a zinc finger nuclease (ZFN) delivery system, which combined DNA aptamers against FokI nuclease domain (FokI) and nanoneedles. Here, we describe how DNA aptamers against FokI were selected and characterized for genome editing applications.

Key words Aptamer, FokI nuclease, Genome editing, Nanoneedles

1 Introduction

Genome editing with site-specific nucleases (SSNs) can modify only the target genes and may be applicable for gene therapy [1–3]. However, the main limitation to the clinical application of genome editing in humans is off-target effects, which lead to double-strand breaks at non-target sites. Researchers have reported that off-target effects are common in genome editing by SSNs because of the uncontrollable duration and amount of SSN expression in the cells due to transfection of the plasmids [3–5]. Therefore, to apply genome editing in clinical use, a controlled system for delivering SSNs is necessary to reduce off-target effects.

Aptamers, the nucleic acid ligands, which specifically bind to the target molecule with high affinity are obtained via selection by systematic evolution of ligands by exponential enrichment (SELEX) [6, 7]. We have developed various aptamer-based sensing systems by taking advantage of their conformational changes in the presence or absence of the target molecule [8–10].

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_12,

We have previously reported that a nanoneedle, 2 nm in diameter, could be inserted into living cells with high efficiency and minimal damage [11]. We have since developed a nanoneedle array containing several tens of thousands of nanoneedles on a 5 mm square silicon chip [12] with an aptamer-based oligonucleotide for protein delivery [13]. Therefore, a combination of aptamer-based oligonucleotides and nanoneedles may achieve the controlled delivery of SSNs to reduce off-target effects in genome editing.

To apply the aptamer to this controlled delivery system, the required function of aptamers is to bind SSNs with high affinity, but release SSNs when introduced into cells. In addition, the aptamers should not be cleaved by SSNs itself. FokI nuclease domain (FokI) is the common cleavage domain of typical SSNs, zinc finger nucleases (ZFN), and transcription activator-like effector nucleases (TALEN). Therefore, we developed the aptamers against FokI [14] to apply them to protein delivery system of SSNs. Here, we describe how we selected and characterized the aptamers against FokI.

2 Materials

Prepare all solutions using distilled water or ultrapure water (prepared by purifying deionized water, to attain a sensitivity of 18 MΩ-cm at 25 °C). Unless indicated otherwise, prepare and store all reagents at room temperature. Diligently follow all waste disposal regulations when disposing waste materials.

Oligonucleotides were diluted using each buffer. The oligonucleotides were then heated at 95 °C for 10 min and gradually cooled to 25 °C for 30 min prior to performing all experiments, including SELEX.

2.1 Selection of DNA Aptamers by SELEX

1. DNA library and primers (*see* **Note 1**).
2. Nitrocellulose membranes.
3. Tris-buffered saline (TBS): 10 mM Tris–HCl, 150 mM NaCl, 1 mM EDTA, pH 7.4 (*see* **Note 2**). Adjust pH using NaOH.
4. TBS-T: TBS supplemented with 0.05% (v/v) Tween20.
5. Blocking solution: 2% (w/v) Bovine serum albumin in TBS-T.
6. Phenol:chloroform:isoamyl alcohol (PCI): 25:24:1, v/v.
7. Chloroform:isoamyl alcohol (CIA): 24:1, v/v.
8. 3 M Sodium acetate, pH 5.2.
9. 100% Ethanol: Store at −20 °C.
10. 1 mg/mL Glycogen.
11. 70% Ethanol: Store at 4 °C.

12. Tris-EDTA buffer (TE): 10 mM Tris–HCl, 1 mM EDTA, pH 8.0.
13. Avidin agarose resin.
14. Column buffer, filter sterilized: 30 mM HEPES, 500 mM NaCl, 5 mM EDTA, pH 7.0.
15. 0.15 M NaOH (*see* **Note 3**).
16. 2 M HCl.
17. pGEX-T®-Easy Vector (Promega).
18. Competent cell, DH5α.

2.2 Aptamer Blotting

1. 5′ Biotinylated oligonucleotides including aptamer sequence.
2. Nitrocellulose membranes.
3. TBS (*see* item in Subheading 2.1).
4. TBS-T (*see* item in Subheading 2.1).
5. Blocking solution (*see* item in Subheading 2.1).
6. NeutrAvidin-HRP.
7. Chemiluminescent substrate: Immobilon Western Chemiluminescent HRP Substrate (Merck Millipore).
8. Image analyzer with chemiluminescence detection.

2.3 Native Polyacrylamide Gel Electrophoresis (Native-PAGE)

1. 20% (w/v) Polyacrylamide gel: Use 30% acrylamide solution (acrylamide:Bis = 29:1) and TBE buffer (1×, as described below).
2. TBE buffer (10×): 890 mM Tris, 890 mM borate, 20 mM EDTA.
3. 6× Loading buffer (*see* **Note 4**).
4. Gel imager with fluorescence detection.

2.4 Aptamer-Antibody Sandwich Assay

1. StreptAvidin-coated white 96-well plate.
2. TBS (*see* item in Subheading 2.1).
3. TBS-T (*see* item in Subheading 2.1).
4. Blocking solution: 4%(w/v) skim milk in TBS-T.
5. Plate shaker.
6. Primary antibody solution: 1:5000 Anti-Histag antibody in TBS-T.
7. Secondary antibody solution: 1:10,000 or 1:50,000 anti-mouse IgG antibody conjugated with alkaline phosphatase (ALP).
8. Chemiluminescent substrate: CDP-Star, ready-to-use (Roche).
9. Plate reader with chemiluminescence detection.

2.5 Enzyme-Linked Oligonucleotide Assay (ELONA)

1. Maxisorp polystyrene 96-well plate.
2. TBS (*see* item in Subheading 2.1).
3. TBS-T (*see* item in Subheading 2.1).
4. Blocking solution (*see* item in Subheading 2.1).
5. Plate shaker (*see* item in Subheading 2.4).
6. NeutrAvidin-HRP (*see* item in Subheading 2.2).
7. Chemiluminescent substrate: BM Chemiluminescence substrate (POD) (Roche).
8. Plate reader with chemiluminescence detection (*see* item in Subheading 2.4).

2.6 In Vitro Digestion Using ZFN

1. 1% Agarose gel: Prepared with TAE buffer (1×, as described below).
2. TAE (50×): 2 M Tris, 1 M CH_3COOH, 50 mM EDTA.
3. 10× NEBuffer4 (New England Biolabs).
4. 10% (w/v) SDS: 1 g of SDS diluted with water and mess up to 10 mL.
5. 6× Loading buffer.
6. SYBR safe.

2.7 Modification of GFP-ZFN on Nanoneedle Array

1. Silicon nanoneedle arrays: Nanoneedles with a length of 20 μm and a diameter of 200 nm. 300 × 300 of nanoneedles are arrayed in 3 × 3 mm^2 area (10 μm pitch) on a 5 mm square chip [12].
2. 100% Ethanol.
3. 1% Hydrogen fluoride.
4. O_2 plasma asher (JPA300; J-SCISSNCE Lab Co. Ltd.).
5. Phosphate-buffered saline (PBS): 2.7 mM KCl, 1.5 mM KH_2PO_4, 138 mM NaCl, 8.1 mM Na_2HPO_4, pH 7.4.
6. Biotinylated BSA: 1 μM Solution in PBS.
7. Streptavidin: 1 μM Solution in PBS.
8. TBS (*see* item in Subheading 2.1).
9. GFP fusion ZFN protein solution: A recombinant protein [13].

2.8 Intracellular Release Assay of GFP-ZFN

1. HEK293 cells.
2. 27φ Glass base dish.
3. Collagen: 0.1% Solution in 0.1 M acetic acid.
4. DMEM: 10% Fetal bovine serum.
5. Lipofectamine 2000 (Thermo Fisher Scientific).

6. Plasmid vector for DsRed2-NES expression: Prepare as reported in the article [11].
7. Opti-MEM (Thermo Fisher Scientific).
8. Nanoneedle array manipulator.
9. Confocal laser scanning microscope.

3 Methods

All procedures were conducted at room temperature unless otherwise specified.

3.1 Selection of DNA Aptamers by SELEX

1. Preparation of the DNA library: Dilute the DNA library with TBS buffer to 1–5 μM. Heat the diluted DNA library at 95 °C for 10 min and gradually cool to 25 °C for 30 min. Dilute the treated DNA library with TBS-T to the working stock concentration.
2. Preparation of the nitrocellulose membrane for negative selection: Cut a nitrocellulose membrane and incubate in blocking solution for 1 h. After incubation, wash the membrane three times with TBS-T.
3. Preparation of the nitrocellulose membrane for positive selection: Cut a nitrocellulose membrane and spot with 20 pmol of FokI. After drying the spot, incubate the membrane in blocking solution for 1 h. After incubation, wash the membrane three times with TBS-T.
4. Negative selection: Incubate the membrane (prepared in **step 2**) in the folded DNA library (prepared in **step 1**) for 2 h. After incubation, remove the nitrocellulose membrane using the tweezers.
5. Positive selection: Incubate the nitrocellulose membrane spotted with FokI (prepared in **step 3**) in the DNA library overnight after negative selection. After incubation, wash the membrane with TBS-T three times. Cut out the FokI immobilized area and put inside a 1.5 mL tube with 200 μL water.
6. Elution of DNA library bound to FokI: Add 600 μL of PCI to the 1.5 mL tube. Vortex the sample and incubate for 10 min. After incubation, centrifuge the sample at 12,000 × *g* for 5 min. Recover the aqueous layer to a new 1.5 mL tube, and add an equal amount of CIA to the 1.5 mL tube. Vortex the sample and incubate for 10 min. After incubation, centrifuge the sample at 12,000 × *g* for 5 min and recover the aqueous layer to a new 1.5 mL tube.
7. Concentration and purification of eluted DNA: The following are added to each sample based on the starting volume of the

sample: one-tenth of the sample amount of 3 M sodium acetate, three times the sample amount of 100% ethanol, and 10 μL of 1 mg/mL glycogen. Mix the sample and incubate for 15 min at −80 °C. After incubation, centrifuge the sample at 15,000 × *g* for 30 min at 4 °C. Discard the supernatant and add 1 mL of cold 70% ethanol. Centrifuge at 15,000 × *g* for 5 min at 4 °C and discard the supernatant. Dry the pellet for 10–15 min. After drying, suspend the pellet using 15 μL of TE.

8. Single-strand preparation: Amplify the collected DNA library using a forward primer and a 5′ biotinylated reverse primer by PCR. Shake the PCR product with avidin agarose resin for 30 min to remove the biotinylated reverse strands. Spin down the resin and remove the supernatant. Add 750 μL of column buffer to the sample, spin down, and then remove the supernatant twice. Add 200 μL of 0.15 M NaOH and shake for 15 min. Spin down the resin and recover the supernatant to a new 1.5 mL tube. Repeat these two steps once. Adjust the pH to 7.0 using 2 M HCl. After adjusting the pH, the collected ssDNA is concentrated by isopropanol precipitation.
9. Using the collected DNA from **step 8**, perform next round of SELEX and repeat until the DNA binds to FokI that is concentrated.
10. Sequence analysis: Amplify selected oligonucleotides using primers without biotinylation. Perform cloning of the PCR product using pGEX-T®-Easy Vector and transform into DH5α. Analyze the DNA sequence inserted into the vector to identify aptamer-candidate sequence.

3.2 Aptamer Blotting

1. Dilute 5′ biotinylated aptamers with TBS. Heat the diluted DNA library at 95 °C for 10 min and gradually cool to 25 °C for 30 min. After the treatment, dilute treated oligonucleotides with TBS-T to 200 nM.
2. Spot 20 pmol of FokI to the nitrocellulose membrane and dry. Block the membrane by incubating with blocking solution for 1 h. Wash the membrane three times with TBS-T.
3. Incubate the membrane in diluted oligonucleotides for 1 h. Wash the membrane three times with TBS-T.
4. Incubate the membrane in 1:50,000 diluted NeutrAvidin-HRP with TBS-T for 30 min. Wash the membrane three times with TBS-T.
5. Add the substrate onto the membrane and allow to stand for 5 min. Detect the chemiluminescence by image analyzer.

3.3 Native Polyacrylamide Gel Electrophoresis (Native-PAGE)

1. 5′ Fluorescein-modified oligonucleotides (5′-fluorescein-TTT-aptamer sequence) are diluted with TBS, not including EDTA. Heat the diluted DNA library at 95 °C for 10 min and gradually cool to 25 °C for 30 min.
2. The following were mixed and incubated at 37 °C for 1 h: 4 μL of each oligonucleotide, 4 μL of FokI (f.c. 500 nM each), 1 μL of 10× NEBuffer 4, and 2 μL of water.
3. Mix the 2 μL of 6× loading buffer with the incubated sample. Load the samples into each well of the 20% (w/v) polyacrylamide gel.
4. Perform the electrophoresis in TBE buffer.
5. Detect the fluorescence using a gel imager.

3.4 Aptamer-Antibody Sandwich Assay

1. Fold the 5′ biotinylated oligonucleotides in TBS as described above.
2. Add 100 μL of 100 nM biotinylated oligonucleotides to each well of a 96-well streptavidin-coated plate. Shake the plate at 800 rpm and incubate for 30 min. The plate is shaken at 800 rpm in all steps of incubation. After incubation, wash the well three times with TBS-T.
3. Add 150 μL of blocking solution to each well and incubate for 1 h. Wash the well three times with TBS-T.
4. Add 100 μL of FokI (0, 50, 100 nM) to each well and incubate for 1 h. Wash the well three times with TBS-T.
5. Add 100 μL of primary antibody solution and incubate for 1 h. Wash the well three times with TBS-T.
6. Add 100 μL of secondary antibody solution and incubate for 1 h. Wash the well three times with TBS-T.
7. Add 100 μL of chemiluminescent substrate and detect the chemiluminescence by a plate reader.

3.5 Enzyme-Linked Oligonucleotide Assay (ELONA)

1. Add 100 μL of FokI diluted with TBS-T to a Maxisorp polystyrene 96-well plate and incubate for 3 h. After incubation, wash the well three times with TBS-T.
2. Add 150 μL of blocking solution to each well and incubate for 1 h. After incubation, wash the well three times with TBS-T.
3. Add 100 μL of folded 5′ biotinylated oligonucleotides (f.c. 0, 10, 50, 100, 250, 500, 1000, 2000 nM) to each well and incubate for 1 h. After incubation, wash the well three times with TBS-T.
4. Add 100 μL of 1:50,000 diluted NeutrAvidin-HRP with TBS-T to each well and incubate for 1 h. After incubation, wash the well three times with TBS-T.
5. Add 100 μL of chemiluminescent substrate to each well. Detect the chemiluminescence by a plate reader.

3.6 In Vitro Digestion Using ZFN

1. Dilute 5′ fluorescein-modified oligonucleotides with TBS, not including EDTA, and fold by heat treatment as described above.
2. Mix each treated oligonucleotide with ZFN_CCR5L in 1xNEBuffer4 and incubate for 1 h.
3. Add the substrate linear plasmid (f.c. 4 ng/μL) to the sample prepared in **step 2** and incubate at 37 °C for 1 h. After incubation, add SDS (f.c. 1%) to each sample to stop the cleavage reaction.
4. Perform electrophoresis using 1% (w/v) agarose gel. Stain the substrate plasmid with SYBR safe.

3.7 Modification of GFP-ZFN on Nanoneedle Array

1. Clean the silicon nanoneedle arrays with an O_2 plasma asher for 5 min at 200 W.
2. Rinse the nanoneedle arrays with ultrapure water and ethanol.
3. Immerse the nanoneedle arrays in 1% hydrogen fluorite solution for 1 min to remove the oxidized layer and rinse with ultrapure water.
4. Drop 20 μL of 1 μM biotinylated BSA solution on the arrays and incubate for 1 h.
5. After rinsing the arrays with PBS, drop 20 μL of 1 μM streptavidin on the arrays and incubate for 1 h. Rinse the arrays with PBS.
6. Fold 5′ biotinylated aptamers (10 μM) by heat treatment as described above.
7. Drop 20 μL of 10 μM aptamers on the arrays and incubate for 1 h for immobilization of the aptamers to the streptavidin-modified array surface.
8. After rinsing with TBS, drop 20 μL of 100 nM GFP-ZFN solution on the arrays and incubate for 30 min to bind GFP-ZFN to the aptamer-modified array surface.
9. Rinse the arrays with TBS.

3.8 Intracellular Release Assay of GFP-ZFN

1. Transfect DsRed-NES2 expression vector into HEK293 cells by use of Lipofectamine 2000 according to the manual.
2. Seed HEK293 cells expressing DsRed2-NES (1×10^5 cells) onto a collagen-coated 27φ glass base dish and culture in DMEM with 10% fetal bovine serum. After incubation in a CO_2 incubator at 37 °C for 24 h, change the medium to Opti-MEM prior to insertion of the nanoneedles into the cell.
3. Attach the GFP-ZFN-bound nanoneedle array chip to a manipulator with an array holder with double-sided adhesive tape. Pierce the cells with the GFP-ZFN-bound nanoneedle array by lowering the chip to the cells using the manipulator.

Position the nanoneedle array approximately 200 μm above the cells and the Z piezo motor was used to approach the cells.

4. Scan images of cells inserted into the nanoneedle from the tip of the nanoneedles upwards by 25 μm at a scan rate of 1 μm/slice (total 26 slices) using a CLSM every 5 min for 30 min.
5. Analyze fluorescence intensity of the stack XY image of the nanoneedles from the tip to the 5 μm upper region using ImageJ. Calculate the average fluorescence intensity of the nanoneedles and estimate the amount of released GFP-ZFN.

4 Notes

1. We used a 66-nucleotide single-stranded DNA library harboring a 24-nucleotide randomized region and primer-binding regions. To introduce biotin to the reverse strand of the DNA library after PCR amplification, the reverse primer was modified with biotin at the 5′ end.
2. We added EDTA to our TBS in order to inactivate FokI. Therefore, during the aptamer selection, the DNA library was not cut by FokI.
3. Either prepare fresh NaOH solution every time or store 0.15 M NaOH at −20 °C.
4. To avoid disruption of aptamer folding, SDS should not be included in the loading buffer.

Acknowledgment

This work was supported by a Grant-in-Aid for Scientific Research (A) (JP26249127) from the Japan Society for the Promotion of Science (JSPS, Japan).

References

1. Carroll D (2011) Genome engineering with zinc-finger nucleases. Genetics 188:773–782
2. Perez EE, Wang J, Miller JC, Jouvenot Y, Kim KA, Liu O, Wang N, Lee G, Bartsevich VV, Lee YL, Guschin DY, Rupniewski I, Waite AJ, Carpenito C, Carroll RG, Orange JS, Urnov FD, Rebar EJ, Ando D, Gregory PD, Riley JL, Holmes MC, June CH (2008) Establishment of HIV-1 resistance in CD4+ T cells by genome editing using zinc-finger nucleases. Nat Biotechnol 26:808–816
3. Cox DB, Platt RJ, Zhang F (2015) Therapeutic genome editing: prospects and challenges. Nat Med 21:121–131
4. Wu X, Kriz AJ, Sharp PA (2014) Target specificity of the CRISPR-Cas9 system. Quant Biol 2:59–70
5. Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas CF 3rd (2014) Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One 9:e85755
6. Tuerk C, Gold L (1990) Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase. Science 249:505–510
7. Ellington AD, Szostak JW (1990) In vitro selection of RNA molecules that bind specific ligands. Nature 346:818–822

8. Yoshida W, Mochizuki E, Takase M, Hasegawa H, Morita Y, Yamazaki H, Sode K, Ikebukuro K (2009) Selection of DNA aptamers against insulin and construction of an aptameric enzyme subunit for insulin sensing. Biosens Bioelectron 24:1116–1120
9. Yoshida W, Sode K, Ikebukuro K (2008) Label-free homogeneous detection of immunoglobulin E by an aptameric enzyme subunit. Biotechnol Lett 30:421–425
10. Yoshida W, Sode K, Ikebukuro K (2006) Aptameric enzyme subunit for biosensing based on enzymatic activity measurement. Anal Chem 78:3296–3303
11. Obataya I, Nakamura C, Han S, Nakamura N, Miyake J (2005) Nanoscale operation of a living cell using an atomic force microscope with a nanoneedle. Nano Lett 5:27–30
12. Matsumoto D, Rao Sathuluri R, Kato Y, Silberberg YR, Kawamura R, Iwata F, Kobayashi T, Nakamura C (2015) Oscillating high-aspect-ratio monolithic silicon nanoneedle array enables efficient delivery of functional bio-macromolecules into living cells. Sci Rep 5:15325
13. Matsumoto D, Nishio M, Kato Y, Yoshida W, Abe K, Fukazawa K, Ishihara K, Iwata F, Ikebukuro K, Nakamura C (2016) ATP-mediated release of a DNA-binding protein from a silicon nanoneedle array. Electrochemistry 84:305–307
14. Nishio M, Matsumoto D, Kato Y, Abe K, Lee J, Tsukakoshi K, Yamagishi A, Nakamura C, Ikebukuro K (2017) DNA aptamers against FokI nuclease domain for genome editing applications. Biosens Bioelectron 15:26–31

Chapter 13

An Improved Genome Engineering Method Using Surrogate Reporter-Coupled Suicidal ZFNs

Jiani Xing, Cunfang Zhang, Kun Xu, Linyong Hu, Ling Wang, Tingting Zhang, Chonghua Ren, and Zhiying Zhang

Abstract

Using engineered nucleases such as zinc-finger nucleases (ZFNs) and TALE nuclease (TALEN) to accomplish genome editing often causes high cellular toxicity because of the consistent expression of artificial nucleases and off-targeting effect. And lacking selection marker in modified cells makes it hard to enrich these positive cells. Here we introduce a method by incorporating a surrogate reporter enrichment into a suicidal ZFN system, which is designed by a pair of ZFN expression cassettes flanked with its target sites. Our data demonstrated that this modified system achieved almost the same ZFN activity as the original method but reduced ~40% toxicity. This new suicidal ZFN expression system coupled with a surrogate reporter not only enables decreased cellular toxicity but also makes the genetic modified cells to be enriched by EGFP analysis.

Key words Zinc-finger nucleases, Cellular toxicity, Double-stranded breaks, Homologous recombination repair, Off-target

1 Introduction

Current genome engineering methodologies are all based on cellular repair of DNA double-stranded breaks (DSBs). When DSBs occur on a given chromosome, the breaks will be repaired spontaneously by nonhomologous end joining (NHEJ) when no donor DNA is available as a template. This repair mechanism will eventually cause the deletion or insertion of one or more base pairs, leading to the disruption of the open reading frame of targeted genes. On the other hand, in the presence of donor DNA, the DSBs can be repaired accurately by homologous recombination repair (HR) [1]. The inherent rate of HR is extremely low but can be dramatically increased when DSBs are introduced at the target site of interest. As DSBs can increase the efficiency of HR-mediated DNA repair, the development of a custom designer nuclease becomes critical in precision genome engineering.

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_13,

Engineered zinc-finger nucleases (ZFN) and TALE nucleases (TALEN) have been widely used as genome engineering tools in many species and cell types. ZFN-mediated gene targeting can reach an efficiency of up to 50% in human stem cells [2]. However, the inherent high cellular toxicity and off-target rates of these two systems remain to be the obstacles for the broader application of these tools. Several strategies have been developed to increase the specificity of ZFNs such as modification of the *Fok*I domain to minimize the toxicity of ZFNs [3]. Nevertheless, the constitutive expression of ZFNs typically leads to low cell viability. Another challenge is that the efficiencies of ZFN-mediated editing in pigs, cattle, and other animals are very low. A surrogate reporter system has been designed by Kim et al. [4] for selection and enrichment of gene-modified cells which may help to solve these problems. In this system, a ZFN target site is inserted between the coding region of mRFP and EGFP, which disrupts the open reading frame (ORF) of EGFP. Thus, the green fluorescence can only be detected when the cognate ZFNs cleave the target site in the reporter vector as well as in the genome. In addition, the green fluorescence can represent the cellular activity of ZFNs. This surrogate reporter system is based on NHEJ repair mechanism; thus only one-third of the modified cells can be enriched, which will underestimate the cellular activity of ZFNs. Moreover, constitutively expressed ZFNs may cleave off-target sites in the genome and cause high cytotoxicity.

Based on the observations above, we developed a new and safer method by designing a suicidal ZFN system that is coupled with a surrogate reporter to overcome the limitation of the previous method [4]. The new system contains a mRFP-EGFP fusion gene, a ZFN expression cassette, and two ZFN target sites between mRFP and EGFP coding sequence, the cleavage which can restore the ORF of EGFP via HR-mediated DNA repair (Fig. 1). The EGFP in our system can be efficiently targeted due to the long (3 kb) homologous arms. Importantly, the ZFN expression cassette will be deleted after ZFNs execute their activity, leading to

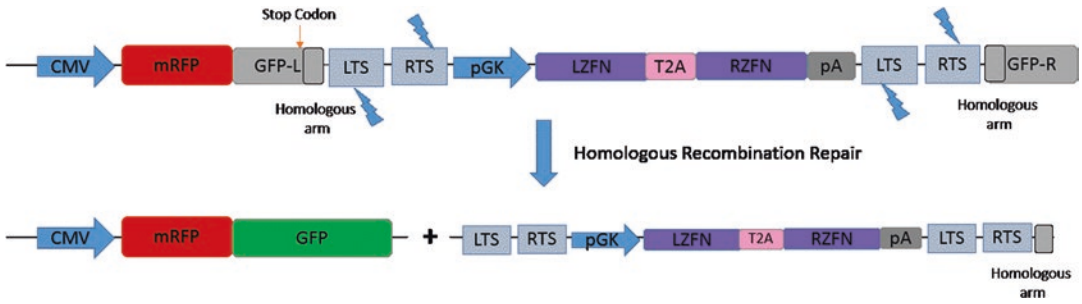


Fig. 1 Schematic diagram of the suicidal ZFN reporter system. *LTS* left target site of ZFN. *RTS* right target site of ZFN

reduced cytotoxicity. Moreover, other artificial endonucleases can also be used in our system.

2 Materials

2.1 Equipment

1. Nanodrop 2000 micro-spectrophotometer (*see* **Note 1**).
2. UV transilluminator.
3. Horizontal and vertical electrophoresis systems.
4. Fluorescence microscope.
5. Flow cytometer.
6. Cell culture incubator.

2.2 Construction of Suicidal ZFN Expression Vector with a Surrogate Reporter

1. 10 mg/mL Ethidium bromide (EB).
2. 1% Agarose gel supplemented with 740 μg/mL EB: Dissolve 100 g agarose in 100 mL deionized water, cool the agarose to 50 °C, and add EB to a final concentration of 740 μg/mL.
3. PCR reagents including dNTPs, DNA polymerase, and reaction buffers (*see* **Note 2**).
4. Restriction enzymes including *Bgl*II, *Sal*I, *Bam*HI, *Xba*I, *Not*I, and others with buffer supplied (*see* **Note 3**).
5. T4 DNA ligase with supplied buffer.
6. PCR primers (Table 1).
7. Gel extraction kit.
8. LB medium: Dissolve 10 g tryptone, 10 g NaCl, and 5 g yeast extract in 950 mL deionized water (*see* **Note 4**).
9. 50 mg/mL Kanamycin stock solution (1000×): Store at −20 °C.
10. 100 mg/mL Ampicillin stock solution (1000×): Store at −20 °C.
11. *E. coli* DH5α competent cells.
12. Plasmid miniprep kit.
13. DNA gel extraction kit.
14. Plasmid pST1374 (Addgene): This is a backbone vector for constructing ZFN expression plasmids.
15. Plasmids pDsRED1-C1 (Takara, Mountain View, California, USA).
16. Plasmid pEGFP-C1 (Takara, Mountain View, California, USA).

Table 1
Primers for constructing the suicidal reporter vector

Primer	Sequence
GFPL-F	5′-GAAGATCTCGATCGAGTACTGGGCCCGTGAGCAAGGGCGAGGAG-3′
GFPL-R	5′-GAAGATCTGTGAGCAAGGGCGAGGAG-3′
GFPR-F	5′-CGCGGATCCGAGCTCGCGGCCGCGTCCGGCGAGGGCGAGGG-3′
GFPR-R	5′-GCTCTAGATTACTTGTACAGCTCGTC-3′
MBS-F	5′-GATCCCGACGGCTCCTTGGAAGACGATGACTACCACGTTACGACGGAAACGTTAGC-3′
MBS-R	5′-GGCCGCTAACGTTTCCGTCGTAACGTGGTAGTCATCGTCTTCCAAGGAGCCGTCGG-3′
RFP-F	5′-TATGCTAGCCACCATGGCCTCCTCCGAGGAC-3′
RFP-R	5′-GCTCTAGACTTAAGGCGCCG-3′
MSTN-F	5′-ACAGCGAGCAGAAGGAAAATG-3′
MSTN-R	5′-CAGCAGTCAGCAGAGTCGTTG-3′

17. Plasmid pcDNA3-mRFP (Takara, Mountain View, California, USA).

2.3 Cell Culture

1. Deionized water.
2. Dulbecco's modified Eagle medium (DMEM).
3. Fetal bovine serum (FBS).
4. 100× Antibiotic/antimycotic solution: 10,000 U/mL of penicillin and 10 mg/mL streptomycin.
5. 0.25% Trypsin-EDTA.
6. Human embryonic kidney 293T (HEK293T) cells, maintained in DMEM supplemented with 10% FBS and 100 U/mL penicillin and 100 μg/mL streptomycin at 37 °C with 5% CO_2.

3 Methods

3.1 Construction of ZFN Suicidal Plasmids

1. PCR amplify EGFP left and right arms using primer pairs GFPL-F/GFPL-R and GFPR-F/GFPR-R, respectively. Perform the PCR in a 50 μL reaction as follows: 5 μL of 10× PCR buffer, 4 μL of 10 mM dNTP mix, 0.5 μL of 100 ng/mL template plasmid pEGFP-C1, 0.5 μL of *Taq* DNA polymerase, and 0.5 μL of each 10 μM forward and reverse primers (Table 1) (*see* **Note 5**).
2. Incubate the reactions in a thermocycler under the following conditions: 94 °C for 5 min, 35 cycles of 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 60 s, and final extension at 72 °C for 10 min. Then keep the reactions at 4° C.
3. Purify PCR products by agarose gel electrophoresis using DNA gel extraction kit.
4. Digest pDsRed1-C1 and EGFPL PCR product, respectively, with *Bgl*II and *Sal*I as follows: 1 μg DNA, 3 μL of 10× Takara QuickCut buffer, 0.5 μL of each enzymes, and water up to 30 μL. Incubate the reaction at 37 °C for 4 h (*see* **Note 6**).
5. Gel purify the digestion products using DNA gel extraction kit following the manufacturer's instructions.
6. Ligate the digestion product in the following reaction: 2.5 μL of 20 ng/μL vector, 3 μL of 50 ng/μL insert, 1 μL of 10× ligation buffer, 0.5 μL T4 DNA ligase, and water up to 10 μL. Incubate the reaction at 16 °C for 4 h (*see* **Note 7**).
7. Transform the ligation products into *E. coli* DH5α competent cells (*see* **Note 8**) as follows: incubate 10 μL of the ligation product with 50 μL of cells on ice for 30 min. Heat shock the competent cells at 42 °C for 60 s and then incubate on ice for 5 min. Recover the cells in LB medium without antibiotics at

37 °C for 1 h. Spread the cells on LB agar plates supplemented with 50 μg/mL kanamycin.

8. The next day, pick single colonies from the plates, miniprep the plasmid DNA, and then verify the sequence of pDsRed1-C1-GFP-L by DNA sequencing.
9. Digest the PCR product of GFP right arm and pDsRed1-C1-GFP-L from **step 8** with *Bam*HI and *Xba*I. Follow **steps 2–8** to obtain pDG plasmid that contains a GFP sequence with an introduced stop codon.
10. Anneal primers MBS-F and MBS-R in the following reaction: 1 μL of each 10 μM forward and reverse primers, 1 μL of 10× Taq buffer, and water up to 10 μL. Incubate the reaction at 95 °C and slowly cool to room temperature.
11. Digest pDG plasmid with *Bam*HI and *Not*I as described in **steps 4** and **5** and ligate in the annealed MBS-F/MBS-R primers as follows: 5 μL of 10 ng/μL vector, 1 μL of annealing insert, 1 μL of 10× ligation buffer, 0.5 μL T4 DNA ligase, and water up to 10 μL. Incubate the reaction at 16 °C for 4 h.
12. Transform the ligation product as described in **steps 7** and **8** and obtain the plasmid pDGS with a *MSTN* target site.
13. PCR amplify the mRFP gene from pCDNA3-mRFP with primers RFP-F and RFP-R in a reaction as described in **steps 1** and **2**.
14. Digest pDGS plasmid and mRFP with *Nhe*I and *Xba*I following the protocol described in **steps 4** and **5**.
15. Ligate the mRFP PCR product into pDGS and transform as described in **steps 6–8** to obtain plasmid pRGS (*see* **Note 9**).
16. Digest the plasmid pRGS and *MSTN* ZFN expression cassette pGK_P-LAZFN-T2A-RZFN-BGHpA [5] with *Bgl*II and *Xho*I following the protocol described in **step 4** and clone the ZFN cassette in pRGS as described in **steps 6–8** to obtain plasmid pRGZS.
17. Clone the second *MSTN* target site in between the EGFP left arm and *MSTN* ZFN cassette in pRGZS with *Not*I and *Bam*HI. Follow **steps 1–8** to obtain the final suicidal plasmid pRGSZS (Fig. 1).

3.2 Validation of the Cellular Activity of ZFNs

1. At 24 h prior to transfection (*see* **Note 10**), seed 1 × 10^5 HEK293T cells on to 24-well plate in DMEM supplemented with 10% (v/v) FBS, 100 U/mL penicillin, and 100 μg/mL streptomycin at 37 °C with 5% (v/v) CO_2.
2. Transfect pRGS alone, regular ZFN system of pRGS, pLZFN (pST1374-LZFN) and pRZFN (pST1374-RZFN) [5], or suicidal reporter plasmid pRGZS or pRGSZS into HEK293T

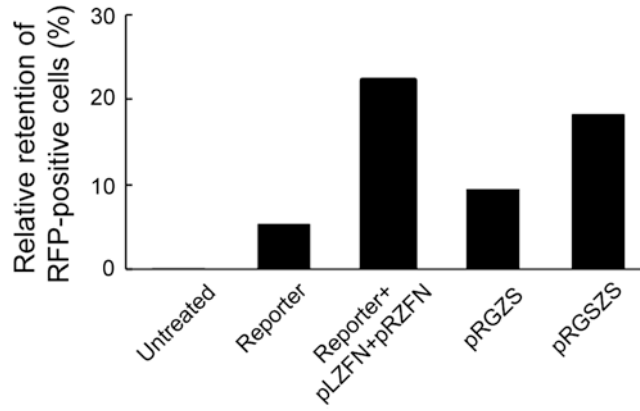


Fig. 2 ZFN activity as determined by flow cytometry. Reporter is constructed from suicidal plasmid where the ZFN expression cassette is removed and is used here to indicate the spontaneous, non-ZFN-induced homologous recombination

cells, respectively, using Lipofectamine 2000 according to the manufacturer's instructions.

3. At 24, 48, or 72 h after transfection, count 3×10^4 cells and analyze RFP- or EGFP-positive cells by flow cytometry (*see* **Note 11**). Aspirate the medium gently, wash the cells twice with 2 mL PBS buffer, then add 0.2 mL of 0.25% trypsin, and incubate for 1 min. Add another 1 mL DMEM medium supplemented with 10% FBS to suspend the cells. Transfection efficiency can be evaluated by the RFP-positive cells and ZFN activity by EGFP-positive cells (Fig. 2).
4. Spin down the cell pellet and extract the genomic DNA using genomic DNA extraction kit following the manufacturer's protocol.
5. PCR amplify the ZFN target site of *MSTN* with primers MSTN-F/MSTN-R (Table 1) using the conditions as described in **steps 1** and **2** of Subheading 3.1.
6. Confirm the gene modification of the ZFN target site by DNA sequencing. ZFN cleavage typically results in double signal peaks around the ZFN cutting site (Fig. 3) [5].

3.3 ZFN-Associated Toxicity Assay

1. Seed HEK293T cells as described in **step 1** in Subheading 3.2.
2. Perform transfection experiments for the following groups with a total amount of 1 μg DNA: negative control, 400 ng pRGS and 600 ng pST1374; ZFN expression control, 400 ng pRGS, 300 ng pLZFN, and 300 ng pRZFN; experimental group, 1 μg pRGSZS.
3. Evaluate the cytotoxicity of ZFNs in suicidal reporter plasmid pRGSZS using Cornu and Cathomen's method [6] as below.

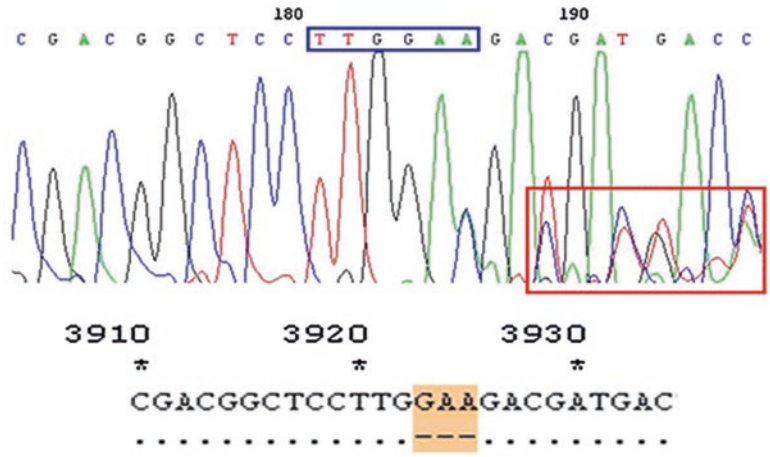


Fig. 3 Detection of ZFN-induced genetic modification by DNA sequencing. Double signal peak is indicated by the red box adjacent to the ZFN target site (blue box)

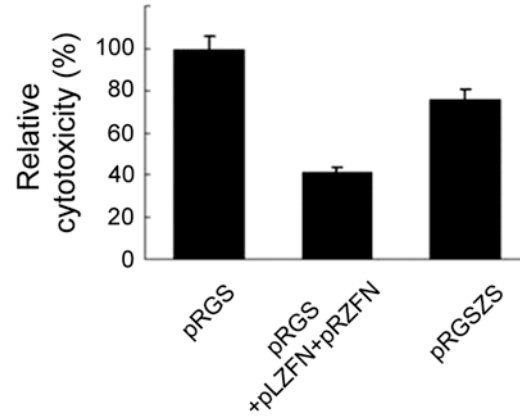


Fig. 4 Cytotoxicity assay of different plasmid systems

4. At 30 h post-transfection, detach and analyze 3×10^4 cells by flow cytometry as described in **step 3** in Subheading 3.2. Record the percentage of RFP-positive population.
5. Reseed the remaining cells and continue to culture for another 5 days.
6. Harvest the cells and analyze the RFP-positive population as described in **step 4**.
7. Calculate the retention of RFP-positive cells over 5 days. Normalize the retention ratio to the negative control. Compare the relative retention ratio as an indicator for the cytotoxicity of different plasmids (Fig. 4).

4 Notes

1. Nanodrop should be calibrated before use.
2. PCR master mix can also be used.
3. Restriction enzymes must be used with correct buffers and reaction temperatures.
4. The pH of LB medium should be adjusted to 7.0 using NaOH.
5. A stop codon is introduced into EGFP sequence to prevent its expression prior to ZFN cleavage.
6. Alternatively, incubate the reaction at 37 °C overnight.
7. Although the T4 DNA ligase manual recommends a ligation time of 20 min, we typically extend the ligation time to 2–4 h to obtain optimum ligation efficiency.
8. Other competent cell strains can also be used.
9. DsRed is replaced with mRFP to reduce overexpression-related cytotoxicity in HEK293T cells. DsRed gene can be used in other cell lines if less cytotoxicity is observed.
10. To achieve optimum transfection efficiency, cells should be seeded on to plates at 24 h prior to transfection.
11. Cells are typically harvested between 48 and 72 h after transfection as GFP or RFP signal is optimum during this period.

Acknowledgments

The authors thank the colleagues in Professor Zhang's lab for their technical assistance and helpful collaboration. We are grateful to financial support from China's Ministry of Agriculture (948 Program 2013-Z27), China's Ministry of Science and Technology (National Science and Technology Major Project 2014ZX0801009B and 973 Program 2011CBA01002), and National Natural Science Foundation of China (NSFC 31172186).

References

1. Lieber MR (2010) The mechanism of double-strand DNA break repair by the nonhomologous DNA end-joining pathway. Annu Rev Biochem 79:181–211
2. Lombardo A, Genovese P, Beausejour CM, Colleoni S, Lee YL, Kim KA et al (2007) Gene editing in human stem cells using zinc finger nucleases and integrase-defective lentiviral vector delivery. Nat Biotechnol 25:1298–1306
3. Pruett-Miller SM, Connelly JP, Maeder ML, Joung JK, Porteus MH (2008) Comparison of zinc finger nucleases for use in gene targeting in mammalian cells. Mol Ther 16:707–717
4. Kim H, Um E, Cho SR, Jung S, Kim H, Kim JS (2011) Surrogate reporters for enrichment of cells with nuclease-induced mutations. Nat Methods 8:941–943
5. Zhang C, Wang L, Ren G, Li Z, Ren C, Zhang T et al (2014) Targeted disruption of the sheep *MSTN* gene by engineered zinc-finger nucleases. Mol Biol Rep 41:209–215
6. Cornu TI, Cathomen T (2010) Quantification of zinc finger nuclease-associated toxicity. Methods Mol Biol 649:237–245

Part III

Delivery Methods of ZFPs

Chapter 14

Non-transgenic Approach to Deliver ZFNs in Seeds for Targeted Genome Engineering

Zoe Hilioti

Abstract

In the post-genomic era, the efficient exploitation of the available information for plant breeding is a pressing problem. The discoveries that DNA double-stranded breaks (DSBs) are both recombinagenic and mutagenic have fuelled the development of targetable zinc-finger nucleases (ZFNs), which act as molecular scissors for the induction of controlled DSBs. These powerful tools are used by researchers to accelerate mutagenesis of the normal gene loci toward the development of useful traits in plants. Seeds contain the embryo, which is a multicellular system representing a micrography of a plant. Therefore, they can serve as a foundation for applying targeted genome engineering techniques. The following single-step method describes how to deliver and express transiently ZFNs in tomato (*Solanum lycopersicum*) seeds using electroporation. Unlike methods that rely on tissue culture and plant regeneration after transformation, the direct delivery of ZFNs to seeds provides a high-throughput breeding technology for safe and site-specific mutagenesis. Tomato is a leading crop in the world and biotechnological advances in this species have great impact.

Key words Tomato, Seeds, Transformation, Electroporation, ZFNs

1 Introduction

Until recent years, all attempts to develop robust and controllable gene targeting techniques in higher plants for precision breeding were inefficient. Traditional plant breeding techniques such as classic breeding and mutations generated through the use of DNA-damaging agents (EMS and X-rays) [1, 2] introduced variation in the genetic makeup of modern varieties. These non-targeted mutagenesis approaches led to the identification of interesting phenotypes and subsequent extensive work on the determination of their genetic basis.

A breakthrough in mammalian and plant cell genetic manipulation was the discovery that generation of a double-strand break (DSB) within a target DNA sequence dramatically increases the frequency of gene targeting anywhere from 50- up to 5000-fold

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_14,

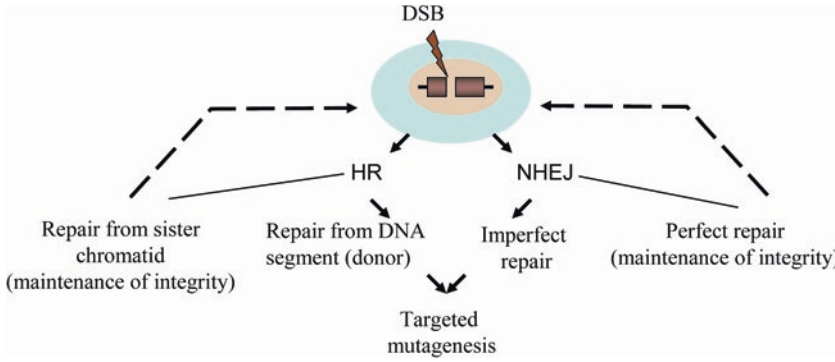


Fig. 1 A scheme to ZFN-induced DSB and repair mechanisms of chromosomal DNA damage. Targeted mutagenesis occurs as a result of (i) homologous recombination (HR) with donor DNA and (ii) imperfect repair during nonhomologous end joining (NHEJ). Correct repair of the DSB site can be accomplished by HR as well as by NHEJ in a cell cycle-dependent manner resulting in the maintenance of genome integrity

[3–5]. This finding generated excitement among researchers that DSBs could be exploited as a molecular tool in plant transformation techniques. DNA cleavage triggers cell-assisted repair of DSBs between highly homologous and identical DNA segments by the homologous recombination (HR) endogenous process [6] or of nonhomologous DNA segments by error-prone nonhomologous end joining (NHEJ) (Fig. 1). During HR, the cell repairs the DSBs using a template with high homology at DSB ends.

In plants, pioneering experiments on DSB repair mechanisms showed that DSBs are predominantly repaired by NHEJ while HR is common in cells undergoing meiosis [7–10]. During meiosis, the precursor diploid cell, which gives rise to four haploid cells, contains template for the DSB repair—the sister chromatid—and the HR-stimulated repair is highly accurate. Usually, during this process, the DSB break is sealed by a DNA sequence containing homology to the DNA sequence on each side of the break. Such sequences can be found from (i) the sister chromatid which is available after DNA replication and in close proximity, (ii) a homologous sequence at close proximity to DSB, and (iiii) an extrachromosomal homologous donor DNA. Through HR an external piece of DNA with homologous regions can be integrated into the corresponding genomic locus, enabling gene editing and creation of gene knock-in, knockout, and point mutations depending on the sequence provided inside the homologous regions.

NHEJ-driven imperfect repair of DSBs leads to mutations as deletions or insertions of DNA sequence at the cleavage site. This may result in frame shift and generation of null alleles. There is, however, the possibility of perfect DSB repair by NHEJ. In contrast, HR repairs DNA in a precise manner based on a DNA template containing DNA homologous to the site of the genetic lesion. It has been shown that germinating seeds are more sensitive than

seedlings to NHEJ-mediated DNA repair indicating that the requirement for NHEJ is developmental stage -specific [11]. Environmental factors such as temperature and light affect the frequency of homologous recombination in plants by an unknown mechanism. Specifically, the rate of homologous recombination is higher in plants grown under suboptimal temperatures (4 or 22 °C) or conditions of short-day exposure [12]. DSBs may occur daily in the DNA of a eukaryotic organism during exposure to ionizing radiation, elevated concentrations of free radicals after pathogen attack, and also DNA replication. A quick and efficient method for generating a direct DSB in a preferred genomic site relies on the design and construction of engineered nucleases with long DNA sequence recognition and DNA binding specificity providing minimal DNA damage. Genome engineering technology requires DNA-binding and -cutting enzymes which can target ideally a unique sequence in the genome and modify individual genes.

In 1996, Chandrasegaran and his team created chimeric restriction enzymes by fusing three zinc fingers to Fok I DNA endonuclease domain [13]. The novel zinc-finger nucleases (ZFNs) are chimeric proteins consisting of engineered zinc fingers (ZFs) associated with a nonspecific cleavage domain from the prototypic type IIS FokI restriction endonuclease (N) [14–16]. The ZF module is at the N-terminus of the chimeric protein. ZFNs require formation of a homodimer for endonuclease activity while the DNA-binding motif specified by the zinc fingers directs ZFN to a specific locus in the genome. Hence, ZFNs act on the DNA in pairs causing a DSB. The co-expression of the two ZFN monomers in the cell is required. The delivery of ZFN monomers to constitute a functional ZFN pair in the same cell may be severely challenged when attempting to deliver in addition to ZFNs, an exogenous DNA template to stimulate HR-driven gene editing in a cell. A potential drawback of the ZFNs could be non-specificity caused by off-target formation of homodimers resulting in off-target DNA cleavage of the genome. To circumvent this problem, pioneer work by two groups [17, 18] developed complementary substitutions in the dimmer interface that prevent homodimer formation while still allowing the heterodimer formation and cleavage of the targeted DNA site. Bioinformatics tools can be used to identify possible high-similarity sequences to target sequence and sequencing analysis ultimately determines if these sites have been modified.

A minimum of three zinc-finger proteins (ZFPs) assembled in tandem are required to recognize a targeted DNA site [19]. Each pair of three finger ZFPs will recognize a DNA sequence of 18 bp, which occurs once every 6.9×10^{10} bp. This, for example, should be very specific for tomato (genome size $<10^{10}$ bp) but not for all flowering plants. With the development of eukaryotic expression systems, ZFNs can be produced by expression of a variable num-

ber of zinc fingers with plant codon-optimized Fok I endonuclease sequence. Properly designed ZFNs are versatile molecular tools that can act as scissors against any DNA sequence.

Overall, the ZFN-mediated gene targeting involves different progressive steps: (i) identification of plausible ZFP recognition sites in the target gene sequence, (ii) selection and design of zinc-finger array (ZFPs), (iii) assembly of the individual zinc fingers into a zing finger array, and (iv) cloning of the ZF-coding sequences in frame with the Fok I nuclease domain. To get the ZFN technology available to research community for broad application in various species, the zinc-finger consortium (http://www.zincfingers.org) developed protocols, software, and research tools for engineering zinc-finger arrays. Once the ZFNs are designed and assembled their activity can be validated in ex planta models such as the single-celled yeast. A yeast-based recombination assay has been developed to test whether the designed ZF proteins function as ZFNs [20, 21]. A modified yeast recombination assay has been developed by Townsend and co-workers [22].

ZFNs can be provided in plant cells as DNA (e.g., plasmid DNA) or mRNA transcripts using various delivery methods. In general, plasmid DNA is stable and easy to produce and has the advantage of regulated expression under different promoters. Naked DNA introduced into cells by direct gene transfer methods can be stably transformed in a fraction of cells under selection pressure. However, the transient expression (selection free) of ZFNs has advantages over the stably expressed ZFNs as it is less toxic to the plant cells and there is no need to eliminate the ZFN encoding DNA in mutant lines. Without selection the transgenic sequences will be eventually degraded by the cellular machinery and the mutated cells will be devoid of foreign DNA. Another application of ZFNs is that if two pairs of ZFNs are used to target sites on the same chromosome for the creation of two concurrent DSBs large regions of the genome may be excised. In this case, two pairs of ZFNs are simultaneously expressed for the removal of transgenes (e.g., selectable marker, reporter gene), promoter sequences, and exons. A drawback of ZFNs is that the affinity and specificity of the engineered ZFP arrays (left and right of target site) are not totally predictable. This is because the ZFPs do not function completely independently in recognizing the target DNA. In the process of DNA recognition and binding the specific attributes of neighboring zinc fingers matter as they do not work in a perfectly modular fashion. These positional effects are addressed in Mani et al. [23]. In some cases, the in vitro assay of ZFN activity poorly correlates with the in vivo outcome. Therefore, the ultimate *in planta* test for evaluating the functionality and specificity of ZFNs is required. Samples can be tested for gene editing at the locus of interest by standard molecular biology tech-

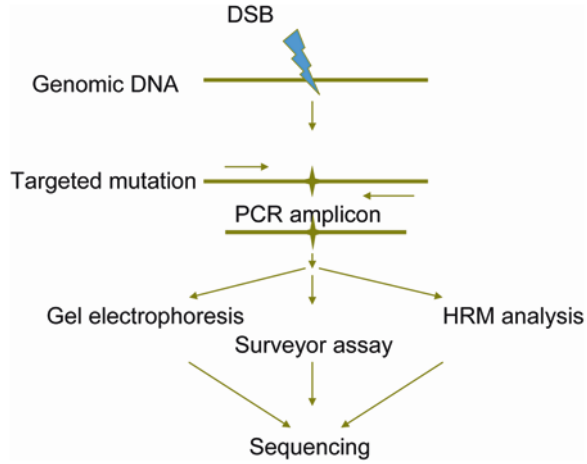


Fig. 2 A screening strategy for genotyping potential genomic mutations after ZFN-induced DSB

niques including PCR amplification, HRM analysis, enzymatic digest, and electrophoresis (Fig. 2). Complete, incomplete, non-specific, or even no cleavage of the DNA are all possible outcomes of the ZFN action. These outcomes may indicate good or bad ZFN design but they can also be explained by the non-proper ZFN expression level, inaccessibility of ZFNs to target region due to chromatin structure, problem with PCR-based genomic amplification, or low detection level in case of plant chimeras. Hence, a negative result from the above assays may not be very informative; a positive result will definitely encourage further analyses. Proper expression of the ZFNs into the target plant cell is the primary factor for both efficacy and specificity of gene targeting. Although plasmid constructs carrying ZFNs were frequently introduced in plant cells by indirect transfer techniques (e.g., *Agrobacterium*, viral vectors), direct DNA transfer to plant cells has been published and optimized for difficult-to-transform plant species and for achieving a desirable level of gene expression. Among those methods regeneration-independent approaches are free of potential problems such as somaclonal variation and they don't require lengthy incubation periods in controlled environments. Typically, direct DNA uptake can be used for both stable and transient expression of chimeric endonucleases. As the efforts of obtaining plant transformation for proof-of-principle approaches shift to approaches for practical precision breeding of economically important plant species, certain combinations of the published methods for delivering DNA (or RNA) may be suitable for approaching a high-throughput level. One important consideration when deciding what plant transformation method to use is

whether the plant is propagated by seeds or vegetative methods. In case of seed-propagated plant, stable transmission of DNA mutation(s) to germline tissue is highly desirable.

Our choice of method for ZFN delivery in plants is DNA electroporation. By exposing cells to short and intense electric pulses the phospholipid bilayer of cells is disrupted and microscopic pores are formed through which DNA is allowed to enter into the cells. The method is quick and inexpensive and can result in high transformation efficiency. Electroporation parameters for different plant tissues have been reported. In general, the optimum electric field is dependent on (a) the pulse length, (b) the composition and temperature of the electroporation medium, (c) the concentration of foreign DNA, and (d) the type of plant system being used regarding the presence or absence of cell wall (seeds, leaf, protoplasts, etc.). Typically, high concentration of plasmid DNA containing the ZFN genes is added to a suspension of plant tissue and the mixture is given a shock with an electric field of 200–600 V/cm. The plant tissue is then grown in tissue culture under selection pressure to select the transformed ones. The method has been used for ZFN delivery into tobacco protoplasts [22]. Our group has used electroporation for ZFN delivery and transient expression into tomato seeds [24].

Tomato genetically is a diploid crop. The target of transformation in seeds is the embryo composed of primordial leaves, axillary buds, and apical meristem. Heritable mutations are those that occur in cells of the apical meristem that contribute to the formation of gametes in flowers. During ZFN-induced mutagenesis, it is possible that the pollen may carry different mutation than the egg and as a result it is advisable to screen individual plant progenies. Mature seeds of homozygous pure line are excellent materials for mutagenesis as they are readily available all year around and do not require lengthy preparation time as starting material for transformation. The primary goal of using the multicellular embryo of seeds to deliver ZFNs is to obtain viable mutant plants that may contain different chimeric tissues and produce adequate M1 seeds to advance to later generations. The formation of chimera requires careful handling and screening for efficient mutation identification and breeding. Genome engineers face the challenge of how to develop a practical and efficient screening procedure of mutation selection from a large pool of variants. The cost of screening may be an important burden for some laboratories especially in cases of low efficiency of targeted mutagenesis. The ideal case scenario is when genotypic variation is linked to phenotypic variation that can be detected early either macroscopically or microscopically. Nonheritable ZFN-induced mutations that result in sterility may be prominent in M_1 and M_2 generations. Other phenotypes, which

are easy to score visually, are earliness, plant height and structure, color variation (e.g., leaves, fruits), and shape changes (e.g., leaves, trichomes, fruits). In self-pollinated tomato plants, when mutations are stabilized (M_2–M_4) and there is sufficiency in seeds, further tests in open field can be conducted (e.g., for yield). A visual selection encourages further detailed examination of the targeted DNA region by PCR-based approaches. In cases that the mutated character is not easily scored, a "genotype-first" approach may be followed. PCR-based amplification techniques which are used to recover potential genomic modification events at the target locus include junction PCR and RLFP assay. The techniques are quick and have different advantages. For example, junction PCR is sensitive but not as quantitative as the RLFP assay and amplifies the integration events only. The RFLP assay, from the other side, is less sensitive but more quantitative than junction PCR and is used for the amplification of both integration and non-integration events. The amplification products from screening primary transgenic plants (and their progenies) are subjected to gel electrophoresis side by side with the amplification products from non-treated plant specimens and the banding pattern is compared. In case a donor DNA strategy is used to induce HR-driven mutagenesis, the knowledge of the engineered DNA and possible alterations (if any) expected in the restriction digestion pattern of the engineered DNA fragment and/or in DNA length can give clues about the expected results. A weak point of the DNA resolution on gel is that DNA modifications that result in small changes in DNA length are going to be missed by this method. To get over this problem, the surveyor assay uses heating of the PCR amplicons followed by slow cooling which allows the formation of heteroduplexes from the annealing/hybridization of wild-type and mutant strands of DNA. The rationale behind this approach is that only the mutated portion of the DNA will remain single stranded leaving the common DNA to both treated and untreated cell pools hybridized. The presence of excess DNA from untreated plants is an important factor that ensures heteroduplex formation. Subsequently, the heteroduplexes are treated with a single-strand-specific endonuclease CEL1 which cuts 3′ of single-base mismatches to yield novel fragments of DNA that can be resolved on a gel. These novel fragments can be further analyzed by sequencing to determine if the mutation is predicted to be silent, missense, or knockout mutation. The endonuclease digestion step is critical for this assay and fraught with potential problems common to enzymatic reactions. Another powerful, fast, and high-throughput genotyping technique, which does not involve an enzymatic digestion, is the high-resolution melt analysis (HRM) after amplification of target sequence by real-time PCR with the incorporation of a fluorescent dye that

specifically binds double-stranded DNA (dsDNA), followed by melt curve analysis of the amplicons. Each PCR product has its own melting temperature which is dependent upon the length and base composition of DNA and heterozygosity of DNA sample. Thus, several DNA samples can be screened at once and the high-throughput processing of the samples minimizes the time, labor, and cost of screening without compromising the level of mutation detection efficiency. These parameters become significant when the nature of the mutations is not known as in the donor-free strategies for genome engineering which are controlled by the NHEJ repair pathway. Non-transgenic approaches may be selection free (foreign DNA is introduced in cells but eliminated prior to or during mitosis process) for the transformants which typically means extensive screening of a heterologous mix of genomic DNAs.

In practice ZFNs offer a valuable toolbox for precision breeding that can be used by researchers to increase genetic diversity in gene loci and improve agronomical traits. Through proper design of ZFNs and delivery method, scientists can navigate the tangle of laws and regulations and deliver acceptable to consumers' agricultural products.

2 Materials

1. Plant expression vector (e.g., pDW1775, Addgene) carrying the left and right ZFN constructs.
2. Plasmid purification kit.
3. A sterile laminar flow hood equipped with a desiccator connected to a vacuum pump.
4. Electroporator.
5. 0.4 cm Electroporation cuvettes.
6. ZFN DNA.
7. Germination buffer: 5% Sucrose, 3% H_3BO_3, 1.3 mM $Ca(NO_3)_2$.
8. Electroporation buffer: 80 mM KCl, 5 mM $CaCl_2$, 10 mM Hepes, 0.5 M mannitol. Adjust to pH 7.2 and filter using a 0.22 μm filter.
9. 6 cm Petri dishes.
10. Seed boxes or Jiffy pots.
11. Forceps, sterile.
12. Plant material: Tomato seeds (e.g., cv Heinz 1706 *see* **Note 1**); for a medium scale you may use approximately 40 seeds with 9 to 10 seeds per cuvette.

3 Methods

Electroporation uses electrical current to temporarily open pores in the membrane of plant cell through which salt-free DNA may pass. Electroporation-based delivery of ZFNs for transient expression and targeted mutagenesis of tomato seeds described herein promoted highly efficient disruption of the *L1L4* gene in *S. lycopersicum* [24]. The protocol employs germinating seeds as starting plant material and assumes basic knowledge about sterile conditions and molecular biology techniques. The procedure described here for transformation of seeds is fast and can be scaled up for transforming greater number of seeds taking into account that the target of tomato production is the fruits (Fig. 3).

3.1 Seed Pre-germination and Direct Introduction of ZFN DNA into Seeds by Electroporation

3.1.1 Day 1

1. Visually inspect mature tomato seeds and choose large undamaged seeds.
2. In a sterile hood, surface sterilize seeds by immersion in 70% ethanol for 20 s. Wash seeds with sterile distilled water and then add 3.5% bleach for 10 min. Wash surface-sterilized seeds three times using sterile water for each wash.
3. After sterilization, use sterile forceps to transfer seeds onto a 6 cm petri dish to germinate for 12 h in germination buffer and place them at 10 °C in the dark.

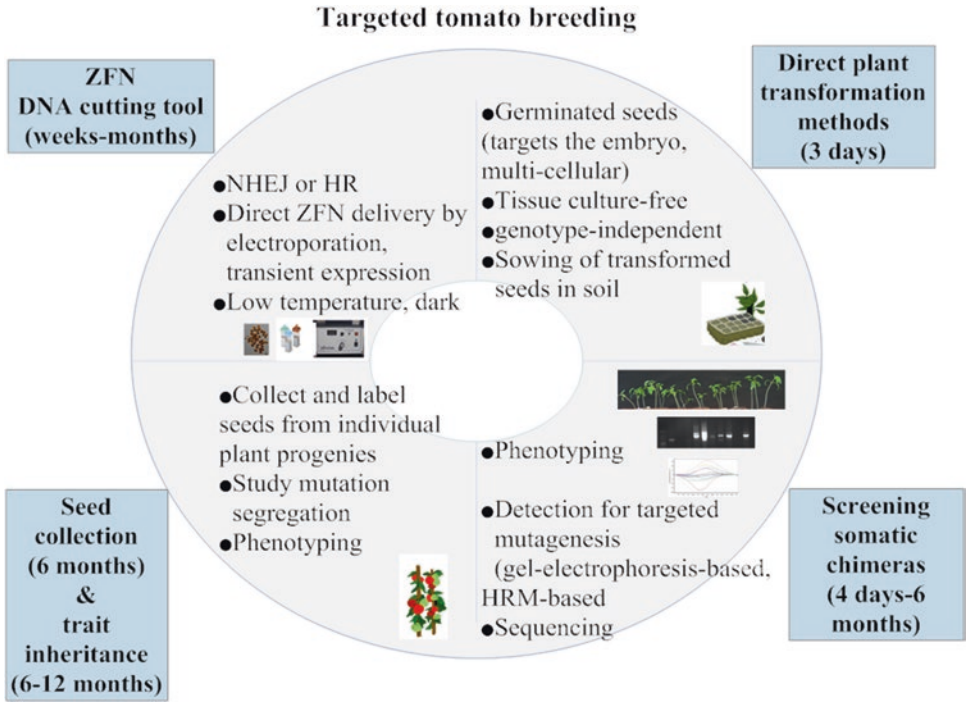


Fig. 3 Protocol for the ZFN-based targeted genome engineering of *S. lycopersicum* without selection

4. For ZFN pair expression supplied as plasmid DNA (two DNA fragments encoding the left and right ZFN) a fresh plasmid purification should be prepared as described by the manufacturer's protocol (mini or midi scale) prior to electroporation. In doing so, inoculate cultures from a single *E. coli* colony for overnight growth at 37 °C and extract DNA the next day.

3.1.2 Day 2

1. Proceed with plasmid DNA extraction (DNA for electroporation must be free of salt) and put DNA on ice.
2. Remove germination buffer from the seeds in petri dish, add filter-sterilized electroporation buffer, apply vacuum conditions (0.09 MPa) for 10 min, and allow the DNA-seed mixture to stand for 1 h in the vacuum chamber. Thereafter, transfer seeds to pre-chilled sterile electroporation cuvettes and suspend them in 200 μL electroporation buffer containing 50 μg of freshly prepared and well-mixed plasmid DNA (25 μg for each ZFN monomer). Cap the cuvette.
3. Put cuvettes containing the seeds and buffer on ice for 10 min prior to electroporation.
4. Wipe cuvettes with tissue paper prior to electroporation. Set the electroporation conditions on a commercial electroporator device, press the button to electroporate the seeds, and apply three electric pulses, each with duration of 4 ms and field strength of 6.25 kV cm^{-1}.
5. After electroporation, place cuvettes on ice for 1 h. Thereafter, transfer the seeds and the buffer solution in petri dishes under the hood area. Add appropriate volume of electroporation buffer or distilled water to keep seeds moist and allow them to grow at 10 °C for 24 h in the dark.

3.1.3 Day 3

1. Sow seeds (*see* **Note 2**) in Jiffy pots or 15 cm pots and place under controlled growth chamber conditions. Jiffy pots have small volume, are easy to handle, and facilitate frequent phenotypic (macroscopic and microscopic) data collection of a large number of transformants.

3.2 Phenotyping and Genotyping Transformants

1. Wait about 7 days for seeds to germinate post-sowing. Record physiological characters throughout development (survival, early germination, seedling height, plant structure, trichomes, number of seeds per fruit) for practical screening as parental plants. They may exhibit chimerism which is easy to detect at young seedling stage (up to 2–3 weeks post-germination) due to rapid growth while it becomes challenging to detect it in mature plants. Record the distinctive characters of plants in terms of their morphology, and behavior under external conditions until the maturation stage.

2. To extract genomic DNA (using your method of choice) and screen for successful targeting events, harvest vegetative samples (e.g., leaves), freeze in liquid nitrogen, and store at −80 °C until needed. Proceed with grinding of plant material to a fine powder using liquid nitrogen and subsequent DNA extraction.
3. Detect on-target mutations in the heterogeneous cell population of chimeric tissues (*see* **Note 3**) by HRM and gel electrophoresis of amplicons using the wild-type DNA as reference. Proceed with DNA extraction and purification of the band of interest from agarose gel. Confirm mutations by sequencing to ultimately determine the outcome of genome editing. Off-target mutations in highly homologous to on-target sites can be checked in the same way. Whole-genome sequencing, however, is the best strategy to evaluate ZFN-induced changes.
4. Collect seeds from individual fruits of each founder plant and study mutation inheritance.

   Bulk seed screening is advised only when the phenotype is well -known. Repeat screening, phenotyping, genotyping, and labeling in succeeding generations until homozygous lines are achieved.

4 Notes

1. The cv Heinz 1706 has been sequenced and serves as a reference genome for tomato [25]. The variety displays cleistogamy allowing self-pollination, which increases homozygosity in progenies.
2. Transient expression of ZFNs in seeds may minimize possible cytotoxicity problems and ensure that foreign DNA would not be incorporated into the cells yielding, thus, non-transgenic plants. In case of ZFN toxicity, however, an alternative approach would be to use variant ZFNs of the obligate heterodimer-type Fok I, which have high fidelity. A mutagenesis approach to generate E490K:I538K and Q486E:I499L variants in pDW1775 containing the wild-type Fok I gene is provided in [24].
3. It is expected that as a result of ZFN function the embryo will be a chimera of the original cells and the altered ones. In the pooled genomic DNA sample from chimeric tissue, diploid tomato cells may carry a heterozygous mutation where only one allele is mutated, some others a biallelic mutation where both alleles are mutated but the sequence of each allele is different, or in others a homozygous mutation may be present where both alleles carry the same mutation. In essence, each

successfully mutated seed may represent a "library" of gene mutations and the heterogeneity of DNA may complicate sequencing efforts in founders.

References

1. Hildering GJ, Verkerk K (1965) Chimeric structure of the tomato plant after seed treatment with EMS and X-rays. The use of induced mutations in plant breeding. Pergamon Press, Oxford, pp 317–320
2. Schoenmakers HCH, Koornneef M, Alefs SJHM, Gerrits WFM, van der Kop D, Cherel I, Caboche M (1991) Isolation and characterization of nitrate reductase-deficient mutants in tomato (Lycopersicon esculentum Mill.). Mol Gen Genet 227:458–464
3. Choulika A, Perrin A, Dujon B, Nicolas JF (1995) Induction of homologous recombination in mammalian chromosomes by using the I-SceI system of Saccharomyces cerevisiae. Mol Cell Biol 15:1968–1973
4. Donoho G, Jasin M, Berg P (1998) Analysis of gene targeting and intrachromosomal homologous recombination stimulated by genomic double-strand breaks in mouse embryonic stem cells. Mol Cell Biol 18:4070–4078
5. Puchta H, Dujon B, Hohn B (1993) Homologous recombination in plant cells is enhanced by in vivo induction of double strand breaks into DNA by a site-specific endonuclease. Nucleic Acids Res 21:5034–5040
6. Jasin M, Moynahan ME, Richardson C (1996) Targeted transgenesis. Proc Natl Acad Sci U S A 93:8804–8808
7. D'Halluin K, Vanderstraeten C, Stals E, Cornelissen M, Ruiter R (2008) Homologous recombination: a basis for targeted genome optimization in crop species such as maize. Plant Biotechnol J 6:93–102
8. Orel N, Kyryk A, Puchta H (2003) Different pathways of homologous recombination are used for the repair of double-strand breaks within tandemly arranged sequences in the plant genome. Plant J 35:604–612
9. Puchta H (2005) The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. J Exp Bot 56:1–14
10. Salomon S, Puchta H (1998) Capture of genomic and T-DNA sequences during double-strand break repair in somatic plant cells. EMBO J 17:6086–6095
11. Riha K, Watson JM, Parkey J, Shippen DE (2002) Telomere length deregulation and enhanced sensitivity to genotoxic stress in Arabidopsis mutants deficient in KU70. EMBO J 21:2819–2826
12. Boyko A, Filkowski J, Kovalchuk I (2005) Homologous recombination in plants is temperature and day-length dependent. Mutat Res 572:73–83
13. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93:1156–1160
14. Cathomen T, Keith Joung J (2008) Zinc-finger nucleases: the next generation emerges. Mol Ther 16:1200–1207
15. Durai S, Mani M, Kandavelou K, Wu J, Porteus MH, Chandrasegaran S (2005) Zinc finger nucleases: custom-designed molecular scissors for genome engineering of plant and mammalian cells. Nucleic Acids Res 33:5978–5990
16. Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. Nat Biotechnol 23:967–973
17. Miller JC, Holmes MC, Wang J, Guschin DY, Lee YL, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA, Gregory PD, Pabo CO, Rebar EJ (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25:778–785
18. Szczepek M, Brondani V, Büchel J, Serrano L, Segal DJ, Cathomen T (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc finger nucleases. Nat Biotechnol 25:786–793
19. Bibikova M, Carroll D, Segal DJ, Trautman JK, Smith J, Kim YG, Chandrasegaran S (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. Mol Cell Biol 21:289–297
20. Doyon Y, M McCammon J, C Miller J, Faraji F, Ngo C, E Katibah G, Amora R, D Hocking T, Zhang L, J Rebar E, D Gregory P, D Urnov F, L Amacher S (2008) Heritable targeted gene disruption in zebrafish using designed zinc finger nucleases. Nat Biotechnol 26:702–708
21. Epinat J-C, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Pâques F, Lacroix E (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. Nucleic Acids Res 31:2952–2962

22. Townsend JA, Wright DA, Winfrey RJ, Fu F, Maeder ML, Joung JK, Voytas DF (2009) High frequency modification of plant genes using engineered zinc finger nucleases. Nature 459:442–445
23. Mani M, Kandavelou K, Dy FJ, Durai S, Chandrasegaran S (2005) Design, engineering, and characterization of zinc finger nucleases. Biochem Biophys Res Commun 335:447–457
24. Hilioti Z, Ganopoulos I, Ajith S, Bossis I, Tsaftaris A (2016) A novel arrangement of zinc finger nuclease system for in vivo targeted genome engineering: the tomato LEC1-LIKE4 gene case. Plant Cell Rep 35:2241–2255
25. Tomato Genome C (2012) The tomato genome sequence provides insights into fleshy fruit evolution. Nature 485:635–641

Chapter 15

Gene Editing in Channel Catfish via Double Electroporation of Zinc-Finger Nucleases

Rex A. Dunham, Ahmed Elaswad, and Zhenkui Qin

Abstract

The traditional approach for gene editing with zinc-finger nucleases (ZFNs) in fish has been microinjection of mRNA. Here, we develop and describe an alternative protocol in which ZFN plasmids are electroporated to channel catfish, *Ictalurus punctatus*, sperm, and embryos. Briefly, plasmids were propagated to supply a sufficient quantity for electroporation. Sperm cells were prepared in saline solution, electroporated with plasmids, and then used for fertilization. Embryos were incubated with the plasmids before performing electroporation just prior to first cell division. Plasmids were then transcribed and translated by embryonic cells to produce ZFNs for gene editing, resulting in mutated fry.

Key words Channel catfish, Electroporation, Zinc-finger nuclease, Plasmids, Sperm-mediated gene transfer, Gene editing, Mutation

1 Introduction

Zinc-finger nucleases (ZFNs) are one of the engineered nucleases that have been used to edit genes in different organisms [1, 2]. ZFN consists of two domains: (a) DNA-binding domain, which recognizes and binds to the DNA target sequence, and (b) DNA-cleaving domain (endonuclease activity, *FokI*), which cuts the DNA at the target site [3]. ZFNs are engineered in pairs, one for each DNA strand. Together, they induce double-strand break (DSB) which will then be repaired by an error-prone DNA repair mechanism, nonhomologous end joining (NHEJ), resulting in insertions, deletions, and frameshift mutations. The applications of ZFN technology include gene knockout [2, 4], insertion [5], and correction [6, 7]. Most of ZFN-based gene editing in fish focused on gene knockout as in, for example, zebrafish (*Danio rerio*) [2, 4], channel catfish (*Ictalurus punctatus*) [8], yellow catfish (*Pelteobagrus fulvidraco*) [9], and rainbow trout (*Oncorhynchus mykiss*) [10].

Electroporation is a technique in which electrical impulses are applied to cells to disrupt the selective permeability and facilitate

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_15,

the passage of large biological molecules such as DNA [11]. The technique is widely used to transform bacterial cells [12, 13], yeast [14], plants [15, 16], and mammalian cells and tissues [11, 17, 18]. Electroporation has many applications in medicine [19], biology [20], cell biology [21], and biotechnology [22]. In fish, electroporation has been routinely used to generate transgenic fish through the introduction of transgene constructs into one-celled embryos such as channel catfish [23, 24], common carp (*Cyprinus carpio*) [23], medaka (*Oryzias latipes*) [25, 26], and zebrafish [23, 27]. Transgene constructs were also electroporated into sperm in a technique called "sperm-mediated gene transfer" [28]. In such a case, sperm will transport the transgene construct and introduce it into the egg during the fertilization process. Transgenic fish species were generated with this technique, for example, silver sea bream (*Sparus sarba*) [29], chinook salmon (*Oncorhynchus tshawytscha*) [30], grass carp (*Ctenopharyngodon idellus*) [31], zebrafish [32] and loach (*Misgurnus anguillicaudatus*) [33]. To increase the possibility of transgene insertion, electroporation of sperm and one-celled embryos can be combined in a single technique called "double electroporation" where sperm is electroporated and used to fertilize the eggs, and then fertilized eggs are electroporated at the one-cell stage [34].

The traditional approach to deliver ZFNs to fish embryos was through microinjection of ZFN-expressing mRNA [2, 4, 35] which will then be translated into functional ZFN domains before targeting the genome. However, plasmid constructs expressing transcription activator-like effector nucleases (TALENs) were electroporated into the ascidian chordate (*Ciona intestinalis*) to knock out multiple genes [36]. In addition, Cas9/gRNA plasmid DNA, Cas9 mRNA/gRNA, and Cas9 protein/gRNA complexes were electroporated into a variety of mammalian cells for gene editing [37]. Here, we describe an alternative protocol that utilizes double electroporation of ZFN-expressing plasmid constructs into the sperm and one-celled embryos to generate gene-edited channel catfish. The protocol is simple and does not require in vitro transcription and many founder fish can be generated (150–200 embryos can be electroporated in a single batch). In this protocol, both sperm and embryos are electroporated with ZFN-expressing constructs. First, sperm is incubated with plasmids for at least 5 min in 0.9% saline solution, and then freshwater is added immediately before electroporation. Addition of freshwater will hydrate and activate the sperm, and increase the uptake of plasmids by the sperm [34]. Second, embryos are incubated with plasmids for at least 10 min before electroporation. ZFN constructs will then be transcribed and translated by the embryonic cells which might affect the level of mosaicism [38]. Qin et al. [8] found no evidence of plasmid DNA integration in the channel catfish genome when

ZFN plasmids were double electroporated. However, there is a possibility of ZFN plasmids integrating into the channel catfish genome generating transgenic gene-edited fish [39]. Therefore, founder fish generated with double electroporation of ZFN plasmids will need to be screened for potential gene integration.

2 Materials

This protocol is optimized for the equipment and materials below. Other alternatives may be used, however, for which further optimization may be required.

1. Electroporator (Fig. 1a, *see* **Note 1**).
2. 37 °C Shaking and non-shaking incubator.
3. Eppendorf centrifuge.
4. 50 mL Tube centrifuge
5. 200–300 mL Centrifuge.
6. 42 °C Water bath.
7. Thermal cycler (optional).
8. ZFN-expressing plasmid constructs: ZFN constructs can be designed and purchased from Sigma-Aldrich (St. Louis, Missouri, USA) [8] or any molecular biology or biotechnology supplier.
9. Lysogeny broth (LB): 1% Tryptone, 0.5% yeast extract, 0.5% NaCl. To prepare 100 L LB broth, weigh the following in a graduated flask: 1 g tryptone, 0.5 g yeast extract, and 0.5 g NaCl and add up to 100 mL distilled water. Adjust the pH to 7.0 with NaOH and autoclave for 20 min. Add 50 ppm kanamycin before use (*see* **Note 2**).
10. Surveyor mutation detection kit (Integrated DNA Technologies, Inc., Coralville, Iowa, USA).
11. Expand™ High Fidelity PLUS PCR System (Sigma-Aldrich, St. Louis, Missouri, USA).
12. 10 cm LB agar plates: 1% Tryptone, 0.5% yeast extract, 0.5% NaCl, 1.2% agar, 50 ppm kanamycin (*see* **Note 3**).
13. TE buffer: 10 mM Tris–Cl, 1 mM EDTA, pH 8.0.
14. 0.9% Saline: 0.9 g NaCl, add 100 mL distilled water.
15. 1.5 mL Eppendorf tubes.
16. Parafilm.
17. One Shot™ TOP10 Chemically Competent *Escherichia coli* cells (Invitrogen, Carlsbad, CA, USA).

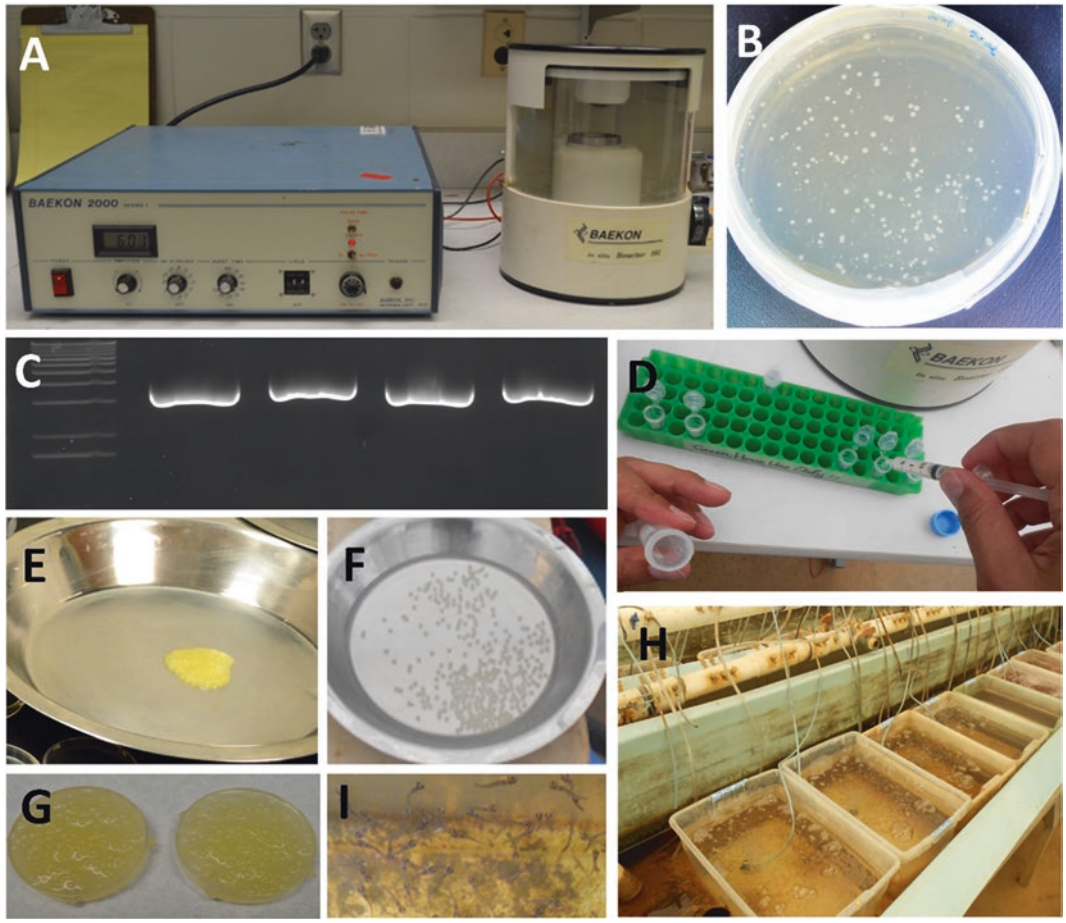


Fig. 1 Gene editing of channel catfish, *Ictalurus punctatus*, via double electroporation of sperm and one-celled embryos. (**a**) Baekon 2000 macromolecule transfer system (Baekon, Inc., Saratoga, CA). (**b**) LB agar plate with single colonies indicates successful transformation of competent cells. (**c**) Inspection of plasmid DNA quality in 1% agarose gel reveals a single band of the expected size for each ZFN plasmid (two ZFN pairs are shown). (**d**) Addition of sperm to the plasmid DNA solution before sperm electroporation. (**e**) Transfer of 150–200 eggs to a greased spawning pan for fertilization. (**f**) Fertilized channel catfish eggs are allowed to harden in the spawning pan. (**g**) Incubation of fertilized eggs with DNA solution 10 min before embryo electroporation. (**h**) Rearing of channel catfish embryos in plastic tubs containing 10 L of Holtfreter's solution with continuous aeration. (**i**) Channel catfish fry swim up after consuming their yolk reserves. Photos by Ahmed Elaswad

18. S.O.C. Medium: 2% Tryptone, 0.5% yeast extract, 10 mM NaCl, 2.5 mM KCl, 10 mM MgCl2, 10 mM $MgSO_4$, 20 mM glucose (provided with competent cells). Alternatively, LB broth without antibiotic can be used.
19. IsoPure™ Plasmid Maxi II Prep Kit (Denville Scientific Inc., Holliston, MA, USA).
20. Holtfreter's solution: 59 mM NaCl, 0.67 mM KCl, 2.4 mM $NaHCO_3$, 0.76 mM $CaCl_2$, 1.67 mM $MgSO_4$ [40] with 10 ppm doxycycline hyclate (*see* **Note 4**).

21. Magnesium sulfate.
22. Sodium chloride.
23. Sodium bicarbonate.
24. Potassium chloride.
25. Calcium chloride (32% solution).
26. Kanamycin stock solution: 50 mg/mL in deionized water (*see* **Note 5**).
27. Ampicillin stock solution: 100 mg/mL in deionized water.
28. 500 mL Centrifuge tubes.
29. 50 mL Falcon tubes.

3 Methods

3.1 Propagation of ZFN Plasmid Constructs

This protocol is optimized for One Shot™ TOP10 Chemically Competent *E. coli* (Invitrogen); however, competent cells from other suppliers can be used according to the manufacturer's instructions. Perform all steps involving competent cells or transformed *E. coli* cells aseptically.

1. Dilute the plasmid constructs (left and right ZFN plasmids) to a final concentration of 50 ng/μL using ultrapure water or TE buffer.
2. On ice, thaw a 50 μL vial of competent cells. Add 1–2 μL of ZFN plasmid construct solution (50–100 ng) to 25 μL of *E. coli* competent cells in a 1.5 mL Eppendorf tube and mix by gentle tapping. Incubate on ice for 30 min. Store the remaining plasmids at −20 °C.
3. Statically, heat-shock the competent cells at 42 °C for 30 s and then put back on ice immediately for 2 min. Alternatively, plasmid constructs can be added to competent cells in a 0.2 Eppendorf PCR tube, incubated on ice for 30 min, heat-shocked in a thermal cycler for 30 s at 42 °C, and then incubated on ice for 2 min.
4. Aseptically, add 250 μL of S.O.C. medium that was prewarmed to room temperature to each tube. Alternatively, LB medium without antibiotic can be used.
5. Seal the tube caps with parafilm to prevent the loss of solution. Place the tubes in an Eppendorf rack and shake horizontally in a shaking incubator at 37 °C for 1 h at 225 rpm.
6. Spread two different volumes on separate labeled LB agar plates that were prewarmed at 37 °C for 30 min. For example, spread 50 μL and 100 μL of transformed competent cells on two separate LB agar plates with antibiotic (*see* **Note 6**).

Incubate the plates in an inverted position at 37 °C for 18–24 h.

7. Check the plates for bacterial growth. Single colonies (Fig. 1b) indicate successful transformation.
8. Transfer a single colony from each agar plate to a sterile 1.5 mL Eppendorf tube containing 500–700 μL LB broth with 50 μg/mL kanamycin. To transfer a single colony, hold a sterile 35 mm 10 μL tip from the wide end with a forceps, open the cover of the agar plate at 45°, and pick a single colony with the fine end of the tip. Place the tip in the Eppendorf tube with LB broth. The tip can be discarded after a few minutes or kept in the tube and discarded after the overnight incubation.
9. Seal the tubes with parafilm and incubate horizontally overnight at 37 °C in a shaker incubator at 225 rpm.
10. Transfer overnight culture of each plasmid separately to a 500 mL sterile glass bottle containing 250 mL LB broth with 50 μg/mL kanamycin. Incubate for 12–16 h at 37 °C in a shaker incubator at 225 rpm.
11. Extract the plasmid DNA using IsoPure™ Plasmid Maxi II Prep Kit (Denville Scientific Inc., Holliston, MA, USA) following the manufacturer's instructions, and elute the DNA in TE buffer.
12. Determine the concentration and purity of the extracted plasmid DNA with a spectrophotometer, and check the quality by loading a few microliters in a 1% agarose gel (Fig. 1c, *see* **Note 7**).

3.2 Preparation of ZFN Plasmids for Electroporation

3.2.1 Preparation of ZFN Plasmids for Sperm Electroporation

If the plasmid concentrations are 500 ng/μL or higher, dilute 50 μg of each ZFN plasmid in 2 mL 0.9% saline to reach a final concentration of 50 μg/mL for both (25 μg/mL each) and store at −20 °C until use (*see* **Note 8**). This amount should be adequate for sperm electroporation to fertilize one batch of eggs. If the plasmid concentrations are less than 500 ng/μL, concentrate the DNA as follows:

1. Add 1/10 volume of 3 M sodium acetate (NaOAc) of pH 5.2 or 5 M NH_4OAc, and 2–2.5 volumes of ice-cold 100% ethanol into DNA solution.
2. Mix and incubate at −20 °C for 1 h.
3. Centrifuge at full speed in a microcentrifuge for 15–20 min.
4. Discard the ethanol, wash the DNA pellet in 70% ethanol, and then air-dry the pellet.
5. Based on the initial concentration, dissolve the pellet in a smaller volume of ultrapure water or TE buffer to achieve a concentration that is 500 ng/μL or higher.

6. Prepare the working solution for sperm electroporation as mentioned above. Prepare the solution in a 5 mL Eppendorf tube or split into two 1.5 mL Eppendorf tubes.

3.2.2 Preparation of ZFN Plasmids for Embryo Electroporation

Dilute the plasmids in TE buffer to reach the same concentration as for sperm electroporation (25 μg/mL of each plasmid, 50 μg/mL for both combined). In this protocol, 3–5 mL of the plasmid solutions is sufficient for one electroporation batch of 150–200 channel catfish eggs. Prepare the plasmid solutions in a 5 mL Eppendorf tube or a 15 mL Falcon tube. Store at −20 °C until use.

3.3 Artificial Spawning of Channel Catfish

For a comprehensive protocol of artificial spawning of channel catfish, *see* Elaswad et al. [38, 41]. Briefly, females are induced to spawn with luteinizing hormone-releasing hormone analog (LHRHa) and eggs are stripped and in vitro fertilized with sperm in batches of 150–200 eggs. A more detailed procedure is described below:

1. Implant channel catfish females in the dorsal musculature with 85 μg LHRHa /kg body weight. Place the females in spawning bags at least 15 cm under water in a flow-through tank with continuous water flow and aeration. Check for ovulation 36 h after implantation.
2. Euthanize channel catfish males by a percussive blow to the head (*see* **Note 9**). Crush and macerate the testes into up to 10 mL 0.9% saline/gram of testes. Fresh sperm is preferred; however, sperm can be stored at 4 °C and used within 24 h.
3. Anesthetize the females that began to ovulate in 100 ppm tricaine methane sulfonate (MS-222) in which the pH was adjusted to 7.0 using sodium bicarbonate, then rinse the female in 0.9% saline, and dry it with a clean towel. Hand-strip the eggs in a 20 cm greased spawning pan.

3.4 Sperm Electroporation and In Vitro Fertilization

1. Thaw the plasmids for sperm electroporation and wait until it reaches room temperature.
2. Add the sperm solution into the tubes containing the plasmids at a ratio of 1–2 sperm drops/1 mL plasmid (Fig. 1d). Mix gently and incubate at room temperature for at least 5 min.
3. Using plastic plant labels, transfer 150–200 eggs to another greased spawning pan for fertilization (Fig. 1e).
4. Adjust the electroporation parameters to 6 kV, 2^7 pulses, 0.8-s burst, four cycles, and 160 μs [23].
5. To electroporate the sperm, pour the DNA/sperm solution into a 7 mL Petri dish, fill it with freshwater, and mix gently. Place the petri dish inside the electroporation chamber, close

the protective glass shield, and move the petri dish stand upward until the solution touches the electrode. Electroporate the sperm and use it immediately for fertilization (*see* **Note 10**).

6. To fertilize the eggs, pour the electroporated sperm solution on the eggs in the greased spawning pan. Mix gently by swirling the eggs. Add more freshwater to slightly cover the eggs, wait for 1 min, then add a few drops of non-electroporated sperm, and mix gently by swirling the eggs (*see* **Note 11**).
7. After a few minutes, add more freshwater and allow the eggs to harden and develop until it is time for embryo electroporation (Fig. 1f).

3.5 Embryo Electroporation

1. Pour freshwater from the spawning pan and transfer the fertilized eggs from the spawning pan to a 7 mL petri dish 10–15 min before the first cell division (*see* **Note 12**).
2. Add enough amount of the DNA solution to totally cover the embryos in the petri dish (3–5 mL is enough) and incubate at room temperature for at least 10 min (Fig. 1g).
3. Place the petri dish with fertilized eggs in the electroporator and perform electroporation with the same parameters as described in **step 4** in Subheading 3.4.

3.6 Embryo Care and Rearing

1. When electroporation is complete, place embryos in plastic tubs containing Holtfreter's solution with 10 ppm doxycycline and continuous gentle aeration. Usually up to a hundred embryos are incubated in 10 L tubs (Fig. 1h). Place tubs in a tank in which water height is less than the height of the tubs (Fig. 1h). This will reduce temperature fluctuations and ensure that all tubs have the same temperature.
2. Depending on the objective of the gene editing, include three replicate tubs with control embryos that were exposed to the same handling stress and electroporated with TE buffer only. Include also three control tubs in which embryos were not electroporated (*see* **Note 13**).
3. Inspect the embryos daily and remove any dead embryos (*see* **Note 14**). To remove dead embryos, reduce the volume of Holtfreter's solution to 1/3 and pick dead embryos with a forceps or a plastic pipette. Avoid rough handling of embryos. If not removed, dead embryos or egg shells will be susceptible to fungal growth.
4. Replace the old Holtfreter's solution with fresh solution daily after removal of dead embryos. Check the temperature of both solutions. Avoid wide variations in temperature. Usually, temperature differences of 1.5 °C or less are acceptable.

5. If fungus infection appears, treat the embryos statically with 35–40 ppm formalin for 30 min and then completely change the solution immediately. Prophylactic formalin treatment of fungus may also be applied beginning the third day after fertilization until 1 day before hatch (*see* **Note 15**).
6. To check embryonic development, randomly pick a few embryos in a petri dish with Holtfreter's solution and visualize under the microscope (*see* Saksena et al. [42] for more information on channel catfish early development). After hatching, remove egg shells to avoid water quality deterioration. Fry then can be reared in Holtfreter's solution without doxycycline.
7. Do not feed the fry in the first few days until they consume their yolk sac and swim up (Fig. 1i). Fry are then fed with artemia nauplii twice per day at a rate of 5 nauplii/mL.
8. Gradually replace Holtfreter's solution with water from the catfish-rearing system that they will be transferred to next. To accomplish this, replace one-third of Holtfreter's solution with water every 1–2 days. Avoid wide temperature changes between both solutions. When Holtfreter's solution is totally replaced with water, move the fry to aquaria with continuous water flow and aeration, and feed them a mixture of artemia nauplii and catfish starter feed (55% protein) for 3–4 days. Then feed the fry with catfish starter feed multiple times daily for a total of 25% of fry body weight per day.

3.7 Mutation Detection

1. Randomly select several electroporated 3–4-day-old embryos. Remove egg shells and yolk, and extract genomic DNA with any commercial kit or other protocol. Mutation detection can also be performed on fingerlings. In this case, tissue samples, e.g., fin or barbel clips, are collected for DNA extraction. Check genomic DNA for integrity with a 1% agarose gel, and for purity and concentration using a spectrophotometer. Include 3–5 full-sib control embryos (*see* **Note 16**).
2. Detect mutations using a mutation detection method that is suitable for the expected mutation such as loss of restriction enzyme cut site, surveyor mutation detection assay, heteroduplex mobility assay, or any other method (*see* Zischewski et al. [43] for more mutation detection methods and the advantages and disadvantages of each method). Here, we use surveyor mutation detection assay for mutation detection, and cloning and DNA sequencing to confirm and determine the type of mutation (*see* **Note 17**).
3. Design and order oligonucleotide primers that are flanking the expected mutation sites (the sites where DSB is expected to occur). For best results, each primer should be at least 150–200 bp away from the expected mutation sites.

4. Prepare the PCR reaction mix in a 0.2 mL PCR tube: up to 20 μL PCR grade water; 1× Expand HiFi PLUS reaction buffer with $MgCl_2$, 0.2 mM dNTP mix, 0.4 μM for each of the forward and reverse primers, 100–300 ng genomic DNA, and 1.25 units of Expand HiFi PLUS enzyme blend. Include three PCR reactions for control embryos.
5. Mix by gentle tapping followed by a brief centrifugation, and then place the tubes in a thermal cycler in the following PCR cycling conditions: initial denaturation at 94 °C for 3 min; 30 cycles of denaturation at 94 °C for 30 s, annealing at optimum temperature for 30 s, and extension at 72 °C for 1 min/kb; and final extension at 72 °C for 10 min.
6. Clone the PCR product in a proper vector, e.g., TOPO™ TA Cloning™ Kit for Sequencing (Invitrogen) following the manufacturer's instructions, and then transfer the vector to One Shot™ TOP10 Chemically Competent *E. coli*. Use 2 μL of the cloning reaction to transform the competent cells and continue as described in Subheading 3.1. Replace kanamycin with 100 ppm ampicillin when using TOPO™ TA Cloning™ Kit for Sequencing (Invitrogen).
7. Pick up 10–20 single colonies from the LB agar plate, culture overnight in LB broth with 100 ppm ampicillin, and then extract and send plasmid DNA for sequencing using vector primers or the same primers that have been used in the PCR. Include several reactions from the full-sib wild-type embryos.
8. To identify ZFN-induced mutations, align the DNA sequences from electroporated embryos to sequences from wild-type embryos using a multiple sequence alignment tool, such as T-COFFEE.
9. Select mutant fish and rear them for further phenotypic evaluation and breed to generate mutant lines with mutations of interest.

4 Notes

1. We use a Baekon 2000 macromolecule transfer system (Baekon, Inc., Saratoga, CA) which is no longer manufactured. Other electroporators that can provide the same electroporation conditions could be used.
2. LB medium with kanamycin can be stored at 4 °C for a few days.
3. For LB agar plates, add 1.2% agar to LB broth formula and autoclave for 20 min. Let it cool down to 55 °C, then add kanamycin to a final concentration of 50 μg/mL, and pour

10–15 mL into each 10 cm petri dish. Allow the agar to solidify for 20 min, seal with parafilm, place in a sealed plastic bag, and store the plates inverted at 4 °C for 2–4 weeks.

4. Holtfreter's solution can be prepared as follows:
   (a) Prepare Holtfreter's stock solution by dissolving 1050 g NaCl, 60 g $NaHCO_3$, and 15 g KCl in 10 L dechlorinated city water.
   (b) Prepare Mg_2SO_4 stock solution by dissolving 150 g of $MgSO_4$ in 500 mL distilled water.
   (c) Prepare doxycycline stock solution by dissolving 1.5 g doxycycline hyclate in 30 mL of distilled water, store in a dark glass bottle at 4 °C, and use within 1 week of preparation.
   (d) Prepare Holtfreter's working solution by mixing the following in 60 L of dechlorinated city water: 2 L Holtfreter's stock solution, 40 mL $MgSO_4$ stock solution, 20 mL $CaCl_2$ (32%) solution, and 12 mL doxycycline stock solution.
5. Store at −20 °C.
6. Spreading two different concentrations will ensure that one plate will have evenly spaced single colonies.
7. Concentrations of 500 ng/μL or higher are adequate for embryo electroporation while for sperm electroporation concentrations of 500 ng/μL or higher are required.
8. We have not tested the effects of plasmid DNA storage time on electroporation success; therefore, we recommend using freshly prepared solutions or solutions prepared within 24 h before electroporation.
9. Anesthesia may have a negative effect on sperm motility.
10. Freshwater activates the sperm. Therefore, electroporated sperm must be used immediately for fertilization, and all the procedures beginning with adding freshwater to the sperm until fertilization must be completed in less than 2 min.
11. Sperm electroporation did not negatively affect the motility or fertilization success in Japanese abalone (*Haliotis diversicolor* supertexta) [44]. However, there is no information on whether electroporation of channel catfish sperm had any impact on fertilization success. Therefore, electroporated sperm is allowed 1 min to fertilize the eggs. Then, non-electroporated sperm is added to ensure fertilization. Both eggs fertilized with electroporated sperm and non-electroporated sperm will be electroporated at around the first cell division.
12. The development of channel catfish embryos is temperature dependent with development faster at higher temperatures. Usually, the first cell division occurs about 90 min after fertilization at 26–27 °C [45].

13. Embryos in which electroporation was performed with TE buffer will be used as controls to determine the effects of electroporation of plasmid constructs on embryo survival and hatching. Non-electroporated embryos will be used as a control for egg and sperm quality, and fertilization success.
14. Dead embryos are white, opaque, and/or swollen due to loss of osmoregulatory control.
15. Channel catfish embryos hatch in 5–8 days depending on water temperature.
16. Genomic DNA from full-sib control embryos will be used as a control for natural genomic variations, such as single-nucleotide polymorphisms (SNPs), during mutation detection and DNA sequencing.
17. The protocol described above resulted in a mutation rate (percent of fry with detectable mutations) of 0–39% depending upon the design of the zinc-finger nuclease sets utilized [8]. The individuals containing the mutations possessed the gene edits in every tissue tested.

References

1. Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. Nat Biotechnol 23:967–973
2. Doyon Y, McCammon JM, Miller JC, Faraji F, Ngo C, Katibah GE, Amora R, Hocking TD, Zhang L, Rebar EJ (2008) Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. Nat Biotechnol 26:702–708
3. Miller JC, Holmes MC, Wang J, Guschin DY, Lee Y-L, Rupniewski I, Beausejour CM, Waite AJ, Wang NS, Kim KA (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. Nat Biotechnol 25:778–785
4. Meng X, Noyes MB, Zhu LJ, Lawson ND, Wolfe SA (2008) Targeted gene inactivation in zebrafish using engineered zinc finger nucleases. Nat Biotechnol 26:695–701
5. Ochiai H, Sakamoto N, Fujita K, Nishikawa M, Suzuki K-i, Matsuura S, Miyamoto T, Sakuma T, Shibata T, Yamamoto T (2012) Zinc-finger nuclease-mediated targeted insertion of reporter genes for quantitative imaging of gene expression in sea urchin embryos. Proc Natl Acad Sci U S A 109:10915–10920
6. Urnov FD, Miller JC, Ya-Li L, Beausejour CM (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435:646–651
7. Sebastiano V, Maeder ML, Angstman JF, Haddad B, Khayter C, Yeo DT, Goodwin MJ, Hawkins JS, Ramirez CL, Batista LF (2011) *In situ* genetic correction of the sickle cell anemia mutation in human induced pluripotent stem cells using engineered zinc finger nucleases. Stem Cells 29:1717–1726
8. Qin Z, Li Y, Su B, Cheng Q, Ye Z, Perera DA, Fobes M, Shang M, Dunham RA (2016) Editing of the luteinizing hormone gene to sterilize channel catfish, *Ictalurus punctatus*, using a modified zinc finger nuclease technology with electroporation. Mar Biotechnol 18:255–263
9. Dong Z, Ge J, Li K, Xu Z, Liang D, Li J, Li J, Jia W, Li Y, Dong X (2011) Heritable targeted inactivation of myostatin gene in yellow catfish (*Pelteobagrus fulvidraco*) using engineered zinc finger nucleases. PLoS One 6:e28897
10. Yano A, Nicol B, Jouanno E, Guiguen Y (2014) Heritable targeted inactivation of the rainbow trout (*Oncorhynchus mykiss*) master sex-determining gene using zinc-finger nucleases. Mar Biotechnol 16:243–250
11. Neumann E, Schaefer-Ridder M, Wang Y, Hofschneider P (1982) Gene transfer into mouse lyoma cells by electroporation in high electric fields. EMBO J 1:841–845
12. Dower WJ, Miller JF, Ragsdale CW (1988) High efficiency transformation of *E. coli* by

high voltage electroporation. Nucleic Acids Res 16:6127–6145

13. Holo H, Nes IF (1989) High-frequency transformation, by electroporation, of *Lactococcus lactis* subsp. *cremoris* grown with glycine in osmotically stabilized media. Appl Environ Microbiol 55:3119–3123
14. Becker DM, Guarente L (1991) High-efficiency transformation of yeast by electroporation. Methods Enzymol 194:182–187
15. Fromm ME, Taylor LP, Walbot V (1986) Stable transformation of maize after gene transfer by electroporation. Nature 319:791–793
16. D'Halluin K, Bonne E, Bossut M, De Beuckeleer M, Leemans J (1992) Transgenic maize plants by tissue electroporation. Plant Cell 4:1495–1505
17. Prausnitz MR, Bose VG, Langer R, Weaver JC (1993) Electroporation of mammalian skin: a mechanism to enhance transdermal drug delivery. Proc Natl Acad Sci U S A 90:10504–10508
18. Chu G, Hayakawa H, Berg P (1987) Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res 15:1311–1326
19. Yarmush ML, Golberg A, Serša G, Kotnik T, Miklavčič D (2014) Electroporation-based technologies for medicine: principles, applications, and challenges. Annu Rev Biomed Eng 16:295–320
20. Potter H (1988) Electroporation in biology: methods, applications, and instrumentation. Anal Biochem 174:361–373
21. Jordan CA, Neumann E, Sowers AE (1989) Electroporation and electrofusion in cell biology. Springer Science & Business Media, New York
22. Kotnik T, Frey W, Sack M, Meglič SH, Peterka M, Miklavčič D (2015) Electroporation-based applications in biotechnology. Trends Biotechnol 33:480–488
23. Powers DA, Hereford L, Cole T, Chen TT, Lin C, Kight K, Creech K, Dunham R (1991) Electroporation: a method for transferring genes into the gametes of zebrafish (*Brachydanio rerio*), channel catfish (*Ictalurus punctatus*), and common carp (*Cyprinus carpio*). Mol Mar Biol Biotechnol 1:301–308
24. Dunham RA, Warr GW, Nichols A, Duncan PL, Argue B, Middleton D, Kucuktas H (2002) Enhanced bacterial disease resistance of transgenic channel catfish *Ictalurus punctatus* possessing cecropin genes. Mar Biotechnol 4:338–344
25. Inoue K, Yamashita S, Hata J-i, Kabeno S, Asada S, Nagahisa E, Fujita T (1990) Electroporation as a new technique for producing transgenic fish. Cell Differ Dev 29:123–128
26. Sarmasik A, Warr G, Chen TT (2002) Production of transgenic medaka with increased resistance to bacterial pathogens. Mar Biotechnol 4:310–322
27. Buono R, Linser P (1992) Transient expression of RSVCAT in transgenic zebrafish made by electroporation. Mol Mar Biol Biotechnol 1:271–275
28. Lavitrano M, Busnelli M, Cerrito MG, Giovannoni R, Manzini S, Vargiolu A (2005) Sperm-mediated gene transfer. Reprod Fertil Dev 18:19–23
29. Lu J-K, Fu B-H, Wu J-L, Chen TT (2002) Production of transgenic silver sea bream (*Sparus sarba*) by different gene transfer methods. Mar Biotechnol 4:328–337
30. Sin F, Bartley A, Walker S, Sin I, Symonds J, Hawke L, Hopkins C (1993) Gene transfer in Chinook salmon (*Oncorhynchus tshawytscha*) by electroporating sperm in the presence of pRSV-lacZ DNA. Aquaculture 117:57–69
31. Zhong J, Wang Y, Zhu Z (2002) Introduction of the human lactoferrin gene into grass carp (*Ctenopharyngodon idellus*) to increase resistance against GCH virus. Aquaculture 214:93–101
32. Khoo H-W, Ang L-H, Lim H-B, Wong K-Y (1992) Sperm cells as vectors for introducing foreign DNA into zebrafish. Aquaculture 107:1–19
33. Tsai H, Tseng F, Liao I (1995) Electroporation of sperm to introduce foreign DNA into the genome of loach (*Misgurnus anguillicaudatus*). Can J Fish Aquat Sci 52:776–787
34. Dunham RA, Winn RN (2014) Production of transgenic fish. In: Pinkert CA (ed) Transgenic animal technology: a laboratory handbook. Elsevier BV, Amsterdam, pp 308–336
35. Foley JE, Maeder ML, Pearlberg J, Joung JK, Peterson RT, Yeh J-RJ (2009) Targeted mutagenesis in zebrafish using customized zinc finger nucleases. Nat Protoc 4:1855–1868
36. Treen N, Yoshida K, Sakuma T, Sasaki H, Kawai N, Yamamoto T, Sasakura Y (2014) Tissue-specific and ubiquitous gene knockouts by TALEN electroporation provide new approaches to investigating gene function in Ciona. Development 141:481–487
37. Liang X, Potter J, Kumar S, Zou Y, Quintanilla R, Sridharan M, Carte J, Chen W, Roark N, Ranganathan S (2015) Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. J Biotechnol 208:44–53
38. Elaswad A, Khalil K, Cline D, Page-McCaw P, Chen W, Michel M, Cone R, Dunham

R (2018) Microinjection of CRISPR/Cas9 protein into channel catfish, *Ictalurus punctatus,* embryos for gene editing. J Vis Exp 131:e56275. doi:10.3791/56275

39. Elaswad A, Dunham R (2017) Disease reduction in aquaculture with genetic and genomic technology: current and future approaches. Rev Aquacult 0:1–23. https://doi.org/10.1111/raq.12205
40. Armstrong J, Duhon S, Malacinski G (1989) Raising the axolotl in captivity. In: Armstrong JB, Malacinski GM (eds) Developmental biology of the axolotl. Oxford University Press, New York, pp 220–227
41. Elaswad A (2016) Genetic technologies for disease resistance research and enhancement in catfish. PhD Dissertation, Auburn University, Alabama, USA
42. Saksena VP, Riggs CD, Yamamoto K (1961) Early development of the channel catfish. Prog Fish Cult 23:156–161
43. Zischewski J, Fischer R, Bortesi L (2017) Detection of on-target and off-target mutations generated by CRISPR/Cas9 and other sequence-specific nucleases. Biotechnol Adv 35:95–104
44. Tsai H-J, Lai C-H, Yang H-S (1997) Sperm as a carrier to introduce an exogenous DNA fragment into the oocyte of Japanese abalone (*Haliotis divorsicolor suportexta*). Transgenic Res 6:85–95
45. Hayat M, Joyce CP, Townes TM, Chen TT, Powers DA, Dunham RA (1991) Survival and integration rate of channel catfish and common carp embryos microinjected with DNA at various developmental stages. Aquaculture 99:249–255

Chapter 16

Delivery of mtZFNs into Early Mouse Embryos

Beverly J. McCann, Andy Cox, Payam A. Gammage, James B. Stewart, Magdalena Zernicka-Goetz, and Michal Minczuk

Abstract

Mitochondrial diseases often result from mutations in the mitochondrial genome (mtDNA). In most cases, mutant mtDNA coexists with wild-type mtDNA, resulting in heteroplasmy. One potential future approach to treat heteroplasmic mtDNA diseases is the specific elimination of pathogenic mtDNA mutations, lowering the level of mutant mtDNA below pathogenic thresholds. Mitochondrially targeted zinc-finger nucleases (mtZFNs) have been demonstrated to specifically target and introduce double-strand breaks in mutant mtDNA, facilitating substantial shifts in heteroplasmy. One application of mtZFN technology, in the context of heteroplasmic mtDNA disease, is delivery into the heteroplasmic oocyte or early embryo to eliminate mutant mtDNA, preventing transmission of mitochondrial diseases through the germline. Here we describe a protocol for efficient production of mtZFN mRNA in vitro, and delivery of these into 0.5 dpc mouse embryos to elicit shifts of mtDNA heteroplasmy.

Key words Mitochondrial disease, mtZFN, Germline, In vitro transcription, Micromanipulation

1 Introduction

Mammalian mitochondria are double membrane-bound organelles originating from bacteria with fundamental roles in energy production, calcium homeostasis, cellular signaling, and apoptosis [1, 2]. Mitochondria have their own small, circular genome (mtDNA) encoding 13 polypeptides of the oxidative phosphorylation (OXPHOS) system as well as 22 tRNAs and 2 rRNAs necessary for the intramitochondrial protein synthesis [3]. In mammals, mtDNA copy number ranges from approximately 1000 copies in somatic cells to approximately 100,000 copies in oocytes [4]. Transmission of mtDNA in mammals occurs exclusively through the maternal line [5]. Mutations in mtDNA have been associated with human diseases, collectively known as mitochondrial diseases. It is believed that 1 in 5000 children develop a mitochondrial disease caused by mutations in the mitochondrial genome [6], and it is estimated that 1 in 200 women are carriers of pathogenic mtDNA mutations [7].

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_16,

Within one cell, mtDNA can be present as either a single mtDNA haplotype (homoplasmic) or a mixture of different mtDNA haplotypes, e.g., wild-type and pathogenic mutant mtDNA (heteroplasmic). Mutations in mtDNA can range from being phenotypically silent to causing devastating diseases that typically affect tissues and organs with high energy demand, such as the brain, heart, and muscle [8]. The phenotypic outcome of an individual harboring a heteroplasmic, deleterious mtDNA mutation depends on the ratio of mutant to wild-type mtDNA in a cell.

Treatment options for mitochondrial diseases are very limited [9], with existing approaches being mainly focused on maintaining general patient health and not being able to correct, eliminate, or compensate for deleterious mtDNA. The current options for preventing transmission of mitochondrial diseases caused by mtDNA mutations are limited to genetic counseling and preimplantation genetic diagnosis (PGD) [10]. However, novel approaches for removing mutant mtDNA, or for transferring mitochondria with wild-type mtDNA into cells, have emerged. Mitochondrial replacement techniques by spindle, pronuclear, or polar body transfer have been reported recently [11–14]. These approaches, by varied means, entail transfer of parental nuclear DNA (nDNA) from an oocyte containing mutant maternal mtDNA into an enucleated donor oocyte with wild-type mtDNA. However, since this technique requires combining genetic material from three individuals, it has raised ethical, safety, and medical concerns [15–17].

Another emerging approach involves the selective degradation of pathogenic mtDNA in a heteroplasmic population [9, 18]. Through altering the ratio of mutated versus wild-type mtDNA followed by restoration of mtDNA copy number, a shift in heteroplasmy can be achieved. In early work, mitochondrially targeted restriction endonucleases (mtREs) have been used to bind and cleave mutant mtDNA sequences with high efficiency and accuracy [19–23]. One example of a mtDNA point mutation that generates a unique restriction site (*Sma*I or *Xma*I) is m.8993T>G, associated with the NARP (neuropathy, ataxia, and retinitis pigmentosa) and MILS (maternally inherited Leigh syndrome) syndromes [24]. Unfortunately, this is the only disease-causing mutation to produce a unique restriction site within mtDNA known so far, triggering the field to explore the use of programmable nucleases to target mtDNA mutations. Zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs), which can be engineered to bind predetermined DNA sequences, have been targeted to mitochondria for the specific elimination of mitochondrial genomes carrying mutations responsible for mitochondrial diseases [25–30]. Mitochondrially targeted ZFNs (mtZFNs) consist of a Cys_2His_2 zinc-finger protein (ZFP) with DNA sequence specificity,

conjugated to a single moiety of the dimeric *Fok*I C-terminal catalytic domain, with additional mitochondrial targeting sequence (MTS) and nuclear export signal (NES) peptides ensuring exclusive mitochondrial localization [31, 32]. TALENs have also been combined with similar components directing them to mitochondria (mitoTALEN) and have been reported to produce similar shifts of mtDNA heteroplasmy in comparable cybrid cell models [33, 34]. Reddy et al. [35] have used mtREs and mitoTALENs for the first time in heteroplasmic NZB/BALBc mouse oocytes as a proof of principle to shift mtDNA heteroplasmy and prevent germline transmission. The heteroplasmic NZB/BALBc mouse model carries two different naturally occurring, nonpathogenic mtDNA haplotypes (NZB and BALBc) [36], of which the BALBc mtDNA contains a unique *Apa*LI site and, therefore, provides an in vivo model for studies of specific mtDNA elimination by use of mtREs and programmable nucleases [19, 33]. Furthermore, in the same paper, it has been shown that mitoTALENs can be used to reduce human mutated mtDNA in artificial mammalian oocytes, which were derived by fusion of patient cells with mouse oocytes, thus representing a potential therapeutic approach for preventing germline transmission of mitochondria diseases caused by mtDNA mutations [35].

Here, we report on methods to produce mtZFN mRNA in vitro and to introduce them into the cytoplasm of one-cell-stage mouse embryos by microinjection. Before starting the protocol, it is assumed that a mouse model bearing a mutation of interest and vectors containing mtZFN specific to this mutation have been acquired. Vectors containing the mtZFN construct should also include either a T2A or an IRES sequence, followed by a reporter gene, such as GFP or mCherry, allowing for monitoring expression of the injected transgene in embryos. Protocols for ZFP design are available [36–41] as well as protocols for the assembly of mtZFN [27]. As an example, we use a mouse model harboring the m.5024C>T point mutation in the *MT-TA* gene, coding for mitochondrial $tRNA^{Ala}$ [42]. General protocols for embryo handling can be found in great detail in Behringer et al. (2013), Manipulating the mouse embryo: a laboratory manual, Fourth Edition [43]. Expression of mtZFNs can be monitored using reporters under an inverted fluorescence microscope. Further downstream methods would include the analysis of heteroplasmy shift in the embryos (pyrosequencing or RFLP) and the determination of mtDNA copy number using qPCR analysis.

2 Materials

2.1 In Vitro Transcription mtZFN of mRNA

2.1.1 Materials

pIRESpuro3: Encodes the IRES sequence and GFP as a marker and is available from TaKaRa Bio (formerly Clontech).

pIRES-mCherry: A pIRESpuro3-based vector, where the GFP-coding sequence has been replaced by the mCherry-coding sequence.

mtZFN(+)-T2A-mCherry: A pcDNA3.1(−)-based vector, in which co-expression of mtZFN and mCherry is facilitated by the T2A sequence [44].

mtZFN(−)-T2A-GFP: A pcDNA3.1(−)-based vector, in which co-expression of mtZFN and GFP is facilitated by the T2A sequence [44].

*Xba*I restriction enzyme.

Formaldehyde, 37% v/v.

Ultrapure agarose.

Ethidium bromide.

10 kb RNA ladder.

RNA gel-loading buffer II.

Phenol:chloroform:isoamyl alcohol, in a v/v proportion of 25:24:1.

Chloroform:isoamyl alcohol in a v/v proportion of 24:1.

3 M NaAc.

2.1.2 Buffers, Solutions, and Kits

HiScribe T7 ARCA mRNA kit with tailing, available from New England Biolabs.

MEGAclear Transcription cleanup kit, available from Thermo Fisher Scientific.

10× MOPS: 200 mM MOPS, 50 mM sodium acetate, 10 mM EDTA, pH 7.0 adjusted with NaOH.

Northern/RNA running buffer: 1× MOPS, 1% v/v formaldehyde, 0.1 μg/mL EtBr.

Denaturing RNA gel: 1× MOPS, 1% formaldehyde, 1.2% agarose.

2.1.3 Equipment

PCR thermocycler.

RNA ScreenTape system, Agilent 2200 TapeStation.

UV-spectrophotometer.

Microwave oven.

Mini-agarose gel apparatus.

Gel-imaging system.

2.2 Microinjection of One-Cell Embryos

2.2.1 Materials

Pregnant mare serum gonadotropin, 5 IU PMSG, available from Intervet/MSD Animal Health.

Human chorionic gonadotropin, 5 IU hCG, available from Intervet/MSD Animal Health.

EmbryoMax® M2 Medium, w/Phenol Red, and Hyaluronidase, available from Merck.

M2 medium, suitable for mouse embryo culture.

KSOM, suitable for mouse embryo culture, produced according to [43].

Oil for embryo culture.

Phosphate-buffered saline, PBS.

Nuclease-free water.

Dumont #5 fine forceps.

Fine scissors.

35 mm Tissue culture dishes.

Cryotube vials.

Micropipette-loading tips.

Transfer and holding pipettes.

Injection capillaries.

Glass depression slide.

2.2.2 Equipment

Stereomicroscope.

Microinjection system.

Heating plate, set to 37 °C.

Humidified incubator with 5% CO_2 and 37 °C.

2.3 Verifying mtZFN Expression in Embryos

96-Well plate, round bottom.

Glass-bottom dish.

2.3.1 Materials

PBS.

Tween-20.

Triton X-100.

Bovine serum albumin, BSA.

2.3.2 Buffers and Solutions

PBS-T: 0.1% Tween v/v in PBS.

Triton X-100: 0.5% v/v in PBS.

PBS-T BSA: 3% BSA w/v in PBS-T.

3 Methods

3.1 In Vitro Transcription of mtZFN mRNA

Perform all work in this protocol in an RNase-free environment. Wear gloves when setting up RNA transcription reactions. Make sure to use nuclease-free tubes and reagents to avoid RNase contamination.

1. To linearize the plasmid, digest 15 μg of each pair of ZFN-T2A-GFP/mCherry plasmids using 50 U of *Xba*I in a total reaction volume of 200 μL for 1 h at 37 °C (*see* **Note 1**). Afterwards purify the linearized plasmid at room temperature by phenol-chloroform extraction by adding 200 μL phenol-chloroform-isoamyl alcohol, vortex for 30 s, and centrifuge for 1 min. Take the top phase and add 200 μL chloroform-isoamyl alcohol. Vortex for 30 s and centrifuge for 1 min. Take the top phase and add 2.5 volume of 100% (v/v) EtOH and 20 μL 3 M NaOAc. Mix and spin for 15 min. Remove the supernatant and wash the pellet with 1 mL 70% (v/v) EtOH. Centrifuge for 2 min, discard EtOH, and let the pellet air-dry. Dissolve the DNA pellet in 5 μL nuclease-free water. Measure the DNA concentration on a UV-spectrophotometer (*see* **Note 2**).
2. For in vitro transcription, thaw the necessary components from the HiScribe T7 ARCA mRNA kit and assemble a 20 μL reaction at room temperature in the following order: Pipette 10 μL 2× ARCA/NTP mix, and add 1 μg linearized template DNA (**step 1**) in a volume of 8 μL of nuclease-free water and 2 μL T7 RNA polymerase mix. Mix thoroughly and incubate at 37 °C for 30 min.
3. For DNase treatment, add 4 units of DNase I, mix well, and incubate at 37 °C for 15 min. Save 1 μL on ice or at −20 °C for later gel electrophoresis.
4. For poly(A)-tail synthesis, set up the poly(A) tailing reaction as follows: add 20 μL nuclease-free water to 20 μL of the in vitro transcription reaction, 5 μL 10× poly(A) polymerase reaction buffer, and 5 μL poly(A) polymerase. Mix thoroughly and incubate at 37 °C for 30 min. Save 1 μL on ice or at −20 °C for later analysis by gel electrophoresis.
5. For mRNA purification, use the MEGAclear Transcription cleanup kit according to the manufacturer's instructions. The only modification of the protocol is the elution step. Elute RNA from the filter by placing the filter cartridge into a new elution tube. Apply 30–35 μL of nuclease-free water to the center of the filter cartridge (*see* **Note 3**). Close the cap of the tube and incubate in a PCR thermocycler at 70 °C for 10 min. Recover eluted RNA by centrifuging for 1 min at room temperature (21,000 × *g*) (*see* **Note 4**).

6. Quantify mRNA using the RNA ScreenTape system, according to the manufacturer's instructions (*see* **Note 5**) or using a UV-spectrophotometer. Each reaction typically yields 15–20 μg of RNA.
7. To verify the integrity and size of mRNA, run the samples on an RNA gel. For two gels, mix 87 mL nuclease-free water with 1.2 g ultrapure agarose. Heat the mixture in a microwave oven to dissolve the agarose. In case of evaporation, add nuclease-free water to the original volume. Add 10 mL of 10× MOPS and 3 mL of 37% (v/v) formaldehyde under a hood and pour the gels immediately, as they will set quickly.
8. Mix 1 μL of the RNA samples from **steps 3**, **4**, and **5** with 1 μL gel-loading buffer II. Heat samples at 55 °C for 10 min to denature, place samples on ice immediately for 2 min, spin down briefly, and load them onto the gel. Run the gel at 80 V/cm for 2.5 h in a 4 °C cold room. Image the gel on a UV table (*see* **Note 6**) (*see* Fig. 1).

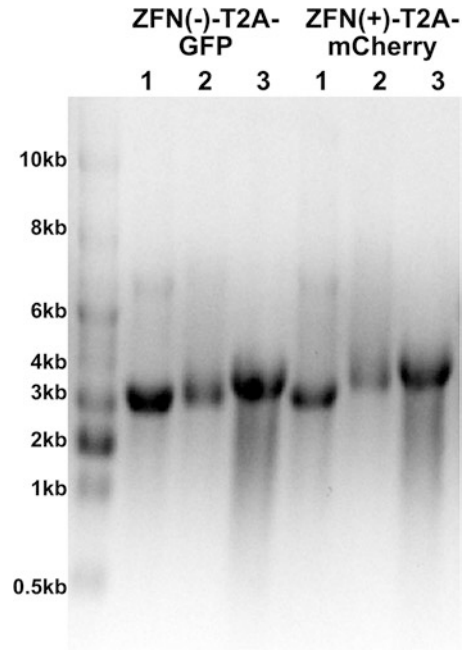


Fig. 1 Gel analysis of mRNA in vitro transcription reactions. 1 μL of mRNA from each of the mtZFN(–)-T2A-GFP or mtZFN(+)-T2A-mCherry constructs. mRNA samples have been loaded following in vitro transcription and DNase treatment (*1*), poly-(A) tailing reaction (*2*), and after cleanup and elution of the mRNA (*3*)

3.2 Microinjection of One-Cell Embryos

3.2.1 Superovulation

1. Inject 0.1 mL of PMSG solution (7.5 IU) intraperitoneal (i.p.) into each female mouse (4–6 weeks old, strain C57BL/6N $tRNA^{Ala}$ C5024T (41), at a light cycle of 7 am to 7 pm) at 4 pm to 5 pm, taking care to avoid the bladder and diaphragm. After 46–48 h, administer 0.1 mL of hCG solution (7.5 IU) i.p. and cross the female mice with C57BL/6N $tRNA^{Ala}$ m.5024C>T stud males overnight in single pairs (*see* **Note 7**).

3.2.2 Recovery of the Embryos

1. To dissect embryo donor females (*see* **Note 8**), the morning after mating took place, euthanize embryo donor females between 10 am and 1 pm by cervical dislocation (*see* **Note 9**). Dissect out the oviducts (42) and place them into a 1.5 mL Eppendorf tube filled with pre-warmed (37 °C) M2 medium for transfer from the animal facility to the laboratory (*see* **Note 10**).
2. To dissect the oviducts, place each oviduct into a drop of 200 μL warm EmbryoMax M2 medium supplemented with hyaluronidase in a 35 mm tissue culture dish. Under a stereomicroscope, tear the oviduct's ampulla (where the oviduct is the most swollen) with a pair of fine forceps and then remove the cumulus complex (*see* **Note 11**). By pipetting the zygotes up and down with a handling pipette, the cumulus cells will fall off the zygotes quickly (*see* **Note 12**).
3. Wash the zygotes through three drops of warm M2 medium (without hyaluronidase) and transfer them into a new dish with a 50 μL drop of fresh, warm M2 medium (*see* **Note 13**). Make sure to avoid the carryover of cumulus cells. Pool the collected embryos from both oviducts and count them. Keep them at 37 °C in the dark until further use (*see* **Note 14**).
4. If transport of embryos is required (e.g., the microinjection apparatus is in a different location than the lab), move the embryos appropriately. Fill a cryotube vial completely with pre-warmed M2 medium and flush out the embryos to the bottom of the tube. Let the embryos settle for a minute to sink to the bottom. Close the tube and wipe away overflowing M2 medium (*see* **Note 15**). During transportation, keep the tubes with the embryos at 37 °C in the dark. Upon arrival to the location of the micromanipulator, prepare a new dish with 50 μL drops of warm M2 medium, cover the drops with mineral oil, and place the dish on a 37 °C warming plate. Take out the cryotube vial and let it stand for 1 min, so the embryos can settle to the bottom. Open the vial and tip the content into a new 35 mm dish. Under the stereoscope, locate the embryos and transfer them into one of the drops of M2 medium. Rinse the tube with 500 μL of warm M2 to wash out any residual embryos (*see* **Note 16**). Let the embryos recover for an hour on the heating block in the dark before performing microinjections.

3.2.3 Microinjection

1. To prepare mRNA for microinjection, mix both mtZFN mRNA and empty vector (EV) mRNA with a final concentration of 100 ng/μL in a total volume of 5 μL 0.1× PBS. Place the tube of mRNA into a 0.5 mL microfuge tube and centrifuge for 10 min at 4 °C just before injection (*see* **Note 17**).
2. To prepare the microinjection chamber, place a 20 μL drop of warm M2 medium in the middle of a depression slide. Cover the medium with a layer of mineral oil to prevent evaporation. Under a stereomicroscope, transfer the first group of embryos (20–30 embryos) to the upper part of the M2 medium drop within the injection chamber by mouth-pipetting. Place the slide into position on the injection microscope. Place the microcapillary for holding zygotes (*see* **Note 18**), filled with fresh, pre-warmed M2 into the connector piece, and set into position in the center of the injection chamber.
3. To prepare the microinjection needle, use a microloader pipette tip to load 1–2 μL of the mtZFN mRNA into the tip of an injection needle, either by eye or under a stereomicroscope (*see* **Note 19**). After loading the injection needle, quickly insert it into the needle holder and place it in the medium, setting into position as above (*see* **Note 20**).
4. For microinjection, use the holding pipette to fix the embryo and place it into the center of the M2 medium drop. Penetrate the zona pellucida with the injection capillary and push forwards into the cytoplasm. Apply injection pressure for 0.5–2 s until visible swelling in the cytoplasm occurs. Carefully withdraw the capillary, place the injected embryo in the lower part of the M2 medium, and proceed until all embryos of one group are injected (*see* **Note 21**). Transfer each group of injected embryos into a single 50 μL drop of M2 medium (covered with mineral oil) in a 35 mm tissue culture dish and incubate the dish at 37 °C on a warming plate.
5. After completion of the microinjections, collect the surviving embryos from the storage dish and wash the embryos, by pipetting them up and down through three 50 μL drops of KSOM (*see* **Note 22**), and place them in 20 μL drops of KSOM under mineral oil in a 37 °C incubator in 5% CO_2 atmosphere.

3.3 Monitoring mtZFN Expression in Embryos

1. For imaging, rinse embryos 18–24 h after injections in M2 medium and place them in 2 μL drops of M2 on a glass-bottom dish, covered with mineral oil.
2. Using an inverted fluorescence microscope, analyze the expression of the reporter genes GFP and mCherry (*see* **Note 23**) (Fig. 2).
3. If the embryos will be used for downstream applications after imaging, collect the embryos and wash them by pipetting up

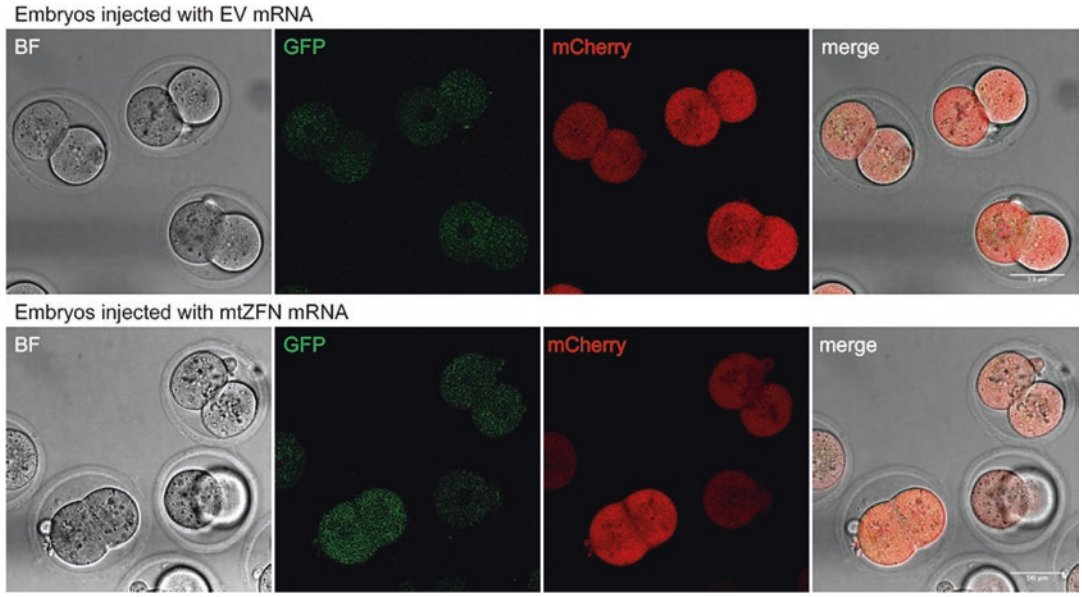


Fig. 2 Two-cell-stage embryos expressing mRNA-encoding mCherry and GFP controls generated by transcription of empty vectors (EV) (top) or expression of mRNA-encoding mtZFN(–)-T2A-GFP and mtZFN(+)-T2A-mCherry (bottom). Polyadenylated mRNA was injected into zygote and incubated for 16 h until the two-cell stage. *BF* bright field

and down through three 50 μL drops of KSOM, finally placing them in 20 μL drops of KSOM under mineral oil in a 37 °C incubator in 5% CO_2 atmosphere.

4 Notes

1. If a vector of different origin is used, digest it with a restriction enzyme that cuts downstream, as close as possible to the 3′ end of the mtZFN-coding region.
2. In order to verify complete digestion of the vector, run the samples on a 0.8% (w/v) agarose gel until you can see separation of the band containing the linearized vector. Excise the band from the gel and extract the DNA using a general gel extraction kit.
3. Instead of nuclease-free water, any RNA storage solution (e.g., Thermo Fisher Scientific, AM7000) can be used as well; however, we noticed that the use of such a solution can occasionally result in blockage of the injection needle. If using water instead, verify RNA integrity, especially upon long-term storage.
4. Elution in 100 μL is possible to achieve a higher yield (as described in the HiScribe T7 ARCA mRNA kit manufacturer's instructions). However, eluting in a lower volume allows for skipping the ethanol precipitation step and, therefore,

saves time while providing sufficient quantities of mRNA for the injections.

5. The RNA Tapestation is a useful and quick way to assess the concentration and quality of the produced RNA. However, discrepancies in RNA concentration values between the RNA Tapestation and the values obtained by analyzing denaturing RNA gel and/or UV-spectrophotometry have been observed.
6. The gel should be helpful in assessing the integrity of the RNA, as well as the successful addition of a poly(A) tail, and any potential losses of RNA during each step.
7. Depending on the age and genetic background of the female mice, as well as the dose and timing of PMSG and hCG injections in relation to the light cycle, the yield of embryos from superovulated mice can vary substantially. Further adjustments may be required for optimal performance. The male should have prior mating experience, and not be older than 20 weeks; otherwise the mating success rate might be low.
8. It is important to know the heteroplasmy level from the mother in order to draw conclusions about the spread in heteroplasmy of the embryos. It is necessary to keep the oviducts from each mother, and later the embryos, separated from each other.
9. The ideal time point for recovery is around 16 h post-hCG injection.
10. Always try and keep the embryos in a habitat as close to their natural environment as possible, i.e., a dark and warm (37 °C) place.
11. If there is excessive tissue from the ovary or the uterus on the oviduct, especially fat cells, remove them first in a separate dish with M2 media before placing them into the M2 drop with hyaluronidase.
12. Do not leave the embryos in M2 medium with hyaluronidase for more than 3–4 min, as this reduces the quality of the embryos. If some embryos still have cumulus cells attached to them, move the other embryos first in M2 and then come back to these.
13. Transferring the embryos into fresh M2 medium will dilute the hyaluronidase carryover. As hyaluronidase can be damaging, it is important to wash thoroughly.
14. A screw cap from a Duran bottle is big enough to cover a 35 mm tissue culture dish. This shields the embryos from light.
15. There shouldn't be any air bubbles in the cryotube, to minimize the possibility for the embryo not to be immersed in the medium.
16. It is expected that not all the embryos will be recovered. Some embryos may not survive the transport, and these need to be discarded.

17. Avoid contamination of the microinjection samples with dust particles, which might block the injection capillaries. Centrifugation of the mRNA just before injecting can help eliminating bigger particles.
18. Holding pipettes are hand-pulled, and the tip is flame polished over a Bunsen flame to a diameter of 80–100 μm.
19. Make sure that there are no air bubbles in the tip of the needle.
20. If the needle is not immersed in media for a period of time, RNA may dry at the very tip, clogging the needle.
21. Split embryos from one mother into a group of embryos injected with mtZFN mRNA and the other group with embryos injected with empty vector mRNA as control. Make sure to have at least ten embryos in each group.
22. KSOM is very sensitive to pH and salt changes, so it is important not to carry over any M2 medium into the KSOM drops.
23. In some cases, when direct detection of the expressed mtZFN is required, immunostaining can be performed. Therefore at 18–24 h after injection, the surviving embryos can be fixed for immunostaining. Prepare a 96-well round-bottom plate by incubating each well that will be used with 50 μL of 3% (w/v) BSA in PBS-T for 20 min, which prevents embryos from getting stuck at the bottom of the well. Remove the BSA. Transfer the embryos from M2 medium to the first row of the 96-well plate in 50 μL PBS-T (0.1% Tween in PBS). From there, transfer the embryos into 4% PFA in the next row and fix them for 15 min at room temperature. Move the embryos into PBS-T and wash 2× 5 min on a rocker, transferring them into a new well with PBS-T for each washing step. Permeabilize the embryos for 20 min in 0.5% (v/v) Triton in PBS in the next row. Wash 3× 5 min in PBS-T. Block for 1 h in PBS-T + 3% (w/v) BSA at room temperature on a rocker. Incubate with primary antibodies in PBS-T + 3% (w/v) BSA overnight at 4 °C. Wash 2× 10 min in PBS-T on the rocker. Incubate with secondary antibodies/DAPI in PBS-T + 3% (w/v) BSA at room temperature for 1 h. Wash 2× 10 min in PBS-T. For best results, image immediately; otherwise keep the embryos at 4 °C . For imaging, rinse the embryos in M2, and image them in drops of M2 on a glass-bottomed dish using an inverted fluorescent microscope.

Acknowledgments

This work was supported by the Medical Research Council, UK (MC_U105697135). We would like to thank Dr. Carlo Viscomi for help and stimulating discussions during the course of this work and the personnel at Phenomics animal care facilities for skillful technical assistance.

References

1. Dyall SD, Brown MT, Johnson PJ (2004) Ancient invasions: from endosymbionts to organelles. Science 304:253–257
2. Weinberg SE, Chandel NS (2015) Targeting mitochondria metabolism for cancer therapy. Nat Chem Biol 11:9–15
3. Anderson S, Bankier AT, Barrell BG, de Bruijn MH, Coulson AR, Drouin J, Eperon IC, Nierlich DP, Roe BA, Sanger F et al (1981) Sequence and organization of the human mitochondrial genome. Nature 290:457–465
4. Shoubridge EA, Wai T (2007) Mitochondrial DNA and the mammalian oocyte. Curr Top Dev Biol 77:87–111
5. Giles RE, Blanc H, Cann HM, Wallace DC (1980) Maternal inheritance of human mitochondrial DNA. Proc Natl Acad Sci U S A 77:6715–6719
6. Schaefer AM, McFarland R, Blakely EL, He L, Whittaker RG, Taylor RW, Chinnery PF, Turnbull DM (2008) Prevalence of mitochondrial DNA disease in adults. Ann Neurol 63:35–39
7. Elliott HR, Samuels DC, Eden JA, Relton CL, Chinnery PF (2008) Pathogenic mitochondrial DNA mutations are common in the general population. Am J Hum Genet 83:254–260
8. Taylor RW, Turnbull DM (2005) Mitochondrial DNA mutations in human disease. Nat Rev Genet 6:389–402
9. Viscomi C, Bottani E, Zeviani M (2015) Emerging concepts in the therapy of mitochondrial disease. Biochim Biophys Acta 1847:544–557
10. Tuppen HA, Blakely EL, Turnbull DM, Taylor RW (2010) Mitochondrial DNA mutations and human disease. Biochim Biophys Acta 1797:113–128
11. Craven L, Tuppen HA, Greggains GD, Harbottle SJ, Murphy JL, Cree LM, Murdoch AP, Chinnery PF, Taylor RW, Lightowlers RN et al (2010) Pronuclear transfer in human embryos to prevent transmission of mitochondrial DNA disease. Nature 465:82–85
12. Paull D, Emmanuele V, Weiss KA, Treff N, Stewart L, Hua H, Zimmer M, Kahler DJ, Goland RS, Noggle SA et al (2013) Nuclear genome transfer in human oocytes eliminates mitochondrial DNA variants. Nature 493:632–637
13. Tachibana M, Amato P, Sparman M, Woodward J, Sanchis DM, Ma H, Gutierrez NM, Tippner-Hedges R, Kang E, Lee HS et al (2013) Towards germline gene therapy of inherited mitochondrial diseases. Nature 493:627–631
14. Wang T, Sha H, Ji D, Zhang HL, Chen D, Cao Y, Zhu J (2014) Polar body genome transfer for preventing the transmission of inherited mitochondrial diseases. Cell 157:1591–1604
15. Hayden EC (2013) Regulators weigh benefits of 'three-parent' fertilization. Nature 502:284–285
16. Hyslop LA, Blakeley P, Craven L, Richardson J, Fogarty NM, Fragouli E, Lamb M, Wamaitha SE, Prathalingam N, Zhang Q et al (2016) Towards clinical application of pronuclear transfer to prevent mitochondrial DNA disease. Nature 534:383–386
17. Vogel G (2014) Assisted reproduction. FDA considers trials of 'three-parent embryos. Science 343:827–828
18. Patananan AN, Wu TH, Chiou PY, Teitell MA (2016) Modifying the mitochondrial genome. Cell Metab 23:785–796
19. Bacman SR, Williams SL, Duan D, Moraes CT (2012) Manipulation of mtDNA heteroplasmy in all striated muscles of newborn mice by AAV9-mediated delivery of a mitochondria-targeted restriction endonuclease. Gene Ther 19:1101–1106
20. Bayona-Bafaluy MP, Blits B, Battersby BJ, Shoubridge EA, Moraes CT (2005) Rapid directional shift of mitochondrial DNA heteroplasmy in animal tissues by a mitochondrially targeted restriction endonuclease. Proc Natl Acad Sci U S A 102:14392–14397
21. Srivastava S, Moraes CT (2001) Manipulating mitochondrial DNA heteroplasmy by a mitochondrially targeted restriction endonuclease. Hum Mol Genet 10:3093–3099
22. Srivastava S, Moraes CT (2005) Double-strand breaks of mouse muscle mtDNA promote large deletions similar to multiple mtDNA deletions in humans. Hum Mol Genet 14:893–902
23. Tanaka M, Borgeld HJ, Zhang J, Muramatsu S, Gong JS, Yoneda M, Maruyama W, Naoi M, Ibi T, Sahashi K et al (2002) Gene therapy for mitochondrial disease by delivering restriction endonuclease SmaI into mitochondria. J Biomed Sci 9:534–541
24. Thorburn DR, Rahman S (1993) Mitochondrial DNA-associated Leigh syndrome and NARP. In: Pagon RA, Adam MP, Ardinger HH, Wallace SE, Amemiya A, LJH B, Bird TD, Fong CT, Mefford HC, RJH S et al (eds) GeneReviews®. University of Washington, Seattle, WA
25. Bacman SR, Williams SL, Pinto M, Peralta S, Moraes CT (2013) Specific elimination of mutant mitochondrial genomes in patient-

derived cells by mitoTALENs. Nat Med 19:1111–1113

26. Gammage PA, Gaude E, Van Haute L, Rebelo-Guiomar P, Jackson CB, Rorbach J, Pekalski ML, Robinson AJ, Charpentier M, Concordet JP, Frezza C, Minczuk M (2016) Near complete elimination of mutant mtDNA by iterative or dynamic dose-controlled treatment with mtZFNs. Nucleic Acids Res 19:7804–7816
27. Gammage PA, Rorbach J, Vincent AI, Rebar EJ, Minczuk M (2014) Mitochondrially targeted ZFNs for selective degradation of pathogenic mitochondrial genomes bearing large-scale deletions or point mutations. EMBO Mol Med 6:458–466
28. Gammage PA, Van Haute L, Minczuk M (2016) Engineered mtzfns for manipulation of human mitochondrial DNA heteroplasmy. Methods Mol Biol 1351:145–162
29. Minczuk M, Kolasinska-Zwierz P, Murphy MP, Papworth MA (2010) Construction and testing of engineered zinc-finger proteins for sequence-specific modification of mtDNA. Nat Protoc 5:342–356
30. Minczuk M, Papworth MA, Miller JC, Murphy MP, Klug A (2008) Development of a single-chain, quasi-dimeric zinc-finger nuclease for the selective degradation of mutated human mitochondrial DNA. Nucleic Acids Res 36:3926–3938
31. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. Proc Natl Acad Sci U S A 93:1156–1160
32. Minczuk M, Papworth MA, Kolasinska P, Murphy MP, Klug A (2006) Sequence-specific modification of mitochondrial DNA using a chimeric zinc finger methylase. Proc Natl Acad Sci U S A 103:19689–19694
33. Bacman SR, Williams SL, Garcia S, Moraes CT (2010) Organ-specific shifts in mtDNA heteroplasmy following systemic delivery of a mitochondria-targeted restriction endonuclease. Gene Ther 17:713–720
34. Hashimoto M, Bacman SR, Peralta S, Falk MJ, Chomyn A, Chan DC, Williams SL, Moraes CT (2015) MitoTALEN: a general approach to reduce mutant mtdna loads and restore oxidative phosphorylation function in mitochondrial diseases. Mol Ther 23:1592–1599
35. Reddy P, Ocampo A, Suzuki K, Luo J, Bacman SR, Williams SL, Sugawara A, Okamura D, Tsunekawa Y, Wu J et al (2015) Selective elimination of mitochondrial mutations in the germline by genome editing. Cell 161:459–469
36. Jenuth JP, Peterson AC, Fu K, Shoubridge EA (1996) Random genetic drift in the female germline explains the rapid segregation of mammalian mitochondrial DNA. Nat Genet 14:146–151
37. Cox DB, Platt RJ, Zhang F (2015) Therapeutic genome editing: prospects and challenges. Nat Med 21:121–131
38. Carroll D, Morton JJ, Beumer KJ, Segal DJ (2006) Design, construction and in vitro testing of zinc finger nucleases. Nat Protoc 1:1329–1341
39. Maeder ML, Thibodeau-Beganny S, Osiak A, Wright DA, Anthony RM, Eichtinger M, Jiang T, Foley JE, Winfrey RJ, Townsend JA, Unger-Wallace E, Sander JD, Muller-Lerch F, Fu F, Pearlberg J, Gobel C, Dassie JP, Pruett-Miller SM, Porteus MH, Sgroi DC, Iafrate AJ, Dobbs D, McCray PB, Cathomen T Jr, Voytas DF, Joung JK (2008) Rapid "open-source" engineering of customized zinc-finger nucleases for highly efficient gene modification. Mol Cell 31:294–301
40. Kim S, Lee MJ, Kim H, Kang M, Kim JS (2011) Preassembled zinc-finger arrays for rapid construction of ZFNs. Nat Methods 8:7
41. Sander JD, Dahlborg EJ, Goodwin MJ, Cade L, Zhang F, Cifuentes D, Curtin SJ, Blackburn JS, Thibodeau-Beganny S, Qi Y, Pierick CJ, Hoffman E, Maeder ML, Khayter C, Reyon D, Dobbs D, Langenau DM, Stupar RM, Giraldez AJ, Voytas DF, Peterson RT, Yeh JR, Joung JK (2011) Selection-free zinc-finger-nuclease engineering by context-dependent assembly (CoDA). Nat Methods 8:67–69
42. Kauppila JH et al (2016) A phenotype-driven approach to generate mouse models with pathogenic mtDNA mutations causing mitochondrial disease. Cell Rep 16:2980–2990
43. Behringer R, Gerstenstein M, Vintersten Nagy K, Nagy A (2013) Manipulating the mouse embryo: a laboratory manual, 4th edn. Cold Spring Harbor Laboratory Press, New York
44. Donnelly ML, Hughes LE, Luke G, Mendoza H, ten Dam E, Gani D, Ryan MD (2001) The 'cleavage' activities of foot-and-mouth disease virus 2A site directed mutants and naturally occurring '2A-like' sequences. J Gen Virol 82:1027–1041

Chapter 17

Stereotaxic Surgery and Viral Delivery of Zinc-Finger Epigenetic Editing Tools in Rodent Brain

Peter J. Hamilton, Carissa J. Lim, Eric J. Nestler, and Elizabeth A. Heller

Abstract

Delivery of engineered zinc-finger proteins (ZFPs) for targeted epigenetic remodeling in rodent brain can be facilitated by the use of viral vector-mediated gene transfer coupled with stereotaxic surgery techniques. Here we describe the surgical protocol utilized by our group which is optimized for herpes simplex virus (HSV) delivery into mouse brain. The protocol outlined herein could be applied for delivery of adeno-associated viruses (AAV) or lentiviruses in both mice and rats. This method allows for the viral expression of engineered DNA-binding factors, particularly engineered ZFPs, and subsequent epigenome editing in rodent brain with excellent spatiotemporal control. Nearly any brain region of interest can be targeted in rodents at every stage of postnatal life. Owing to the versatility, reproducibility, and utility of this technique, it is an important method for any laboratory interested in studying the cellular, circuit, and behavioral consequences of in vivo neuroepigenetic editing with synthetic ZFP constructs.

Key words Viral mediated gene transfer, Zinc-finger proteins, Neuroepigenome editing, Stereotaxic surgery, Rodent brain

1 Introduction

Stereotaxic surgery is a powerful method used to manipulate the brain of living animals. This technique allows researchers to consistently and accurately target deep structures of the rodent brain through the use of a stereotaxic brain atlas, which provides the coordinates of a given brain area relative to Bregma, an anatomical landmark on the rodent's skull. Stereotaxic coordinates for rodent brain regions of interest can be determined from *The mouse brain in stereotaxic coordinates* [1] and *The rat brain in stereotaxic coordinates* [2]. Facilitated through the use of a stereotaxic instrument, one can perform this surgery on large numbers of anesthetized animals to reliably and accurately access structures within the rodent brain.

Combining this approach with viral mediated gene transfer, which has been widely used as a means to introduce transgenes to

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_17,

intact brain tissue [3], researchers have been successful in delivering engineered ZFP neuroepigenome editing tools to deposit gene-locus-specific modifications in vivo to alter neural function and animal behaviors [4–6]. ZFP epigenetic editing tools, which can exogenously introduce chromatin modifications at a single genomic locus within neurons or even a single type of neuron in an injected brain region [4–8], are necessary to establish the causal relevance of such mechanisms to gene expression and neural function. Given the fact that regulation of epigenetic landscapes is central to neuropsychiatric health and disease [9, 10], it is crucial to combine epigenetic editing techniques with in vivo inquiry in the brains of awake and behaving animals. The technique of viral expression of ZFP epigenetic editing tools in rodent brain using stereotaxic surgery techniques facilitates the exploration of the causal impact of the targeted chromatin modifications in these neurobiological contexts.

2 Materials

2.1 Reagents

1. Ketamine and xylazine
2. 70% Ethanol
3. 100% Acetone
4. 10% Bleach solution
5. Alcohol prep wipes
6. Sterile ocular lubricant
7. Sterile PBS
8. Sterile normal saline
9. Purified virus (HSV, AAV, lentivirus)
10. Betadine antiseptic
11. Bupivacaine HCl local anesthetic

2.2 Instruments and Materials

1. Dual-small-animal stereotaxic instrument (such as Kopf Model 902)
2. Fine Science Surgical Tools, including but not necessarily limited to scalpel, scissors, forceps
3. Laboratory scale
4. Bead sterilizer
5. Electric hair shaver
6. Sterile tip cotton swabs
7. Biohazard bags
8. Low-binding, 0.65 mL microcentrifuge tube

9. Needles and syringes for IP injection of anesthetics and analgesics
10. Absorbent lab bench diapers
11. Handheld dental drill and 0.6 mm burr
12. Hamilton syringes (5 μL Catalog #84851) with Hamilton small-gauge RN needles (33-gauge Catalog #7762-06)
13. Tissue adhesive, surgical clips, or surgical sutures
14. Temperature-regulated heating pads and/or heat lamp

3 Methods

3.1 Stereotaxic Surgery

1. Position the stereotaxic instrument (*see* **Note 1**) under a heat lamp. Make sure that the surgical area is cleaned with 70% ethanol and surgical instruments are cleaned and properly sterilized. We find that a bead sterilizer works well for this purpose. Cover the surgical area with absorbent lab bench diapers. All procedures should be performed in accordance with your institution's biosafety and animal use guidelines.
2. Place Hamilton syringes in arms of the stereotaxic instrument and clear any blockages by drawing and expelling 100% acetone five times. Subsequently draw and expel sterile PBS five times to remove any residual acetone. Draw the maximum volume of sterile PBS into the Hamilton syringe, taking care to include no bubbles. Swing the stereotaxic arms to move the Hamilton syringes out of the way of the workspace in the center of the instrument.
3. Anesthetize animals with a ketamine/xylazine mixture (100 mg/kg ketamine and 5 mg/kg xylazine in sterile normal saline) delivered via intraperitoneal injection. The animal should reach surgical anesthesia within 5–10 min, and should not respond to a light pinch to the hind paw (*see* **Notes 2** and **3**).
4. Cover the anesthetized animal's eyes with sterile ocular lubricant to keep them moist during the surgery.
5. Shave the fur off of the top of the animal's skull and clean the surface of the skin with alcohol prep wipes. Apply Betadine antiseptic using sterile tip cotton swabs.
6. Place the animal in the stereotaxic instrument. To do so, carefully place one ear bar in the ear canal, secure the bar, and hold the animal in place as the other ear bar is placed and secured. The animal should not be able to move laterally. Next, secure the mouth in the incisor adapter of the stereotaxic instrument, taking care that the tongue is not pinched in the adapter or blocking the airway. The nose clamp can be gently tightened to firmly secure the animal's head in position

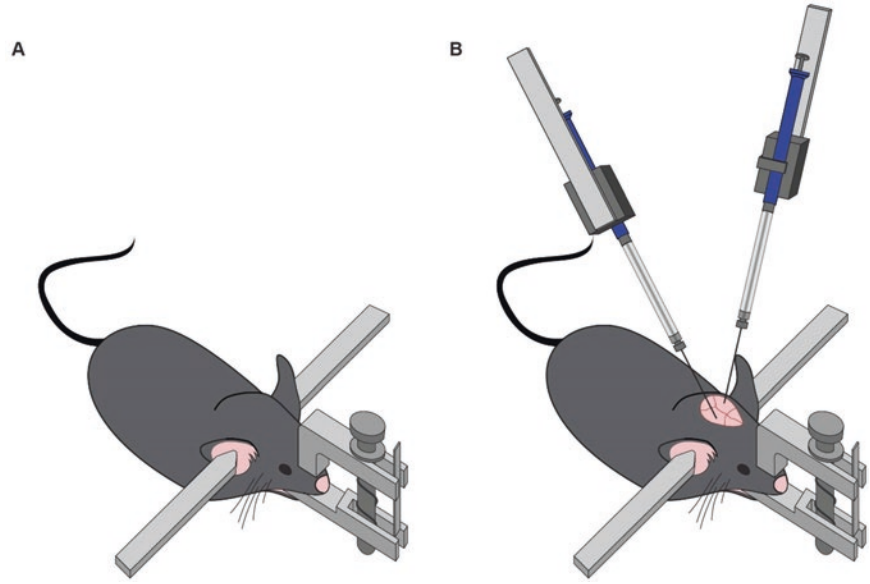


Fig. 1 Correct placement of rodent's head within stereotaxic instrument and surgical procedure for viral delivery. (**a**) A cartoon depicting the fixation of animal's head within the stereotaxic instrument. The ear bars are securely in place, preventing lateral movement of the skull. The incisor adapter restricts vertical movement, with the nose clamp gently tightened into place. (**b**) Upon surgically exposing the stereotaxic landmarks on the skull, the stereotaxic coordinates are measured relative to Bregma. Hamilton syringes are used to deliver the viral solution to desired regions within the animal's brain via small burr holes in the animal's skull

(*see* **Notes 4** and **5**). Visually inspect the head and make adjustments to the pitch of the incisor adapter to make sure that the head is level (Fig. 1a).

7. Make a midline incision to the top of the animal's skull with small surgical scissors or a scalpel. Use small surgical clips to gently keep the incision open, providing access to the skull. Optionally, sterile saline can be used with sterile swabs to clean the skull to aid in visualization of stereotaxic landmarks on the skull (Fig. 2).
8. Measure the *z* coordinates of Bregma and Lambda on the animal's skull and adjust the position of the head with the incisor adapter until they become equal. This serves to level the skull. Adjust the pitch of the ear bar to ensure that the skull is completely flat.
9. Position the tip of the Hamilton syringes to Bregma and record the *x*, *y*, and *z* coordinates on the vernier scale located on the arms of the stereotaxic instrument. Subtract the coordinates of the targeted brain region to calculate the site of targeted viral injection. These coordinates can be determined from a stereotaxic brain atlas (*see* Subheading 1 and **Note 6**). Note that the angle of the stereotaxic arm is an important

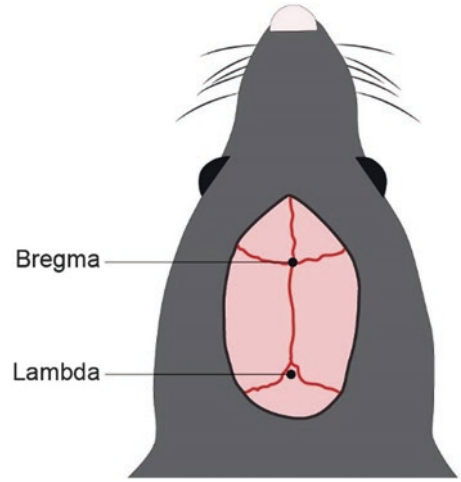


Fig. 2 Stereotaxic landmarks on the skull. The diagram above depicts the stereotaxic landmarks Bregma and Lambda on the exposed surface of the rodent's skull

consideration when determining the coordinates for targeting a desired brain region.

10. Position the tip of the Hamilton syringes according to the calculated *x* and *y* coordinates. Using a dental drill with a 0.6 mm burr, thin the area of the skull directly under the Hamilton syringe tip. Do not apply excessive downward force, as it may result in drilling through the skull and damaging the surface of the brain. Lower the Hamilton syringe on the *z* coordinate until it slides through the thinned skull, and raise the Hamilton syringe above the surface of the skull.
11. Proper safety attire and handling techniques should be applied based on the biosafety level of the virus being used (*see* **Note 7**). Defer to your institutional biosafety requirements for proper safety attire and handling techniques. The use of HSV vectors for our ZFP-mediated epigenome engineering experiments necessitates the use of a lab coat, gloves, and goggles when handling the virus. Place a viral aliquot in a low-binding, 0.65 mL microcentrifuge tube on wet ice, allowing it to thaw.
12. Taking care not to alter the *x* or *y* coordinates, expel 2.5 μL of sterile PBS from the Hamilton syringe. This volume will accumulate on the tip of the syringe, indicating unobstructed flow through the syringe tip (*see* **Note 8**). This volume can be removed with a sterile tip cotton swab. Draw the plunger up by an increment of 0.5 μL to introduce a small air bubble into the barrel of the syringe. This serves to separate the viral solution

from the sterile PBS. Finally, pull up the desired volume of virus to inject (typically 0.5–1 μL) and place the microcentrifuge tube back on wet ice.

13. Slowly lower the Hamilton syringe through the burr hole in the animal's skull to the calculated *z* coordinate to the desired injection site within the brain (Fig. 1b).
14. Deliver the viral solution by lowering the plunger of the Hamilton syringe at a rate of 0.1 μL per minute. Once the full volume of the viral solution has been dispensed, wait for 5 min for the virus to diffuse through the tissue (*see* **Note 9**).
15. To avoid backflow of the virus to the surface of the brain, *slowly* raise the Hamilton syringe out of the skull.

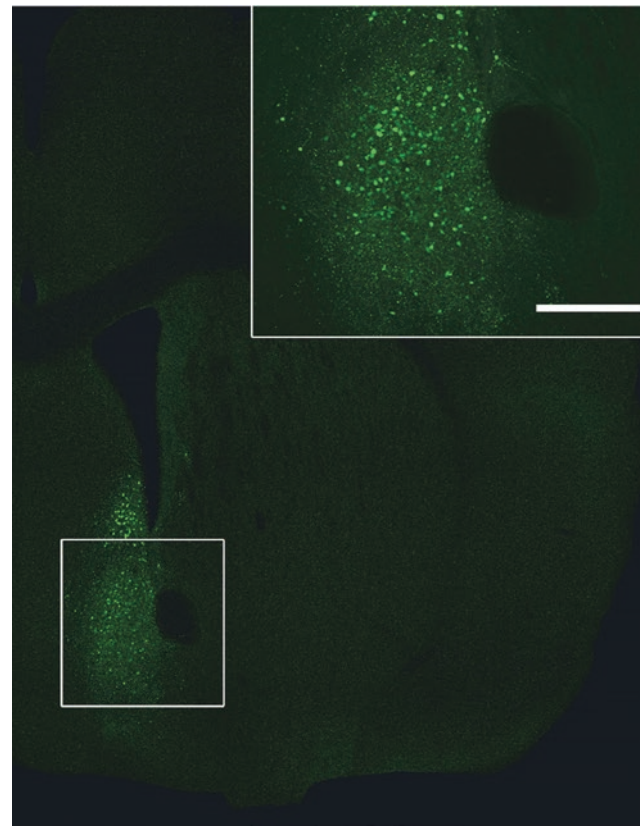


Fig. 3 Stereotaxic delivery of HSV to discrete brain regions. HSV expressing GFP under the control of the CMV promoter was stereotaxically injected into the nucleus accumbens (NAc) of a mouse to demonstrate the transduction efficiency and spread of the HSV viral vectors (image previously published [13]). These viral vectors are capable of co-expressing neuroepigenome editing constructs under the control of distinct promoters. The injection was performed at a 10° lateral angle at +1.6 anterior/posterior, +1.5 mediolateral, and −4.4 dorsal/ventral coordinate relative to Bregma. Scale bar is 300 μm

16. Expel the remaining contents of the Hamilton syringe into a flask containing a 10% bleach solution and use an alcohol prep wipe to remove any material that may have accumulated on the syringe tip.
17. Remove the animal from the stereotaxic instrument and close the incision via surgical suture or tissue adhesive. Small burr holes (less than 1 mm in diameter) do not need to be covered with bone wax. Apply antibiotic ointment to the wound and inject the local anesthetic bupivacaine subcutaneously near the wound, to reduce discomfort during the recovery period.
18. Place the animal in a clean cage that is warmed either by a temperature-regulated heating pad or a heat lamp until the animal fully recovers. This should take approximately 20 min, depending on the duration of the surgery.
19. Return the animal to a clean age with moistened food pellets for easy access to food. Monitor the animal's recovery, looking for any signs of distress which can include a lack of grooming, wound scratching, inflammation, altered locomotion, or reduced weight gain.
20. Clean the Hamilton syringes with 100% acetone and sterile PBS and according to the manufacturer's instructions (*see* **Note 10**). Discard the lab bench diapers into a biohazard receptacle and clean the workspace with 70% ethanol.
21. The time to maximal in vivo expression of our HSV-delivered, engineered transcription factors is approximately 2–3 days and persists through days 8–10 (Fig. 3). During this window, any number of molecular or behavioral experiments can be performed.

3.2 Validation of Neuroepigenomic Editing Tools

It is essential to validate the use of chromatin-modifying tools in several ways to ensure their effectiveness and selectivity in vivo [4, 5, 11]. Depending on your application, the essential validations may vary. The outline below provides a general list of suggested validations.

1. Validate expression of the tool (e.g., a fusion of a chromatin-modifying moiety to ZFP) in the brain in vivo (*see* **Note 11**). This includes validating selective expression in neurons when using a neurotrophic vector like HSVs, as well as selective expression within a single type of neuron if using a Cre-dependent vector in a mouse line that expresses Cre recombinase in a given cell type.
2. Validate that the epigenetic editing tool produces the designated chromatin modification at the targeted locus, for example, that the –p65 effector domain induces histone acetylation, that –G9a induces H3 Lys9 dimethylation, that –Tet1 induces DNA hydroxymethylcytosine, and that –Dnmt induces DNA methylation.

3. Determine whether the designated epigenetic modification is associated with altered expression of the targeted gene. This could be performed through the utilization of quantitative PCR, Western blot analysis, or other techniques to quantify expression levels of the product of the targeted gene. This analysis demonstrates that the targeted, ZFP-mediated epigenetic reprogramming at a given locus is sufficient to regulate gene expression through a physiologically relevant mechanism.
4. Study whether the designated chromatin modification is associated with other forms of epigenetic regulation, transcription factor binding, or changes in chromatin architecture. This line of inquiry is important to convey that the targeted epigenetic editing is capable of functionally integrating into the endogenous biological mechanisms of epigenetic regulatory events. These analyses could be performed with chromatin immunoprecipitation (ChIP) of transcription factors or epigenetic marks in question followed by quantitative PCR or next-generation sequencing, among other approaches.
5. Validate that the neuroepigenomic editing tool is selective for the targeted locus (*see* **Note 12**).

4 Notes

1. Ensure that the stereotaxic frame and accessories including the ear bars and incisor adaptor are appropriate for the type of animal to receive surgery.
2. If the animal does not reach a sufficient level of surgical anesthesia after 10–15 min, inject an additional 20% dose of ketamine/xylazine. Closely monitor the animal to confirm that the anesthesia deepens.
3. If the animal begins to awaken during surgery, remove the animal from the stereotaxic instrument and reapply the anesthesia. The early signs of an animal awakening from anesthesia include twitches of the large facial whiskers and twitching of the tail. With careful monitoring, this occurrence can be avoided.
4. It is essential that the animal be firmly secured in the ear bars. Visually validate that the ear bars are in the ear canal and not pinching the jaw, neck, or skull. Animals appropriately positioned in the ear bars will be able to move their snout up and down in the incisor adapter, but will not be able to move side to side.
5. If the animal is not securely placed in the stereotaxic instrument, then it is possible that the skull's position will shift when

drilling burr holes. This invalidates all recorded coordinates. To be sure this does not occur, when the animal is first positioned in the instrument; apply light pressure to the skull with a sterile tip cotton swab. If the animal's skull shifts in response to the pressure, resecure the animal within the stereotaxic instrument.

6. The stereotaxic coordinates provided in atlases are optimized for adult, male animals. If experiments involve varying from these average metrics, it becomes important to validate and/or alter targeting coordinates through pilot experiments. In short, use the stereotaxic atlas coordinates as initial values, perform surgeries, and validate viral targeting with fluorescent microscopy. Adjust the stereotaxic coordinates as needed.
7. Selecting the appropriate viral vector for delivery of neuroepigenomic editing tools is paramount to the success of these experiments. Each viral vector varies in its spread, packaging capacity, tropism, transgene expression timing, and duration of expression. These variables should be carefully considered, and pilot studies should be performed to empirically validate viral function.
8. If the Hamilton syringe clogs, it will prevent dispersion of the virus into the brain. Always visually confirm that the flow from the syringe is not impeded by expelling a very small volume of the viral solution back into the microcentrifuge tube before lowering into the rodent brain. If the syringe does not appear to work, expel the contents of the syringe into a 10% bleach solution, clean with 100% acetone and/or replace the Hamilton needle, and restart the process of loading viral solution into the Hamilton syringe.
9. The Hamilton syringe tip is left in place during the 5-min rest after delivering virus in order to prevent backflow of viral solution up the needle track. However, we have found it beneficial during this time to *slightly* retract the Hamilton syringe along the *z*-axis (<1 mm) to provide a small space for the more even dispersion of the viral solution in the tissue.
10. If using surfactant-based cleaning solutions for cleaning and maintaining Hamilton syringes (as is often suggested according to the manufacturer's instructions), be extremely vigilant to thoroughly remove all traces of soap, as it can be damaging to viral function and titers. There should be no formation of soap bubbles when pipetting sterile PBS.
11. Introducing an epitope tag onto the neuroepigenome editing tool facilitates more convenient immunohistochemical approaches for validating expression, among other uses.
12. To date, validation of selectivity of targeting has been accomplished primarily by identifying regions of the genome that are

most highly homologous to the targeted region and demonstrating lack of changes of chromatin modifications of these other regions and lack of altered expression of any nearby genes. Ultimately, it is advantageous to demonstrate direct and selective binding of the epigenetic editing tool to the targeted locus and to no other genomic location. This has been successfully performed in cultured cell lines [11, 12]. However, this is challenging, particularly in brain in vivo, since every cell only has two specific binding locations and the quantity of infected tissue can be limiting.

References

1. Franklin KBJ, Paxinos G (2008) The mouse brain in stereotaxic coordinates, 3rd edn. Elsevier/Academic Press, Amsterdam; New York
2. Paxinos G, Watson C (2007) The rat brain in stereotaxic coordinates, 6th edn. Elsevier, Amsterdam; Boston
3. Neve RL (2012) Overview of gene delivery into cells using HSV-1-based vectors. Curr Protoc Neurosci Chapter 4:Unit 4.2
4. Heller EA, Cates HM, Pena CJ, Sun H, Shao N, Feng J, Golden SA, Herman JP, Walsh JJ, Mazei-Robison M, Ferguson D, Knight S, Gerber MA, Nievera C, Han MH, Russo SJ, Tamminga CS, Neve RL, Shen L, Zhang HS, Zhang F, Nestler EJ (2014) Locus-specific epigenetic remodeling controls addiction- and depression-related behaviors. Nat Neurosci 17:1720–1727
5. Heller EA, Hamilton PJ, Burek DD, Lombroso SI, Pena CJ, Neve RL, Nestler EJ (2016) Targeted epigenetic remodeling of the Cdk5 gene in nucleus accumbens regulates cocaine- and stress-evoked behavior. J Neurosci 36:4690–4697
6. Hamilton PJ, Burek DJ, Lombroso SI, Neve RL, Robison AJ, Nestler EJ, Heller EA (2018) Cell-type specific epigenetic editing at the Fosb gene controls susceptibility to social defeat stress. Neuropsychopharmacology 43: 272–284
7. de Groote ML, Verschure PJ, Rots MG (2012) Epigenetic editing: targeted rewriting of epigenetic marks to modulate expression of selected target genes. Nucleic Acids Res 40:10596–10613
8. Jamieson AC, Miller JC, Pabo CO (2003) Drug discovery with engineered zinc-finger proteins. Nat Rev Drug Discov 2:361–368
9. Sweatt JD (2013) The emerging field of neuroepigenetics. Neuron 80:624–632
10. Pena CJ, Bagot RC, Labonte B, Nestler EJ (2014) Epigenetic signaling in psychiatric disorders. J Mol Biol 426:3389–3412
11. Liu XS, Wu H, Ji X, Stelzer Y, Wu X, Czauderna S, Shu J, Dadon D, Young RA, Jaenisch R (2016) Editing DNA methylation in the mammalian genome. Cell 167(233-247):e217
12. Thakore PI, D'Ippolito AM, Song L, Safi A, Shivakumar NK, Kabadi AM, Reddy TE, Crawford GE, Gersbach CA (2015) Highly specific epigenome editing by CRISPR-Cas9 repressors for silencing of distal regulatory elements. Nat Methods 12:1143–1149
13. Engmann O, Labonte B, Mitchell A, Bashtrykov P, Calipari ES, Rosenbluh C, Loh YE, Walker DM, Burek D, Hamilton PJ, Issler O, Neve RL, Turecki G, Hurd Y, Chess A, Shen L, Mansuy I, Jeltsch A, Akbarian S, Nestler EJ (2017) Cocaine-induced chromatin modifications associate with increased expression and three-dimensional looping of Auts2. Biol Psychiatry 82:794–805

Chapter 18

In Vivo Applications of Cell-Penetrating Zinc-Finger Transcription Factors

Chonghua Ren, Alexa N. Adams, Benjamin Pyles, Barbara J. Bailus, Henriette O'Geen, and David J. Segal

Abstract

Artificial transcription factors based on zinc finger, TALE, and CRISPR/Cas9 programmable DNA-binding platforms have been widely used to regulate the expression of specific genes in cultured cells, but their delivery into organs such as the brain represents a critical challenge to apply such tools in live animals. In previous work, we developed a zinc-finger-based artificial transcription factor harboring a cell-penetrating peptide (CPP) that could be injected systemically, cross the blood–brain barrier, and alter expression of a specific gene in the brain of an adult mouse. Importantly, our mode of delivery produced widespread distribution throughout the brain. Here we describe methods for the production and purification of the factor, testing CPP activity in cells, and testing CPP activity in mice.

Key words Engineered zinc-finger protein, Cell-penetrating peptide, TAT peptide, Animal models, Artificial transcription factors

1 Introduction

Artificial transcription factors (ATFs) are based on the attachment of transcriptional effector domains to programmable DNA-binding platforms such as zinc fingers (ZFs), transcription activator-like effectors (TALEs), or catalytically inactive clustered regularly interspaced short palindromic repeats/dead Cas9 (CRISPR/dCas9). These tools are capable of activating or repressing specific genes as has been described extensively [1–4]. Widespread delivery of these gene regulators to the organs such as brain remains a significant challenge for the study and treatment of medical conditions.

We previously reported the systemic delivery of a purified ATF protein as a potential therapeutic approach for the treatment of Angelman syndrome (AS) [5]. AS is a rare neurological genetic disorder caused by loss or mutation of the maternal copy of *UBE3A* in the brain. Due to brain-specific genetic imprinting at this locus the paternal copy of *UBE3A* is silenced, resulting in the complete

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_18,

loss of UBE3A expression in brain neurons of patients. The ATF was designed to unsilence the paternal *UBE3A* by inhibiting an antisense transcript that is responsible for the silencing. However, as with many gene therapy methods, a significant challenge to actual therapeutic use is an efficient means of delivering the ATF to the many neurons in the brain.

One method for protein delivery into cells is the use of cell-penetrating peptides (CPPs). These short peptides are positively charged sequences that can be added to various proteins to facilitate their translocation across cellular membranes, usually by hijacking the normal process of receptor-mediated endocytosis [6, 7]. The cationic class of CPPs contains clusters of arginine and lysine residues, such as Antp (RQIKIWFQNRRMKWKK), derived from the third helix of the Antennapedia protein homeodomain from Drosophila [8], and Tatp (GRKKRRQRRR), derived from HIV-1 TAT transcription-activating protein [9]. Successful delivery using these CPPs had been shown both in vitro and in vivo with several full-length proteins that retain their biological activity [6, 7, 10–12], including engineered ZF proteins [13–15].

For our AS studies, we combined the ATF with a CPP in a construct composed of an N-terminal maltose-binding protein for purification, the CPP consisting of the 10 aa transduction domain of the HIV transactivator protein (TAT, residues 48–57), an mCherry red fluorescent protein to aid in protein solubility and visualization, an HA epitope tag for detection, an SV40 nuclear localization signal to ensure nuclear delivery, an engineered zinc-finger protein (designated "S1"), and a KRAB transcriptional repression domain (Fig. 1) [5]. The purified fusion protein was able to enter cells both in vitro (Fig. 2) and in vivo (Fig. 3). In both cases, a protein containing all the components except the TAT CPP was still able to enter cells, but with considerably reduced efficiency. In mice, we observed that intraperitoneal (i.p.) or subcutaneous (s.c.) injection of the purified protein at 160–200 mg/kg was able to cross the BBB and distribute widely throughout the brain. Importantly, significant activation of Ube3a expression in the brain was observed after a 4-week treatment period.

In principle, a CPP could be used to deliver ATFs designed to target other promoters or DNA elements in the brain or other organs by using a different ZF protein, TALEs or dCas9. Exchanging the KRAB transcriptional repression domain with an activation domain (e.g., VP64 or p300) or epigenetic modifiers (e.g., DNMT3A or G9A) could produce tools for activating gene expression and altering epigenetic information, respectively. Cellular internalization of CPPs has been observed for a large number of cell types and tissues, although the efficiencies vary depending on the CPP, the cargo, and the target cell type [16, 17]. In fact, it has recently been shown that since ZFs carry a net positive

B

Maltose binding protein (MBP)
MKIEEGKLVIWINGDKGYNGLAEVGKKFEKDTGIKVTVE
HPDKLEEKFPQVAATGDGPDIIFWAHDRFGGYAQSGLLA
EITPDKAFQDKLYPFTWDAVRYNGKLIAYPIAVEALSLI
YNKDLLPNPPKTWEEIPALDKELKAKGKSALMFNLQEPY
FTWPLIAADGGYAFKYENGKYDIKDVGVDNAGAKAGLTF
LVDLIKNKHMNADTDYSIAEAAFNKGETAMTINGPWAWS
NIDTSKVNYGVTVLPTFKGQPSKPFVGVLSAGINAASPN
KELAKEFLENYLLTDEGLEAVNKDKPLGAVALKSYEEEL
AKDPRIAATMENAQKGEIMPNIPQMSAFWYAVRTAVINA
ASGRQTVDEALKDAQTNSSSNNNNNNNNNNLG(IEGRIS
EFGSGAP)

TAT cell penetrating peptide
GRKKRRQRRR(VDM)

mCherry red fluorescent protein
VSKGEEDNMAIIKEFMRFKVHMEGSVNGHEFEIEGEGEG
RPYEGTQTAKLKVTKGGPLPFAWDILSPQFMYGSKAYVK
HPADIPDYLKLSFPEGFKWERVMNFEDGGVVTVTQDSSL
QDGEFIYKVKLRGTNFPSDGPVMQKKTMGWEASSERMYP
EDGALKGEIKQRLKLKDGGHYDAEVKTTYKAKKPVQLPG
AYNVNIKLDITSHNEDYTIVEQYERAEGRHSTGGMDELY
K(SR)

HA epitope tag
YPYDVPDYA(NSPGIPGMG)

SV40 nuclear localization signal (NLS)
PKKKRKVG(RLE)

Zinc finger protein S1
PGEKPYMCAECGKSFSRSDDLVRHQRTHTGEKP
YKCPECGKSFSDCRDLARHQRTHTGEKP
YKCPECGKSFSQRAHLERHQRTHTGEKP
YKCPECGKSFSREDNLHTHQRTHTGEKP
YKCPECGKSFSRSDDLVRHQRTHTGEKP
YKCPECGKSFSTSGNLTEHQRTHTG(AAATL)

KRAB transcriptional repression domain
VTFKDVFVDFTREEWKLLDTAQQIVYRNVMLEKYKNLVS
LGYQLTKPDVILRLEKGEEPWLVEREIHQETHP*

Fig. 1 Design of TAT-S1 ATF. (**a**) Diagram of TAT-S1 ATF indicating individual protein domains. (**b**) Protein sequence of TAT-S1 ATF. Individual domains are indicated. Intervening sequences are shown as grey sequences in parenthesis

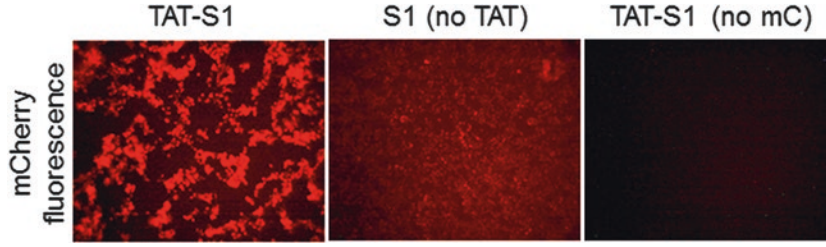


Fig. 2 TAT-S1 ATF in cultured HEK293 cells. mCherry fluorescent images of cells transduced with 11 μM TAT-S1 ATF, S1 ATF with no TAT, or TAT-S1 ATF with no mCherry

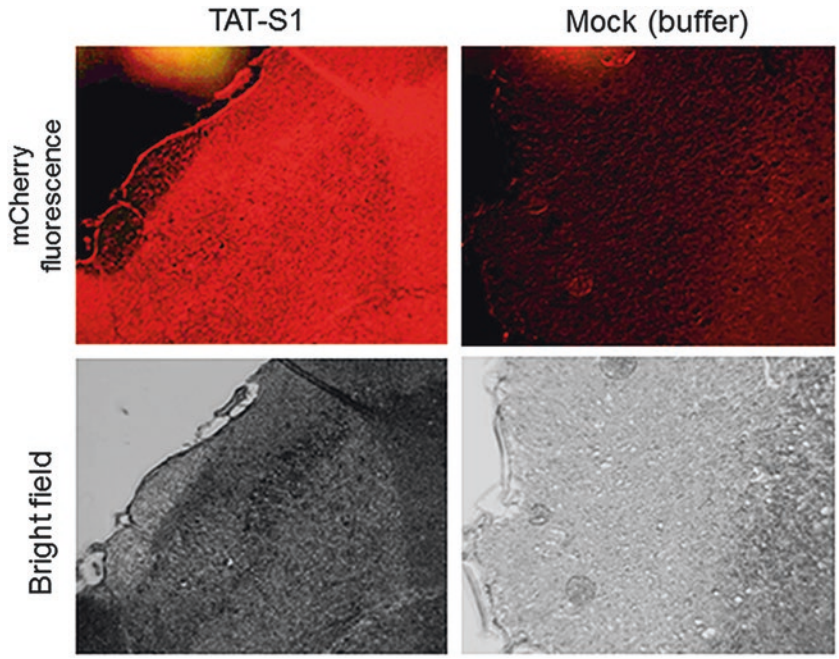


Fig. 3 TAT-S1 ATF in mouse brain. Sectioned tissues (50 μm) were harvested after 3 days of injections (160–200 mg/kg s.c.) with purified TAT-S1 ATF, or elution buffer (Mock) that contains no ATF as a negative control. Top panel, mCherry fluorescent images of mouse brain cortex; bottom panel, bright field

charge they can penetrate cells in the absence of additional peptides [18, 19]. It is therefore important to test CPPs empirically to determine their utility for a particular application. The method that follows can be used to quickly determine if a CPP increases cellular uptake more than a zinc-finger protein alone, in vitro (Subheading 2) and in vivo (Subheading 3). Although the mCherry fluorescence could be directly visualized in mouse tissues, an antibody staining protocol is provided in Subheading 3 that could be used for weaker fluorophores.

2 Materials

2.1 Expression and Purification of CPP-ZF ATFs in Bacteria

1. Construct a prokaryotic expression vector expressing a CPP-ZF ATF. There are a number of possibilities, but a good starting point might be to use TAT-S1 ATF construct described in our published study. A modified pMAL-c2X (New England Biolabs, Ipswich, MA) with an expression cassette containing (1) an N-terminal maltose-binding protein (MBP) for purification, (2) a TEV1 protease cleavage site, (3) a cell-penetrating peptide consisting of the 10 aa transduction domain of the HIV-transactivator protein (TAT, residues 48–57), (4) mCherry red fluorescent protein to aid in protein solubility

and visualization, (5) an HA epitope tag for detection, (6) an SV40 nuclear localization signal to ensure nuclear delivery, (7) a six-finger zinc-finger DNA-binding protein designated "S1," and (8) a KRAB transcriptional repression domain. "S1 (no TAT)" and "S1 no mCherry" do not contain the TAT or mCherry components, respectively. Vectors expressing these constructs are available from the authors upon request. The complete sequence of the full-length ATF TAT S1 protein is provided in Fig. 1 (*see* **Note 1**).

2. Chemically competent NEB5α *E. coli* bacteria (New England Biolabs): Store at −80 °C (*see* **Note 2**).
3. Carbenicillin antibiotic stock at 100 mg/mL in H_2O: Keep stock at −20 °C.
4. 10 cm Plates of Luria Broth (LB) agar (*see* **Note 3**) supplemented with carbenicillin at 50 μg/mL: Prepare no more than 1 month in advance and store at 4 °C.
5. 1× Luria Broth medium (*see* **Note 3**).
6. Isopropyl β-D-1-thiogalactopyranoside (IPTG) stock at 0.5 M in H_2O: Keep stock at −20 °C.
7. Zinc chloride ($ZnCl_2$) stock at 1 M in H_2O: Store at room temperature.
8. Zinc buffer A (ZBA): 10 mM Tris base, 90 mM KCl, 1 mM $MgCl_2$, 100 μM $ZnCl_2$. Adjust pH to 8.5 using HCl (*see* **Note 4**). Store at room temperature.
9. Bacterial culture shakers at both 37 and 4 °C.
10. Centrifuge for bacterial cultures.
11. Microfluidizer (Microfluidics model M-110Y).
12. Chromatography columns (~100 mL volume) with valves.
13. Amylose resin (New England Biolabs, E8021L).
14. Maltose.
15. Dithiothreitol (DTT) stock at 1 M in H_2O: Aliquot in 0.5–2 mL volumes and store at −20 °C.
16. Elution buffer: ZBA, 1 M maltose.
17. Centricon Plus-70 spin concentrators (Millipore, Billerica, MA, UFC710008).
18. Refrigerated tabletop centrifuge.
19. Nalgene Rapid-Flow sterile disposable filter units, 0.2 μm (Thermo Fisher).
20. HEK-Blue LPS Detection Kit (InvivoGen, San Diego, CA, rep-lps2).
21. 4–20% Mini-PROTEAN TGX Precast Protein Gels (Bio-Rad, Hercules, CA, 4561096).

22. Apparatus and materials for performing Coomassie-stained SDS-PAGE.
23. Nanodrop UV spectrophotometer, or equivalent.
24. Glycerol, sterilized by autoclave.

2.2 Testing CPP Activity in Cell Culture

1. Purified ATF protein in elution buffer, 30% glycerol, 5 mM DTT.
2. The human cell line HEK293 (ATCC CRL-1573, *see* **Note 5**).
3. Water-jacketed CO_2 37 °C incubator.
4. Class II biosafety cabinet.
5. Six-well polystyrene tissue culture plates.
6. 75 cm^2 Tissue culture flasks.
7. DMEM medium.
8. Bovine calf serum: 50 mL Aliquots are stored at −20 °C.
9. Pen–Strep contains 100 U/mL of penicillin and 100 μg/mL streptomycin: 5 mL Aliquots are stored at −20 °C.
10. DMEM complete medium: 1 L DMEM medium supplemented with 50 mL bovine calf serum (10% final) and 5 mL of Pen–Strep. Store at 4 °C.
11. Phosphate-buffered saline, sterile for tissue culture.
12. Trypsin-EDTA.
13. Hemocytometer for counting cells.
14. 10% bleach solution.
15. Zeiss Axiovert 135 inverted epifluorescent microscope with filters for mCherry imaging (Filter Set 15, Excitation: 546 nm; Emission: LP 590 nm) and a digital camera.

2.3 Testing CPP Activity in Mice

1. Purified ATF protein in elution buffer, 30% glycerol, 5 mM DTT.
2. C57BL/6 mice of either sex, approximately 8 weeks of age.
3. One milliliter syringe with 25-gauge hypodermic needle.
4. Isoflurane and appropriate apparatus.
5. CO_2 chamber or other apparatus for humane euthanasia.
6. TBS: 50 mM Tris–Cl, pH 7.5, 150 mM NaCl.
7. Buffered paraformaldehyde: 10% (v/v) in PBS.
8. 30% Sucrose in water.
9. Tissue-Tek CRYO-OCT Compound (Thermo Fisher 14-373-65).
10. Glass microscopy slides.
11. Glass coverslips.
12. Leica CM 1850 UV cryotome (Nussloch, Germany).

13. Superblock (Thermo Fisher).
14. Goat serum.
15. Triton X.
16. Apex Antibody labeling kit with Alexa Fluor 555.
17. Anti-HA primary antibody: Pre-label with the Apex Antibody labeling kit.
18. 4′,6-Diamidino-2-phenylindole (DAPI).
19. Prolong Gold (Thermo Fisher, P10144).
20. Leica DM6000B epifluorescent microscope with deconvolution software and filter cubes to image mCherry (TxRed 4040B, excitation: 562 nm; emission: 624 nm) and DAPI (DAPI 5060B, excitation: 377: emission 447).

3 Methods

3.1 Expression and Purification of CPP-ZF ATFs in Bacteria

1. Transform the vector into NEB5α *E. coli* by standard heat-shock methods.
2. Plate the transformed bacteria on LB agar supplemented with 50 μg/mL carbenicillin and incubate overnight.
3. Pick a single colony to inoculate into 5 mL of LB medium supplemented with 50 μg/mL carbenicillin. Incubate overnight with shaking at 37 °C.
4. Inoculate the 5 mL overnight cultures into 800 mL of LB medium supplemented with 50 μg/mL carbenicillin. Incubate with shaking overnight at 37 °C.
5. At optical density ~1.0, induce protein expression by moving the culture to 4 °C (*see* **Note 6**) and adding IPTG to 0.75 mM and 1 mL of zinc chloride. Shake gently at 4 °C for 4 days.
6. To release the protein from the bacteria, pellet by centrifugation at 1663 × *g* (3200 rpm) for 15 min in a refrigerated tabletop centrifuge. Resuspend the culture in 30 mL of cold ZBA (*see* **Note 7**). Apply the suspension to a microfluidizer according to the manufacturer's instructions (*see* **Note 8**). Keep the lysate on ice.
7. Pellet the insoluble fraction by centrifugation at 1663 × *g* for 15 min in a refrigerated tabletop centrifuge. Keep supernatant for column purification and discard pellet.
8. Prepare gravity flow amylose resin purification columns by applying enough amylose resin to produce a 30 mL compact bed in the columns (*see* **Note 9**). All steps can be performed at room temperature. Wash the columns twice with four-column volumes of deionized water, and then one-column volume of ZBA. Finally, apply the microfluidized supernatants to the columns.

The initial drip rate should be approximately two drops/s, but will decrease as the solution in the column decreases.

9. Dissociate and elute the protein in ~150 mL of elution buffer supplemented with 5 mM DTT (add DTT just before use). Keep eluted samples on ice.
10. Concentrate the eluate to 16 mg/mL using a Centricon Plus-70 in a refrigerated centrifuge at 2000 × *g* for ~60 min (depending on the initial concentration) at 4 °C. For maximum tolerated dose studies, additional higher concentrations may be desired.
11. Sterilize the protein sample using Nalgene Rapid-Flow sterile disposable filter units to remove any residual bacteria. Purified, sterilized protein samples are routinely checked by an endotoxin kit (e.g., HEK-Blue LPS Detection Kit) to assure that no detectable endotoxins are present.
12. Measure protein concentration using a Nanodrop UV spectrophotometer at A280, blanking with elution buffer. The procedure typically yields 4 g total protein/L of culture for S1-KRAB, but values will likely change for other ATFs. Evaluate protein integrity by SDS polyacrylamide gel electrophoresis using a 4–20% TGX precast protein gel followed by staining with Coomassie blue. Usually there is a 44 kD band corresponding to free MBP. Protein concentrations for injections refer to the full-length + free-MBP band intensities, of which only half was considered to be the 100 kD full-length protein.
13. For storage of the proteins, add glycerol to 30% and DTT to 5 mM. This typically decreases the concentration from 16 to 12 mg/mL total protein. Store at −20 °C (*see* **Note 10**). Note that it is difficult to measure the protein concentration after addition of glycerol, so concentration is measured in the previous step.

3.2 Testing CPP Activity in Cell Culture

1. Obtain prior approval from the Institutional Biosafety Committee of the investigator's institution before any work using human cells is performed.

Day 0: Seed the cells in the six-well plates

2. Prepare the HEK293 cells to seed on six-well plates. Aspirate the medium from the cells growing in a 75 cm^2 flask.
3. Wash the cells with 5 mL of PBS. Swirl the PBS around to wash away remaining medium, and then aspirate the PBS.
4. Add 2 mL of trypsin-EDTA and swirl the flask to ensure that the entire surface is covered. Incubate at room temperature for 1 min.
5. Resuspend the cells using 8 mL of DMEM complete medium.

6. Place 15 μL of the above suspension onto a hemocytometer. Count the number of cells using an inverted microscope. After use, decontaminate the hemocytometer in 10% final bleach solution.
7. Add 400,000–600,000 cells to each well.
8. Incubate for 24 h in a 37 °C CO_2 incubator before protein transduction.

Day 1: Protein transduction

9. The cells should be about 60% confluent at the time of protein transduction. Replace the complete medium with 2 mL DMEM without supplements.
10. Add ATF proteins directly to the wells. Varying concentrations should be used. The results shown in Fig. 2 required 100 μL of the 12 mg/mL purified protein, for a final concentration of about 11 nM per well (*see* **Note 11**).
11. Incubate for 2 h in a 37 °C incubator.
12. Add an additional 2 mL of DMEM complete medium.
13. Incubate the cells in a 37 °C incubator for 24 h.

Day 2: Imaging

14. To prepare the cells for imaging, remove the media from wells.
15. Wash twice with 2 mL of PBS to remove remaining red indicator dye from the DMEM complete medium. Image in 2 mL of PBS.
16. Image using a Zeiss Axiovert 135 inverted epifluorescent microscope with filters for mCherry imaging. Example images are shown in Fig. 2.

3.3 Testing CPP Activity in Mice

ATF injections

1. Obtain prior approval from the Institutional Animal Care and Use Committee of the investigator's institution before any work on animals is performed.
2. Mice should be anesthetized with 4% isoflurane before injection (*see* **Notes 12**). Inject three mice with TAT-S1 ATF at a dose of 160–200 mg/kg *s.c.* (*see* **Note 13**). Also inject three mice with elution buffer, 30% glycerol, and 5 mM DTT ["Mock"] as a negative control. When mice are ambulatory, they can be returned to their cages.
3. Repeat **step 2** for 2 consecutive days, for a total of three injection days.
4. Humanly euthanize mice on the fourth day. Harvest brain or other organs to be studied. Tissues that are not used immediately for immunohistochemistry should be flash frozen.

Preparation of tissue sections

5. Wash one brain hemisphere in 1× TBS, and then fix in 10% buffered paraformaldehyde overnight.
6. Place the tissue in 30% sucrose for 3 days.
7. Following brain saturation, freeze the tissue in Tissue-Tek CRYO-OCT Compound, and then section on a Leica cryotome in 50 μm thick sections. Transfer the sections to glass microscope slides.

Immunohistochemistry (if florescence cannot be observed from tissues directly)

8. Block using Superblock for 1 h, followed by 10% goat serum and 0.3% Triton X in 1× TBS for 2 h.
9. Aspirate slides, and then stain with primary anti-HA (that had been directly labeled with the Apex Antibody labeling kit) at 1:150 in 5% goat serum and 0.15% Triton X.
10. Incubate at 4 °C overnight.
11. Wash slides three times with 1× TBS.
12. Stain with DAPI for 10 min.
13. Wash three times with 1× TBS.

Imaging

14. Mount with coverslip using Prolong Gold.
15. Image using a Leica DM6000B epifluorescent microscope using filter cubes for imaging DAPI and mCherry. Example images are shown in Fig. 3.

4 Notes

1. The expression and purification methods described here will likely require significant optimization for any new type of zinc-finger protein or effector domain used. Also, at the time of this writing, it is not clear if protein domains such as MBP and mCherry are required for full function of the ATF.
2. BL21 Star cells could also be used for protein expression. The ATF S1-KRAB seems to express equally well in both BL21 Star and NEB5a cells. However, others, such as TALE proteins, express much better in BL21.
3. Any source of LB medium is usually acceptable. However, we have found for some ATFs (not S1-KRAB) that LB from some vendors produced a dramatic reduction in yield, which was restored by using LB from VWR.

4. A good guide to the appropriate pH for the ZBA is the protein's isoelectric point, which can be calculated by any of several online services (e.g., http://web.expasy.org/protparam/). Some optimization may be necessary.
5. For safety, human cells should always be handled using the philosophy of Universal Precautions, as if they were contaminated with a blood-borne pathogen such as HIV or hepatitis virus B. This includes biosafety level 2 conditions, such as appropriate personal protective equipment and a biosafety cabinet.
6. The cold-temperature induction was critical to obtaining high yields of the TAT-S1 ATF protein. Induction at room temperature or 37 °C was far less efficient. This unusual requirement was fortuitously observed when several standard induction conditions were tried and proved unsatisfactory. Concurrently, a culture that had been accidentally left at 4 °C for several days turned noticeably pink. This was an indication that the protein was being expressed, at least the mCherry domain. Experimental refinement of methods resulted in the reported production protocol. However, versions of the ATF that contained other effector domains required induction at an optical density of 0.6, and incubation for only 4 h at 37 °C.
7. Protease inhibitors are not used when purifying the ATF S1-KRAB. They are sometimes used with other factors, especially if degradation appears as a significant issue. However, we have generally been cautious out of concern for undesired effects of residual protease inhibitors in animals.
8. Use of the microfluidizer was found to be critical for obtaining full-length protein. Sonication and freeze/thaw techniques produced fragmented proteins.
9. The amylose resin can be reused up to 14 times by washing. In some cases, yield seemed to increase using resin that had been used and washed.
10. Proteins were originally stored at −80 °C, but later studies showed less protein fragmentation due to freeze/thaw when the ATF was stored at −20 °C.
11. Some optimization of the protein concentration applied to the cells may be necessary. It is useful to test a range of concentrations initially.
12. In addition to avoiding accidental autoinoculation by trying to inject a moving mouse, anesthesia before injection also provides a transient immunosuppression that prevents an acute immune response to the protein when the bolus is injected. Anesthesia is thus highly recommended.

13. A power calculation would be performed first to determine how many mice would be required to observe the molecular phenotype (i.e., the change in Ube3a expression). This calculation requires an estimation of the variance in the live fluorescence assay. Since the variance may differ for different proteins, testing three mice here can provide information on the variance that can be used to make a more informed calculation for a sufficiently powered experiment. In our experience, three mice per group are typically sufficient.

Acknowledgments

We thank Jennifer Trang Nguyen for her assistance with experiments. We thank Enoch Baldwin and Sarah Lockwood for expert advice and discussions in developing these methods. This work was supported by the NIH (NS071028), the Angelman Syndrome Foundation, and the Foundation for Angelman Syndrome Therapeutics. B.J.B. was also funded by an NSF fellowship (0707429) and a grant to UC Davis from the Howard Hughes Medical Institute through the Med into Grad Initiative (56005706) and a CTSC pilot study (TR000002).

Competing Interests: The authors declare no competing interests.

References

1. Blancafort P, Segal DJ, Barbas CF 3rd (2004) Designing transcription factor architectures for drug discovery. Mol Pharmacol 66:1361–1371
2. Polstein LR, Perez-Pinera P, Kocak DD, Vockley CM, Bledsoe P, Song L, Safi A, Crawford GE, Reddy TE, Gersbach CA (2015) Genome-wide specificity of DNA binding, gene regulation, and chromatin remodeling by TALE- and CRISPR/Cas9-based transcriptional activators. Genome Res 25:1158–1169
3. Thakore PI, Black JB, Hilton IB, Gersbach CA (2016) Editing the epigenome: technologies for programmable transcription and epigenetic modulation. Nat Methods 13:127–137
4. Thakore PI, Gersbach CA (2016) Design, assembly, and characterization of TALE-based transcriptional activators and repressors. Methods Mol Biol 1338:71–88
5. Bailus BJ, Pyles B, McAlister MM, O'Geen H, Lockwood SH, Adams AN, Nguyen JT, Yu A, Berman RF, Segal DJ (2016) Protein delivery of an artificial transcription factor restores widespread Ube3a expression in an Angelman syndrome mouse brain. Mol Ther 24:548–555
6. Mae M, Langel U (2006) Cell-penetrating peptides as vectors for peptide, protein and oligonucleotide delivery. Curr Opin Pharmacol 6:509–514
7. Wagstaff KM, Jans DA (2006) Protein transduction: cell penetrating peptides and their therapeutic applications. Curr Med Chem 13:1371–1387
8. Derossi D, Joliot AH, Chassaing G, Prochiantz A (1994) The third helix of the Antennapedia homeodomain translocates through biological membranes. J Biol Chem 269:10444–10450
9. Vives E, Brodin P, Lebleu B (1997) A truncated HIV-1 Tat protein basic domain rapidly translocates through the plasma membrane and accumulates in the cell nucleus. J Biol Chem 272:16010–16017
10. Schwarze SR, Ho A, Vocero-Akbani A, Dowdy SF (1999) In vivo protein transduction: delivery of a biologically active protein into the mouse. Science 285:1569–1572
11. Dietz GP, Bahr M (2004) Delivery of bioactive molecules into the cell: the Trojan horse approach. Mol Cell Neurosci 27(2):85–131

12. Gong B, Cao Z, Zheng P, Vitolo OV, Liu S, Staniszewski A, Moolman D, Zhang H, Shelanski M, Arancio O (2006) Ubiquitin hydrolase Uch-L1 rescues beta-amyloid-induced decreases in synaptic function and contextual memory. Cell 126:775–788
13. Yun CO, Shin HC, Kim TD, Yoon WH, Kang YA, Kwon HS, Kim SK, Kim JS (2008) Transduction of artificial transcriptional regulatory proteins into human cells. Nucleic Acids Res 36:e103
14. Mino T, Mori T, Aoyama Y, Sera T (2007) Development of protein-based antiviral drugs for human papillomaviruses. Nucleic Acids Symp Ser (Oxf) 51:427–428
15. Mino T, Mori T, Aoyama Y, Sera T (2008) Cell-permeable artificial zinc-finger proteins as potent antiviral drugs for human papillomaviruses. Arch Virol 153:1291–1298
16. Maiolo JR, Ferrer M, Ottinger EA (2005) Effects of cargo molecules on the cellular uptake of arginine-rich cell-penetrating peptides. Biochim Biophys Acta 1712:161–172
17. Mueller J, Kretzschmar I, Volkmer R, Boisguerin P (2008) Comparison of cellular uptake using 22 CPPs in 4 different cell lines. Bioconjug Chem 19:2363–2374
18. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 9:805–807
19. Gaj T, Liu J, Anderson KE, Sirk SJ, Barbas CF 3rd (2014) Protein delivery using Cys2-His2 zinc-finger domains. ACS Chem Biol 9:1662–1667

Chapter 19

Manufacturing and Delivering Genome-Editing Proteins

Jia Liu, Ya-jun Liang, Pei-ling Ren, and Thomas Gaj

Abstract

Genome-editing technologies have revolutionized the biomedical sciences by providing researchers with the ability to quickly and efficiently modify genes. While programmable nucleases can be introduced into cells using a variety of techniques, their delivery as purified proteins is an effective approach for limiting off-target effects. Here, we describe step-by-step procedures for manufacturing and delivering genome-modifying proteins—including Cas9 ribonucleoproteins (RNPs) and TALE and zinc-finger nucleases—into mammalian cells. Protocols for combining Cas9 RNP with naturally recombinogenic adeno-associated virus (AAV) donor vectors for the seamless insertion of transgenes by homology-directed genome editing are also provided.

Key words ZFNs, TALENs, CRISPR, RNP, Protein delivery, Genome editing

1 Introduction

Programmable nucleases—including the RNA-guided Cas9 endonuclease from type II clustered regularly interspaced short palindromic repeats (CRISPR)-associated (Cas) systems, transcription activator-like effector (TALE) nucleases (TALENs), and zinc-finger nucleases (ZFNs)—have revolutionized biomedical research and biotechnology. These technologies can be configured to recognize a specific genomic site to induce a DNA double-strand break (DSB) that is processed by cellular DNA repair mechanisms, enabling a range of genomic modifications, including gene knockout via non-homologous end joining (NHEJ) [1, 2] and the site-specific insertion of transgenes by homology-directed repair (HDR) [3–5].

ZFNs and TALENs—among the first widely used tools for genome editing—are fusions of the DNA cleavage domain from the FokI restriction endonuclease and engineered Cys_2-His_2 zinc-finger [6] and TALE DNA-binding proteins [7], respectively. An individual zinc finger binds to 3–4 nucleotides (with occasional target site overlap from an adjacent zinc-finger domain) [8], whereas a TALE domain recognizes only a single base pair (bp) [9, 10], with

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3_19,

the specificity of each TALE conferred almost entirely by two hypervariable amino acid residues in the center of each domain (dubbed the repeat variable diresidues). Importantly, because the FokI nuclease domain cleaves DNA as a dimer, two ZFN and TALEN monomers are required to form an activate nuclease.

Unlike ZFNs and TALENs, CRISPR-Cas systems rely on a single guide RNA (sgRNA) molecule to direct a Cas nuclease, such as Cas9 or Cas13a, to a specific DNA [11] or RNA sequence [12, 13], respectively. CRISPR-based tools can thus be harnessed for genome editing with only minimal molecular engineering [14–17]. For example, the only major requirement for targeting DNA using CRISPR-Cas9 is the presence of a conserved protospacer adjacent motif (PAM) located at the 3′-end of the Cas9 target site [11].

Nuclease-encoding gene(s) can be introduced into cells using viral and nonviral methods. Nonviral delivery, for instance, can be performed using electroporation [18] and nucleofection [19] but also with chemical methods, including cationic liposomes [20] and charged polymers [21]. These approaches are easy to use but limited to certain cell lines or associated with toxicity [22]. In addition to nonviral methods, viral vectors, including lentivirus, adeno-associated virus (AAV), and adenovirus, can be harnessed to facilitate the expression of nuclease-encoding gene(s) in more difficult-to-transfect cell types and in vivo [23]. AAV vectors, in particular, have emerged as effective tools for delivering CRISPR-Cas9 to the central nervous system [24–26], among other tissues [27].

Regardless of the delivery strategy employed, persistent transient expression of nuclease-encoding gene(s) can lead to increased off-target (OT) effects [28]. One approach for improving nuclease specificity is minimizing the amount of time that a nuclease is exposed to the cell. Crucially, when delivered directly into cells as purified proteins, genome-modifying nucleases are rapidly degraded and induce fewer OT modifications compared to methods that rely on transient expression from nucleic acids [29].

Each of the three major genome-editing platforms can be delivered into cells as proteins [30]. ZFNs—which were the first technology used for DNA-free genome editing—can naturally cross cell membranes [29] due to the intrinsic cell-penetrating capabilities of zinc-finger proteins [31, 32] (notably, their internalization efficiency can be increased following the fusion of multiple nuclear localization sequences [NLS] [33]). In addition to this mode of entry, ZFN proteins can be engineered to enter cells via receptor-mediated endocytosis [34] or packaged into retroviral [35] and lentiviral [36] protein-based platforms. Similarly, TALEN proteins can be introduced into cells via lentiviral based particles [36] or cell-penetrating peptides (CPPs), which are tethered onto the TALEN protein surface [37] or genetically fused to their termini [38]. Moreover, both Cas9 and its sgRNA can be delivered to cells as a preformed ribonucleoprotein (RNP) using nucleofection

[39, 40] or transient transfection [41], though CPPs [42], nanoparticles [43–46], and small-molecule-induced osmocytosis [47] can also facilitate RNP delivery. Additionally, RNPs can be easily complemented with highly recombinogenic single-stranded DNA oligonucleotides [48] and AAV donor vectors [48, 49] for homology-directed genome editing, which expands on the capabilities of the protein-based genome-editing toolbox.

Here, we provide detailed step-by-step procedures for manufacturing and delivering genome-modifying proteins, including RNPs, TALEN, and ZFN proteins. We also provide protocols for combining Cas9 RNP with AAV donor vectors for homology-directed repair (HDR)-mediated genome editing. These procedures are broadly applicable and can be used to modify cells within 10–14 days.

2 Materials

2.1 Molecular Cloning

1. Centrifuge.
2. DNA imaging system.
3. UV-Vis spectrophotometer.
4. PCR thermocycler.
5. Sterile 1.5 mL microcentrifuge tubes.
6. 5 mL Polystyrene tubes.
7. Sterile 15 mL conical tubes.
8. Sterile 50 mL conical tubes.
9. 0.2 mL Flat-cap PCR tubes.
10. Bacterial (e.g., pET-28B; EMD Biosciences) and mammalian expression vectors (e.g., pcDNA 3.1; Life Technologies): Note that the expression vectors encoding Cas9 and the TALENs and ZFNs used here can be obtained from J.L.
11. Deionized water (dH_2O).
12. Nuclease-free water.
13. High-fidelity PCR DNA amplification system.
14. *Nco*I restriction enzyme.
15. *Xho*I restriction enzyme.
16. *Xba*I restriction enzyme.
17. Master buffer for restriction enzymes.
18. T4 DNA ligase with buffer.
19. Plasmid purification kit.
20. PCR amplicon purification kit.
21. DNA gel extraction kit.

22. T7 RNA polymerase in vitro transcription kit.
23. Phenol:chloroform:isoamyl alcohol solution (25:24:1, v/v/v) (Sigma-Aldrich).
24. 3 M Sodium acetate stock solution (dissolve 24.6 g of anhydrous sodium acetate powder in 100 mL of dH_2O): Adjust pH to 5.2 and then autoclave. Can be stored at room temperature (RT) for up to 1 year.
25. 1 kilobase (kb) DNA ladder.
26. Agarose.
27. 10× TAE running buffer.
28. 6× DNA-loading dye.
29. Ethidium bromide (*see* **Note 1**).

2.2 Bacterial Culture

1. Kanamycin.
2. LB medium (Gibco).
3. Bacto-agar.
4. LB agar plates (1 L): Add 15 g of bacto-agar to 1 L of LB medium. Sterilize by autoclave and pipet 22 mL of LB agar to a sterile petri dish. Can be stored at 4 °C for up to 1 month.
5. Chemically competent TOP10 cells.
6. Chemically competent BL21 (DE3) cells.

2.3 Protein Purification

1. Sonicator.
2. 0.22 μm Low-protein-binding filter.
3. 0.45 μm Low-protein-binding filter.
4. Polypropylene gravity-flow purification column.
5. Protein concentrator: 10 kDa MWCO.
6. Protein concentrator: 30 kDa MWCO.
7. HEPES stock solution: 200 mM, pH 8.0. Dissolve 52.1 g HEPES sodium salt in dH_2O. Adjust volume to 1 L and pH to 8.0 and autoclave. Store at 4 °C for up to 6 months.
8. Tris–HCl stock solution: 1 M. Dissolve 12.1 g Tris base in 100 mL dH_2O and autoclave. Store at room temperature for up to 1 year.
9. NaCl stock solution: 5 M. Dissolve 29.2 g NaCl in 1 L dH_2O and autoclave. Store at room temperature for up to 1 year.
10. $MgCl_2$ stock solution: 1 M. Dissolve 20.3 g $MgCl_2$-$6H_2O$ in 1 L dH_2O and autoclave. Store at room temperature for up to 1 year.
11. $ZnCl_2$ stock solution: 9 mM. Dissolve 1.23 g $ZnCl_2$ in 1 L dH_2O and autoclave. Store at room temperature for up to 1 year.
12. Glucose.

13. β-Mercaptoethanol (β-ME): 55 mM (*see* **Note 2**).
14. Glycerol stock solution: 50%, v/v. Mix glycerol with dH_2O and autoclave. Store at room temperature for up to 6 months.
15. Protease inhibitor cocktail: 25× Stock solution. Dissolve one protease inhibitor cocktail tablet in 2 mL dH_2O and sterilize using a 0.22 μm filter. Store at −20 °C for up to 1 week.
16. Imidazole stock solution: 2 M. Dissolve 13.6 g imidazole in 100 mL dH_2O and autoclave. Store at room temperature for up to 1 year.
17. D, L-Dithiothreitol (DTT) stock solution: 1 M. Dissolve 1.54 g in 10 mL dH_2O. Prepare 1 mL aliquots. Sterilize using 0.22 μm filter and store at −20 °C for up to 6 months.
18. L-Arg stock solution: 1 M. Dissolve 17.4 g in 100 mL dH_2O. Adjust pH to 7.4 and sterilize using 0.22 μm filter (*see* **Note 3**). Store at room temperature for up to 6 months.
19. Isopropyl-β-D-1-thiogalactopyranoside (IPTG).
20. Nickel agarose resins for purification of His-tagged proteins.
21. Protein-binding buffer (1 L): Add 100 mL of 200 mM HEPES stock solution, 1 mL of 1 M $MgCl_2$ stock solution, 29.2 g NaCl, and 100 mL glycerol into 799 mL dH_2O. For ZFN-binding buffer, add 10 mL of 9 mM $ZnCl_2$ stock solution and an additional 87.7 g NaCl. Adjust volume to 1 L and autoclave. Store at 4 °C for up to 6 months.
22. Lysis buffer: Add 1 mM β-ME and 1× protease inhibitor cocktail to binding buffer before use (*see* **Note 4**). Store at 4 °C for up to 24 h.
23. Wash buffer A: Add 2.5 mL of 2 M imidazole to 1 L of binding buffer. Store at 4 °C for up to 6 months.
24. Wash buffer B: Add 17.5 mL of 2 M imidazole to 1 L of binding buffer. Store at 4 °C for up to 6 months.
25. Elution buffer: Dissolve 20.4 g imidazole into 1 L of binding buffer. Store at 4 °C for up to 6 months.
26. Storage buffer: Add 100 mL of 200 mM HEPES stock solution, 1 mL of 1 M $MgCl_2$ stock solution, 29.2 g NaCl, and 100 mL glycerol into 799 mL dH_2O. For ZFN storage buffer, add 100 mL of 1 M L-Arg. For Cas9 protein storage buffer, add 1 mL of 1 M DTT stock solution. Store at 4 °C for up to 2 weeks.
27. 4–20% Tris-glycine SDS-PAGE gel: Purchased from commercial vendor.
28. 2× Protein-loading dye.
29. Protein marker.
30. Bovine serum albumin (BSA): 10 mg/mL. Dissolve 10 mg BSA powder in 1 mL dH_2O.

31. PCR amplicons carrying designated nuclease-targeting sequence.
32. In vitro cleavage reaction buffer (10×): 500 mM NaCl, 10 mM Tris–HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9. Mix 1 mL of 5 M NaCl solution, 0.1 mL of 1 M Tris–HCl solution, 0.1 mL of 1 M $MgCl_2$ solution, and 0.1 mL of 1 M DTT solution with water up to 10 mL. Adjust pH to 7.9 and autoclave. Store at −20 °C for up to 2 months.

2.4 Mammalian Cell Culture

1. Tissue culture hood.
2. Tissue culture incubator.
3. Hemocytometer.
4. Leica microscope.
5. Sterile serological pipets: 5, 10, and 25 mL.
6. 10 cm Petri dish.
7. Tissue culture plate: 24-well, 12-well, and 6-well plates.
8. Ethanol: 75%. Mix 25 mL dH_2O with 75 mL ethanol. Store at room temperature (25 °C) for up to 1 month (*see* **Note 5**).
9. Complete DMEM medium: Add 50 mL of fetal bovine serum (FBS) and 5 mL of 100 U/mL of penicillin-streptomycin to 500 mL DMEM medium. Store solution at 4 °C for 3 months. Pre-warm the medium to 37 °C before use.
10. Complete RMPI 1640 medium: Add 50 mL of FBS and 5 mL of 100 U/mL of penicillin-streptomycin to 500 mL RPMI 1640 medium. Store solution at 4 °C for 3 months. Pre-warm the medium to 37 °C before use.
11. Trypsin-EDTA, 0.05% (v/v).
12. Dulbecco's phosphate-buffered saline, no calcium, no magnesium (DPBS).
13. CD4+ primary T cells: Purchased from commercial vendor.
14. CD3/CD28 human T-cell activation beads.
15. Recombinant interleukin-2 (rIL-2).
16. Poly-L-lysine solution.
17. Poly-lysine-coated culture plates: Add 250 mL poly-lysine solution to each well of a 24-well culture plate and incubate at 37 °C for 1 h. Remove solution by aspiration and wash twice with 0.5 mL of DPBS. Dry poly-lysine-treated plates at room temperature for 2 h. Store coated plates at 4 °C for up to 1 month (*see* **Note 6**).

2.5 Transfection Reagents

1. 4D-Nucleofector System (Lonza; core unit, cat. no. AAF-1002B; X unit, cat. no. AAF-1002X).
2. P3 Primary Cell 4D-Nucleofector kit (Lonza, cat. no. V4XP-3032).

3. EGFP reporter cell line: House-made HEK293-derived cells (*see* **Note 7**).
4. Sodium phosphate buffer stock solution: 200 mM, pH 5.5. Dissolve 23.4 g sodium phosphate monobasic in 1 L dH_2O. Adjust pH to 5.5 and autoclave. Store at 4 °C for up to 6 months.
5. Sodium hydroxide: 1 M. Dissolve 4 g NaOH in 100 mL dH_2O and autoclave. Store solution at room temperature for up to 6 months.
6. Peptide stock solution of Cys-Npys-modified peptides: 2.3 mM. Dissolve 1 mg peptides in 250 μL dH_2O. Sterilize using a 0.22 μm filter and store at −20 °C for up to 6 months.
7. Genomic DNA extraction kit.
8. T7E1 endonuclease and buffer.
9. PCR TA cloning kit.
10. Flow cytometry buffer: Supplement DPBS with 2% FBS.

3 Methods

3.1 Expression Vector Construction

1. PCR amplify the genes encoding the "left" and "right" ZFN and TALEN monomers or the gene encoding the Cas9 nuclease from *Streptococcus pyogenes* (SpCas9) using primers encoding 5′ *Nco*I and 3′ *Xho*I restriction sites (*see* Table 1). Carry out the PCR reaction using 5 ng of template DNA, 5 μL of 10× PCR buffer, 0.2 μM of each primer, 0.2 mM of dNTPs, and 2 U of high-fidelity DNA polymerase with water to 50 μL. Use the following cycle: initial denaturation at 95 °C for 5 min; 30 cycles of 95 °C for 30 s; 55 °C for 30 s; and 72 °C for 2 min (for ZFNs) or 4 min (for TALENs) or 5 min (for Cas9); final extension at 72 °C for 10 min (*see* **Note 8**).
2. Purify the PCR products from Subheading 3.1, **step 1**, using a PCR purification kit.
3. Separately incubate pET-28b and the nuclease-encoding genes from Subheading 3.1, **step 2**, with restriction enzymes in separate 50 μL reactions containing 2 μg of DNA; 5 μL of 10× restriction enzyme buffer; and 20 U each of *Nco*I and *Xho*I. Incubate at 37 °C for 3 h.
4. Mix the restriction digests from Subheading 3.1, **step 3**, with 10 μL of 6X DNA-loading dye. Resolve the DNA by gel electrophoresis using a 1% agarose gel. Excise the desired band from the agarose gel, and isolate the DNA using a gel extraction kit, following the manufacturer's instructions (*see* **Note 9**). Determine DNA concentration using a spectrophotometer by measuring $Abs_{260} \times 50$ μg/mL.

Table 1
Primer sequences for nuclease cloning

Primer name	Sequence (5′–3′)
*Nco*I ZFN Fwd	aaaCCATGGatgggtcatcatcatcatcacggtggcagcccgaaaaaaaaacgcaaa
*Xho*I ZFN Rev	aaaCTCGAGttaaaagtttatctcgccgtt
*Nco*I TALEN Fwd	aaaCCATGGatgatgggtcatcatcatcatcacggtggcagcgactacaaagaccatgacggt
*Xho*I TALEN Rev	aaaCTCGAGttaaaagtttatctcgccgttatt
*Nco*I SpCas9 Fwd	aaaCCATGGatgggcagcagccccaagaagaagaggaaggtgggcggctccatggataagaaatactca
*Xho*I SpCas9 Rev	aaaCTCGAGttaatgatgatgatgatgggagccgcccactttgcgtttcttttttcggggagccgcc

Note: Restriction sites are capitalized

5. Ligate the purified nuclease-encoding gene(s) from Subheading 3.1, **step 4**, with the digested pET-28 vector from Subheading 3.1, **step 4**, using an insert-to-vector molar ratio of 6-to-1 in a 10 μL reaction containing ~50 ng of digested DNA, 400 U of T4 DNA ligase, and 1 μL of 10× T4 ligase buffer. Incubate at RT for 1 h.
6. Thaw 100 μL of chemically competent TOP10 cells on ice and mix gently with 10 μL of the ligation reaction from Subheading 3.1, **step 5**. Keep cells on ice for 30 min. Incubate the mixture at 42 °C for 60 s using a water batch, and then put the cells on ice for 30 s. Transfer the cells to a culture tube containing 1 mL of LB medium and incubate the cells for 1 h at 37 °C with shaking at 250 rpm. Spread 100 μL of the bacterial cell culture on a LB agar plate with 50 μg/mL kanamycin and incubate overnight at 37 °C.
7. The following day, inoculate 4 mL of LB medium containing 50 μg/mL kanamycin with one colony from the LB agar plate and culture overnight at 37 °C with shaking at 250 rpm.
8. Purify the nuclease expression vectors using a plasmid miniprep kit per the manufacturer's instructions and confirm the identity of the nuclease-encoding gene(s) by DNA sequencing (*see* **Note 10**).

3.2 In Vitro Transcription of the sgRNA

Note that the following procedures must be carried out in RNase-free environment.

1. PCR amplify the sgRNA. Carry out the PCR reaction using 5 ng of template DNA (e.g., an sgRNA-encoding mammalian expression vector); 5 μL of 10× PCR buffer with $MgCl_2$; 0.2 μM each of forward primer (e.g., 5′-GAAATTAATACGACTCACTATAGGNNNNNNNNNNNNNNNNNNGTTTTAGAGCTAGAAATA-3′, T7 promoter is under-

lined; 20 bp targeting sequence donated by N) and reverse primer (5′-AGCACCGACTCGGTGCCA-3′); 0.2 mM dNTP; 2 U of high-fidelity DNA polymerase; and water up to 50 μL. Use the following cycling program: initial denaturation at 95 °C for 5 min; 30 cycles of 95 °C for 30 s; 55 °C for 30 s; and 72 °C for 40 s; final extension at 72 °C for 10 min.

2. Purify the PCR products by gel extraction and determine DNA concentration using a spectrophotometer by measuring $Abs_{260} \times 50$ μg/mL (*see* **Note 11**).
3. Transcribe the sgRNA by a "run-off" reaction using a T7 in vitro transcription kit. Carry out a 20 μL reaction containing 1 μg of purified PCR product; 2 μL of 10× reaction buffer; 7.5 mM of each dNTP; and 1.5 μL of T7 RNA polymerase mix. Incubate at 37 °C for 16 h (*see* **Note 12**).
4. Upon completion of the reaction, add 70 μL of nuclease-free water, 10 μL of 10× DNase I buffer, and 2 μL of DNase I (RNase free). Mix well and incubate for 15 min at 37 °C.
5. Add 80 μL of nuclease-free water and 20 μL of 3 M sodium acetate to the DNase I-treated transcription reaction, followed by 200 μL of phenol:chloroform:isoamyl alcohol solution and vortex.
6. Centrifuge the mixture at 12,000 × *g* for 5 min, and transfer the upper aqueous phase to a new microcentrifuge tube.
7. Add 200 μL of chloroform and centrifuge at 12,000 × *g* for 5 min to remove residual phenol. Transfer the supernatant to a new microcentrifuge tube.
8. Repeat chloroform extraction procedure once more as described in Subheading 3.3, **step 7**.
9. Add 2.5 volumes of ethanol to the chloroform-extracted sgRNA. Store at −20 °C for 12 h.
10. Centrifuge at 12,000 × *g* for 30 min at 4 °C. Remove the supernatant and then add 500 μL of ice-cold ethanol (75%) to remove residual salt.
11. Centrifuge at 12,000 × *g* for 5 min at 4 °C and air-dry the pellet.
12. Dissolve the pellet in 20 μL of nuclease-free water (*see* **Note 13**). Determine sgRNA concentration using a spectrophotometer by measuring $Abs_{260} \times 40$ μg/mL (*see* **Note 14**). Store sgRNA at −80 °C for up to 1 year.

3.3 Nuclease Protein Manufacturing and Quality Control

1. Thaw 50 μL of chemically competent BL21(DE3) cells on ice and mix gently with 100 ng of sequence-verified nuclease-encoding expression vectors from Subheading 3.1, **step 7**. Transform as described in Subheading 3.1, **step 6**.

2. The following day, inoculate 20 mL of LB medium containing 50 μg/mL kanamycin with one colony from the LB agar plate from Subheading 3.3, **step 1**, and culture overnight at 37 °C.
3. The following day, transfer 20 mL of the overnight starter culture from Subheading 3.2, **step 2**, into 1 L of LB medium containing 50 μg/mL kanamycin, 200 mM NaCl, and 0.2% glucose (*see* **Notes 15** and **16**).
4. Induce protein expression once the culture reaches an OD_{600} of 0.8 with 0.1 mM of IPTG (*see* **Note 17**). After induction, culture ZFN- and TALEN-expressing bacterial cultures at RT for an additional 4 h, and Cas9-expressing culture at 18 °C for an additional 12 h (*see* **Note 18**).
5. Harvest cells by centrifugation at 5000 × *g* for 10 min at 4 °C. Discard the supernatant. Do not dry cell pellets (*see* **Note 19**).
6. Resuspend the cell pellets from Subheading 3.2, **step 5**, with 20 mL of lysis buffer (*see* **Notes 20** and **21**). Transfer the suspension to a fresh collection tube.
7. Lyse the cells by sonication using the following settings: 50% power output and 2-min process time with 5-s on and 10-s off intervals (*see* **Note 22**).
8. Centrifuge the cell lysate at 25,000 × *g* for 30 min at 4 °C and transfer the supernatant into a fresh collection tube. Filter the supernatant through a 0.45 μm low-protein-binding filter (*see* **Note 23**).
9. Add 1 mL of nickel agarose resins (50% slurry in 30% ethanol) to the filtered lysate and incubate on a rotisserie or shaking platform for 30 min at 4 °C.
10. Rinse an empty polypropylene gravity-flow purification column with 5 mL of lysis buffer and then transfer the protein-bound slurry into the column with the bottom cap attached.
11. Remove the bottom cap and discard the flow-through. Wash the column with 20 mL of wash buffer A and 5 mL of wash buffer B.
12. Elute the nuclease protein with ten fractions of 0.5 mL elution buffer (*see* **Note 24**).
13. Analyze fractions by SDS-PAGE by mixing 5 μL of each elution with 5 μL of 2× protein-loading dye. Boil samples at 95 °C for 10 min and resolve on a 4–20% Tris-glycine gel. Combine the fractions with the highest purity.
14. Buffer exchange the combined fractions with storage buffer and concentrate the proteins to 400–800 μL using an Amicon Ultra-15 Centrifugal Filter Unit (note: use 10 kDa MWCO for

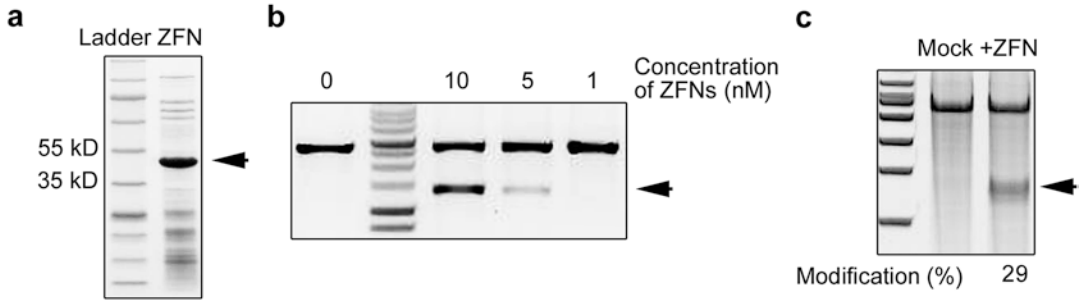


Fig. 1 Genome editing in primary CD4+ T cells using *CCR5*-targeting ZFN proteins. (**a**) SDS-PAGE of purified ZFN protein. Arrowhead indicates the predicted molecular weight of the ZFN protein. (**b**) In vitro cleavage assay assessing ZFN protein activity. (**c**) Frequency of indel formation in primary CD4+ T cells following ZFN protein delivery. (**b**, **c**) Arrowheads indicate expected cleavage product

ZFN proteins, and 30 kDa MWCO for Cas9 and TALEN proteins). Clarify the proteins through 0.22 μm low-protein-binding filters.

15. Visualize purified nuclease proteins and a BSA standard by SDS-PAGE. Determine protein concentration by densitometry using the BSA standard curve (*see* **Note 25**). Sample SDS-PAGE displaying purified nuclease proteins is shown in Fig. 1.
16. Aliquot 250 μL of each concentrated proteins to 1.5 mL microcentrifuge tubes and flash freeze samples using liquid nitrogen and store at −80 °C for up to 1 year (*see* **Note 26**).
17. Determine nuclease protein activity using an in vitro cleavage assay. Carry out the reaction with 1 μL of 10× in vitro assay buffer (*see* **Note 27**), 200 ng of PCR amplicon encoding the nuclease target site, 1 μL of 10 mg/mL BSA, and increasing concentrations (100 nM to 1 nM) of ZFN, TALEN, and RNP (*see* **Note 28**) with dH_2O up to 20 μL. A representative in vitro cleavage assay is shown in Fig. 2.

3.4 Nuclease Protein Delivery

In this section, we describe protocols for delivering nuclease proteins into cells for genome editing. Importantly, owing to their distinct characteristics, ZFNs, TALENs, and RNPs are introduced into cells through different procedures, which are described here. We also provide a protocol for site-specific gene insertion via HDR-mediated genome editing using RNP and AAV donor delivery.

3.4.1 ZFN Protein Delivery to Primary CD4+ T Cells

1. Seed 1 × 10^6 CD4+ T cells into one well of a 24-well plate containing 1 mL of complete RPMI 1640 medium with 25 μL CD3/CD28 human T-cell activation beads and 50 U rIL-2 for stimulation and expansion (*see* **Note 29**). Incubate at 37 °C with 5% CO_2 and full humidity.

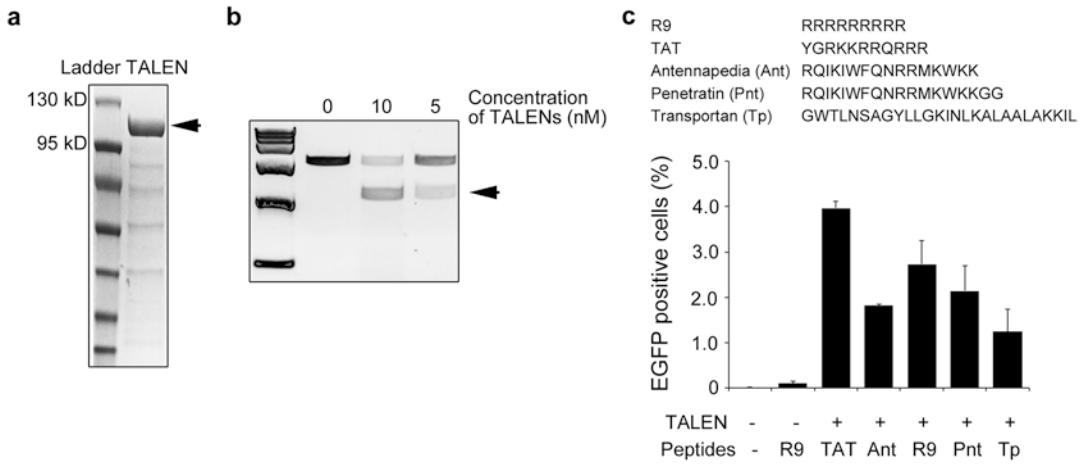


Fig. 2 Genome editing in HEK293T reporter cells using purified TALEN proteins. (**a**) SDS-PAGE of purified TALEN protein. Arrowhead indicates predicted molecular weight of the TALEN protein. (**b**) In vitro cleavage assay assessing TALEN protein activity. (**c**) Flow cytometry analysis of genome editing in HEK293T reporter cells following the delivery of TALEN proteins conjugated with various cell-penetrating peptide (CPP) domains. (**b**, **c**) Arrowheads indicate expected cleavage product

2. At 48 h after activation, harvest cells by centrifugation at 300 × *g* for 10 min, discard supernatant, and wash the cells once with serum-free DMEM medium (*see* **Note 30**).
3. Resuspend 2×10^5 cells in 250 μL serum-free DMEM containing 2 μM each of "left" and "right" ZFN proteins (*see* **Note 31**) supplemented with 90 μM $ZnCl_2$ and 100 mM L-Arg (*see* **Note 32**). Incubate cells at 37 °C for 1 h.
4. Harvest cells by centrifugation at 300 × *g* for 10 min and discard supernatant. Resuspend cells with 500 μL of complete RPMI 1640 medium containing 50 U/mL rIL-2.
5. Incubate cells at 30 °C for 24 h (*see* **Note 33**) and then 37 °C for an additional 24 h.
6. Upon completion of the treatment, centrifuge cells at 300 × *g* for 10 min, discard supernatant, and extract the genomic DNA of ZFN-treated cells using a genomic DNA extraction kit according to the manufacturer's instruction. Store genomic DNA at −80 °C for further analysis.

3.4.2 TALEN Protein Delivery to HEK293 Reporter Cells

1. Seed HEK293 cells onto a 24-well plate at a density of 1 x 10^5 cells per well and incubate at 37 °C with 5% CO_2 and full humidity.
2. At 24 h after seeding, conjugate cell-penetrating peptides (CPP) to the surface of the TALEN proteins by incubating 3.3 μM each "left" and "right" TALEN proteins with 100 μM

of Npys-modified CPPs and 1× complete protease inhibitor cocktail in 100 mM sodium phosphate buffer, pH 5.5, in a 75 μL reaction. Incubate for 2 h at RT (*see* **Note 34**).

3. Adjust the pH of the reaction by adding 7.5 μL of 1 M sodium hydroxide and mix the solution well with 175 μL of serum-free DMEM medium (*see* **Note 35**).
4. Remove medium from each well of cultured cells and wash cells once with serum-free medium (SFM). Gently remove SFM and apply the entire TALEN protein solution onto cells. Incubate cells at 37 °C for 2 h (*see* **Note 36**).
5. Replace medium with complete DMEM and incubate cells at 30 °C for 24 h, and then at 37 °C for additional 24 h for flow cytometry analysis.

3.4.3 RNP Delivery via Electroporation

1. Seed Hepa 1-6 cells onto a 24-well plate at a density of 1×10^5 cells per well. Incubate at 37 °C with 5% CO_2 and full humidity.
2. At 24 h after seeding, form the RNP complex by mixing 16 μg of Cas9 protein with 4 μg of sgRNA (molar ratio of 1:1) in a maximum volume of 2 μL. Incubate at RT for 10 min to allow the complex to form.
3. Remove the medium from the cultured cells, and remove the cells from the plate using trypsin-EDTA (0.05%; v/v). Neutralize the trypsin-EDTA by adding four volumes of complete DMEM medium, and then centrifuge the cells at 200 × *g* for 5 min. Resuspend 2×10^5 cells in 20 μL nucleofection solution containing 16.4 μL of SF Cell Line Solution and 3.6 μL of Supplement 1. Mix the cell suspension with 2 μL RNP complex (*see* **Note 37**).
4. Transfer the cells to a 16-well nucleocuvette strip and electroporate using a 4D-Nucleofector System with the manufacturer's program EH-100.
5. Immediately after nucleofection, add 100 μL of complete DMEM medium to each well and transfer the cells into a 24-well plate containing 0.5 mL of DMEM medium. Incubate at 37 °C with 5% CO_2 and full humidity.
6. At 48 h after electroporation, centrifuge cells at 200 × *g* for 10 min, discard supernatant, and extract the genomic DNA using genomic DNA extraction kit. Store DNA samples at −80 °C for further analysis.

3.4.4 AAV Donor Delivery for HDR-Mediated Genome Editing

1. Clone the AAV donor vector. Step-by-step cloning procedures to generate AAV donor templates have been described by our laboratory [50].
2. Package and purify the AAV donor vector. Step-by-step protocols for manufacturing AAV vectors have also been described by our laboratory [51].

3. Seed cells onto a 96-well plate at a density of 2×10^5 cells per well in serum-containing media with 200 ng/mL of nocodazole (*see* **Note 38**).
4. After 30 min, add AAV donor vector onto cells in the presence of 200 ng/mL of nocodazole (*see* **Notes 39** and **40**).
5. After 16 h, dissociate cells with trypsin and centrifuge at $400 \times g$ for 3 min.
6. Wash cells once with PBS and resuspend in 20 μL of Nucleofector Solution SF with 10 μL of RNP.
7. Immediately after nucleofection, add 100 μL of serum-containing medium to nucleofected wells and transfer cells into a fresh 96-well plate for further analysis.

3.5 Genome-Editing Quantification

In this section, we describe protocols for quantifying editing with the T7E1 endonuclease assay, DNA sequencing, and flow cytometry.

3.5.1 T7E1 Endonuclease Assay

1. Amplify the genome site targeted by the nuclease protein using nested PCR (*see* **Note 41**). First, carry out an external PCR reaction using 1 μg of genomic DNA, 5 μL of 10× PCR buffer, 0.4 μM each of the forward and reverse primers, 0.2 mM dNTP, 5% DMSO, and 2 U of high-fidelity DNA polymerase with water to 50 μL. Use the following cycling program: initial denaturation at 95 °C for 5 min; 20 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 90 s; and final extension at 72 °C for 5 min. Next, carry out an internal PCR reaction using 2 μL of the PCR product from the external PCR reaction as the template, 5 μL of 10× PCR buffer, 0.2 μM each of the forward and reverse primers, 0.2 mM dNTP, and 2 U of high-fidelity DNA polymerase with water to 50 μL. Use the following cycling program: initial denaturation at 95 °C for 5 min; 30 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 40 s; and final extension at 72 °C for 5 min (*see* **Note 38**). Store the PCR product at −20 °C for up to 6 months.
2. Verify amplification by resolving 5 μL of the PCR reaction on a 1% agarose gel as described in Subheading 3.1, **step 4** (*see* **Note 39**).
3. Generate mismatched duplex DNA for the T7E1 assay by denaturing and reannealing 30 μL of the PCR amplicon using the following cycle: 95 °C for 10 min; descend from 95 to 85 °C (at a rate of −2 °C/s); 1 min at 85 °C; 85 to 75 °C (−0.3 °C/s); 1 min at 75 °C; 75 to 65 °C (−0.3 °C/s); 1 min at 65 °C; 65 to 55 °C (−0.3 °C/s); 1 min at 55 °C; 55 to 45 °C (−0.3 °C/s); 1 min at 45 °C; 45 to 35 °C (−0.3 °C/s); 1 min at 35 °C; 35 to 25 °C (−0.3 °C/s); 1 min at 25 °C; hold at 4 °C (*see* **Note 42**).

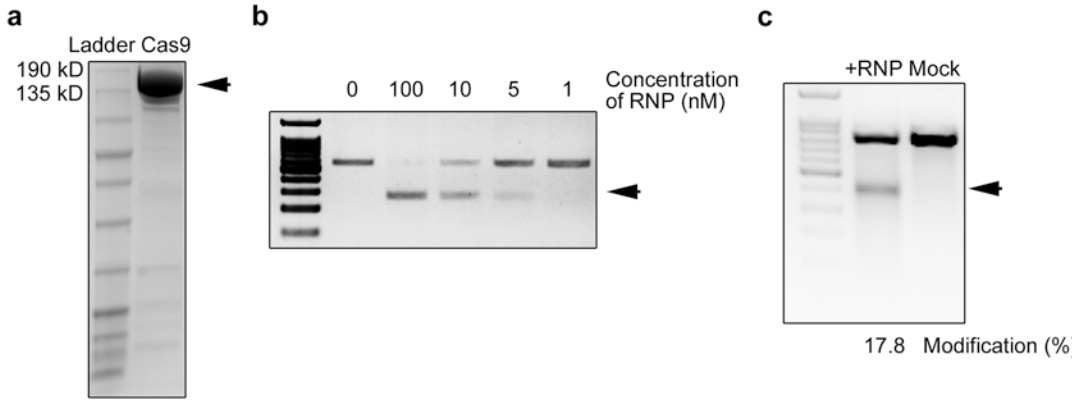


Fig. 3 Cas9-mediated modification of the mouse dystrophin gene in Hepa 1-6 cells. (**a**) SDS-PAGE of purified Cas9 protein. Arrowhead indicates predicated molecular weight of Cas9 protein. (**b**) Assessment of Cas9 RNP activity using in vitro cleavage. Arrow, cleavage product. (**c**) Frequency of indel formation in the mouse dystrophin gene in Hepa 1-6 cells following RNP delivery. (**b**, **c**) Arrowheads indicate expected cleavage product

4. Mix 10 μL of heteroduplex DNA with 5 U of T7E1 endonuclease and 1× cleavage buffer in 20 μL. Incubate the reaction at 37 °C for 30 min.
5. Resolve the T7E1 reaction product on a 1% agarose gel as described in Subheading 3.1, **step 4**.
6. Visualize the agarose gel using an agarose gel-imaging system and measure the intensity of each band. The percent gene modification can be determined by measuring the fraction of parental band cleaved at the anticipated location, as described [52]. An example of a T7E1 analysis is shown in Fig. 3.

3.5.2 DNA Sequence Analysis

1. Ligate the PCR products from Subheading 3.5.1, **step 1**, into TA cloning vector using a PCR TA cloning kit per the manufacturer's instructions. Incubate the reaction at RT for 30 min.
2. Transform the ligation reaction into chemically competent TOP10 cells as described in Subheading 3.1, **step 6**.
3. Pick individual clones from the LB agar plate, and culture them overnight in LB medium. Purify the plasmid by miniprep and submit it for Sanger sequencing using the primer recommended by the PCR TA cloning kit. Alternatively, submit the LB agar plate to a qualified service for direct colony sequencing. An example of Sanger sequencing results confirming genome editing is shown in Fig. 3.

3.5.3 Quantifying Genome Editing in Reporter Cells Using Flow Cytometry

1. Detach cells from Subheading 3.4.2, **step 5**. First, remove the culture medium and wash cells once with DPBS. Next, add 200 μL of trypsin-EDTA to each well, and incubate at 37 °C for 2 min.

2. Resuspend cells in flow cytometry buffer and gently mix by pipetting.
3. Use untreated reporter cells to draw a gate for flow cytometry. Adjust the parameters (i.e., FSC and SSC) according to instrument manual. Specifically, for EGFP reporter cells, use the FITC channel to analyze signal.
4. Perform the analysis on nuclease-treated cells, and quantify the percentage of fluorescent cells. Note that using our previously described EGFP-293 reporter cell line, approximately one-third of nuclease-induced indels restores EGFP fluorescence. An example of this analysis is shown in Fig. 3.

4 Notes

1. Ethidium bromide is potentially carcinogenic and should be handled and deposited of in accordance with institutional guidelines and regulations.
2. βME is hazardous if swallowed or inhaled.
3. If L-Arg is difficult to dissolve, adjust the pH to 7.4 while stirring the solution over gentle heat.
4. Both β-ME and protease inhibitor cocktail should be freshly prepared and added to lysis buffer immediately before use.
5. Ethanol is flammable and should be stored and handled under appropriate conditions.
6. To ensure efficient cell attachment, allow plates to dry for 2 h prior to use.
7. The HEK293 reporter cell line used in our protocol harbors a frame-shift-disabled EGFP gene containing a strategically placed TALEN target site. Following TALEN-induced DNA cleavage, approximately one-third of indels are expected to restore the EGFP reading frame and fluorescence.
8. We recommend optimizing the PCR conditions if nonspecific amplification occurs. Alternatively, nuclease-encoding gene(s) can be sub-cloned from a separate expression vector.
9. To ensure that agarose gel slices are completely dissolved in solubilization buffer, invert the tube every 5 min. If the solution becomes pink or red after the agarose gel slice is dissolved, add 3 M sodium acetate dropwise until the solution becomes yellow.
10. We recommend confirming the full sequence of each cloned nuclease-encoding gene(s). Mutations can dramatically affect the activity of a nuclease.

11. Isolating high-quality DNA template for RNA transcription is essential for in vitro transcription.
12. Short RNA transcripts require a 16-h incubation period for maximum transcription.
13. Resuspending the pellet in nuclease-free water is critical for preventing sgRNA degradation.
14. In our experience, sgRNA concentrations >15 μg/μL are optimal for genome editing.
15. When expressing ZFN proteins, supplement the medium with 90 μM $ZnCl_2$.
16. Addition of 0.2% glucose and 200 mM NaCl to bacterial culture media can prevent leaky expression and inhibit nuclease activity, respectively. Both factors can affect cell growth.
17. Protein yield and purity are dependent on the induction conditions. We recommend monitoring cultures every 30 min until an OD_{600} of 0.8 is reached.
18. Cas9 protein yield can be improved by culturing bacteria at temperatures less than 22 °C after induction.
19. Cell pellets can be stored for up to 1 week at −20 °C or 1 month at −80 °C with no loss of nuclease activity.
20. When purifying ZFN proteins, add 90 μM $ZnCl_2$ to the lysis buffer.
21. Insufficient resuspension of the cell pellet can negatively affect cell lysis by sonication.
22. Prevent overheating by sonicating the cells on ice. Multiple sonication cycles may be necessary to completely lyse cells.
23. If the supernatant goes through the filter too slowly, consider a second centrifugation to remove remaining cell debris in the supernatant.
24. For ZFN proteins only, add 50 μL of 1 M L-Arg to each fraction (final L-Arg concentration: 100 mM) immediately after elution. Failure to add L-Arg can lead to ZFN protein precipitation.
25. To ensure maximum dosage into cells, ZFN and TALEN proteins should be concentrated to ~40 μM and Cas9 protein to ~90 μM (15 μg/μL).
26. Avoid repeating freeze-thawing to prevent damaging proteins.
27. For ZFNs, supplement the reactions with 90 μM $ZnCl_2$ and 100 mM L-Arg.
28. Prepare RNP complex by incubating Cas9 protein and sgRNA at room temperature for 10 min.

29. Activation of $CD4^+$ T cells is required to achieve maximum ZFN-mediated gene modification rates. Proper cell-to-bead ratio is critical for cell activation.
30. Activate T cells for no more than 72 h. Prolonged activation can decrease the efficiency of genome editing.
31. Maximum modification is observed using between 0.5 and 4 μM ZFN proteins. Depending on the purity of ZFN proteins, low levels of cytotoxicity could be observed at high protein concentrations.
32. Lack of L-Arg will dramatically decrease genome-editing efficiency.
33. Transient incubation at 30 °C is necessary for ZFN proteins to achieve optimal rates of gene editing.
34. The molar ratio of Npys-modified CPP to TALEN protein must be between 8-to-1 and 15-to-1 for efficient cellular internalization and genome-editing activity.
35. Depending on the purity of the TALEN proteins and the quality of the peptide, precipitation may be observed after the conjugation reaction. Precipitates can be removed by centrifuging the reaction solution at 5000 × *g* for 2 min.
36. In contrast to ZFN proteins, TALEN proteins should be incubated with cells for ~2 h for efficient internalization.
37. The total volume of the RNP should not exceed 2 μL since a protein solution that exceeds 10% of the nucleofection volume (20 μL) can reduce transfection efficiency.
38. Nocodazole enhances HDR-mediated genome editing by arresting cells at G2/M phase.
39. The transduction efficiency of AAV can vary depending on the target cell type and the AAV serotype being used. It may be necessary to screen for the optimal AAV variant.
40. High doses of AAV (multicity of infection >100,000) are necessary for efficient transduction in vitro.
41. Nested PCR was designed to yield ~1 kb and ~300–500 bp amplicons for the external and internal reactions, respectively.
42. The absence of PCR side products is critical for preventing off-target heteroduplexation of DNA.

Acknowledgments

This work was supported by Natural Science Foundation of China (No. 31600686 to J.L.) and ShanghaiTech University (Startup fund to the Laboratory of ADC Chemistry).

References

1. Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. Genetics 161:1169–1175
2. Santiago Y, Chan E, Liu PQ, Orlando S, Zhang L, Urnov FD, Holmes MC, Guschin D, Waite A, Miller JC, Rebar EJ, Gregory PD, Klug A, Collingwood TN (2008) Targeted gene knockout in mammalian cells by using engineered zinc-finger nucleases. Proc Natl Acad Sci U S A 105:5809–5814
3. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. Science 300:763
4. Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. Science 300:764
5. Urnov FD, Miller JC, Lee YL, Beausejour CM, Rock JM, Augustus S, Jamieson AC, Porteus MH, Gregory PD, Holmes MC (2005) Highly efficient endogenous human gene correction using designed zinc-finger nucleases. Nature 435:646–651
6. Kim YG, Cha J, Chandrasegaran S (1996) Hybrid restriction enzymes: zinc finger fusions to *Fok* I cleavage domain. Proc Natl Acad Sci U S A 93:1156–1160
7. Miller JC, Tan S, Qiao G, Barlow KA, Wang J, Xia DF, Meng X, Paschon DE, Leung E, Hinkley SJ, Dulay GP, Hua KL, Ankoudinova I, Cost GJ, Urnov FD, Zhang HS, Holmes MC, Zhang L, Gregory PD, Rebar EJ (2011) A TALE nuclease architecture for efficient genome editing. Nat Biotechnol 29:143–148
8. Gersbach CA, Gaj T, Barbas CF 3rd (2014) Synthetic zinc finger proteins: the advent of targeted gene regulation and genome modification technologies. Acc Chem Res 47:2309–2318
9. Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. Science 326:1501
10. Boch J, Scholze H, Schornack S, Landgraf A, Hahn S, Kay S, Lahaye T, Nickstadt A, Bonas U (2009) Breaking the code of DNA binding specificity of TAL-type III effectors. Science 326:1509–1512
11. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science 337:816–821
12. Abudayyeh OO, Gootenberg JS, Konermann S, Joung J, Slaymaker IM, Cox DB, Shmakov S, Makarova KS, Semenova E, Minakhin L, Severinov K, Regev A, Lander ES, Koonin EV, Zhang F (2016) C2c2 is a single-component programmable RNA-guided RNA-targeting CRISPR effector. Science 353:aaf5573
13. Abudayyeh OO, Gootenberg JS, Essletzbichler P, Han S, Joung J, Belanto JJ, Verdine V, Cox DBT, Kellner MJ, Regev A, Lander ES, Voytas DF, Ting AY, Zhang F (2017) RNA targeting with CRISPR-Cas13. Nature 550:280–284
14. Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA, Zhang F (2013) Multiplex genome engineering using CRISPR/Cas systems. Science 339:819–823
15. Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE, Church GM (2013) RNA-guided human genome engineering via Cas9. Science 339:823–826
16. Cho SW, Kim S, Kim JM, Kim JS (2013) Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nat Biotechnol 31:230–232
17. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J (2013) RNA-programmed genome editing in human cells. Elife 2:e00471
18. Chu G, Hayakawa H, Berg P (1987) Electroporation for the efficient transfection of mammalian cells with DNA. Nucleic Acids Res 15:1311–1326
19. Hamm A, Krott N, Breibach I, Blindt R, Bosserhoff AK (2002) Efficient transfection method for primary cells. Tissue Eng 8:235–245
20. Felgner PL, Gadek TR, Holm M, Roman R, Chan HW, Wenz M, Northrop JP, Ringold GM, Danielsen M (1987) Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc Natl Acad Sci U S A 84:7413–7417
21. Boussif O, Lezoualc'h F, Zanta MA, Mergny MD, Scherman D, Demeneix B, Behr JP (1995) A versatile vector for gene and oligonucleotide transfer into cells in culture and in vivo: polyethylenimine. Proc Natl Acad Sci U S A 92:7297–7301
22. Roursgaard M, Knudsen KB, Northeved H, Persson M, Christensen T, Kumar PEK, Permin A, Andresen TL, Gjetting T, Lykkesfeldt J, Vesterdal LK, Loft S, Moller P (2016) *In vitro* toxicity of cationic micelles and liposomes in cultured human hepatocyte (HepG2) and lung epithelial (A549) cell lines. Toxicol In Vitro 36:164–171
23. Liu J, Shui SL (2016) Delivery methods for site-specific nucleases: achieving the full poten-

tial of therapeutic gene editing. J Control Release 244:83–97

24. Swiech L, Heidenreich M, Banerjee A, Habib N, Li Y, Trombetta J, Sur M, Zhang F (2015) *In vivo* interrogation of gene function in the mammalian brain using CRISPR-Cas9. Nat Biotechnol 33:102–106
25. Tervo DG, Hwang BY, Viswanathan S, Gaj T, Lavzin M, Ritola KD, Lindo S, Michael S, Kuleshova E, Ojala D, Huang CC, Gerfen CR, Schiller J, Dudman JT, Hantman AW, Looger LL, Schaffer DV, Karpova AY (2016) A designer AAV variant permits efficient retrograde access to projection neurons. Neuron 92:372–382
26. Gaj T, Ojala DS, Ekman FK, Byrne LC, Limsirichai P, Schaffer DV (2017) In vivo genome editing improves motor function and extends survival in a mouse model of ALS. Sci Adv 3:eaar3952
27. Gaj T, Epstein BE, Schaffer DV (2015) Genome engineering using adeno-associated virus: basic and clinical research applications. Mol Ther 24:58–464
28. Pruett-Miller SM, Reading DW, Porter SN, Porteus MH (2009) Attenuation of zinc finger nuclease toxicity by small-molecule regulation of protein levels. PLoS Genet 5:e1000376
29. Gaj T, Guo J, Kato Y, Sirk SJ, Barbas CF 3rd (2012) Targeted gene knockout by direct delivery of zinc-finger nuclease proteins. Nat Methods 9:805–807
30. Liu J, Gaj T, Yang Y, Wang N, Shui S, Kim S, Kanchiswamy CN, Kim JS, Barbas CF 3rd (2015) Efficient delivery of nuclease proteins for genome editing in human stem cells and primary cells. Nat Protoc 10:1842–1859
31. Gaj T, Liu J, Anderson KE, Sirk SJ, Barbas CF 3rd (2014) Protein delivery using Cys2-His2 zinc-finger domains. ACS Chem Biol 9:1662–1667
32. Gaj T, Liu J (2015) Direct protein delivery to mammalian cells using cell-permeable Cys_2-His_2 zinc-finger domains. J Vis Exp 97:e52814
33. Liu J, Gaj T, Wallen MC, Barbas CF 3rd (2015) Improved cell-penetrating zinc-finger nuclease proteins for precision genome engineering. Mol Ther Nucleic Acids 4:e232
34. Chen Z, Jaafar L, Agyekum DG, Xiao H, Wade MF, Kumaran RI, Spector DL, Bao G, Porteus MH, Dynan WS, Meiler SE (2013) Receptor-mediated delivery of engineered nucleases for genome modification. Nucleic Acids Res 41:e182
35. Bobis-Wozowicz S, Galla M, Alzubi J, Kuehle J, Baum C, Schambach A, Cathomen T (2014) Non-integrating gamma-retroviral vectors as a versatile tool for transient zinc-finger nuclease delivery. Sci Rep 4:4656
36. Cai Y, Bak RO, Mikkelsen JG (2014) Targeted genome editing by lentiviral protein transduction of zinc-finger and TAL-effector nucleases. elife 3:e01911
37. Liu J, Gaj T, Patterson JT, Sirk SJ, Barbas CF 3rd (2014) Cell-penetrating peptide-mediated delivery of TALEN proteins via bioconjugation for genome engineering. PLoS One 9:e85755
38. Ru R, Yao Y, Yu S, Yin B, Xu W, Zhao S, Qin L, Chen X (2013) Targeted genome engineering in human induced pluripotent stem cells by penetrating TALENs. Cell Regen 2:5
39. Kim S, Kim D, Cho SW, Kim J, Kim JS (2014) Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. Genome Res 24:1012–1019
40. Lin S, Staahl BT, Alla RK, Doudna JA (2014) Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. elife 3:e04766
41. Zuris JA, Thompson DB, Shu Y, Guilinger JP, Bessen JL, Hu JH, Maeder ML, Joung JK, Chen ZY, Liu DR (2015) Cationic lipid-mediated delivery of proteins enables efficient protein-based genome editing *in vitro* and *in vivo*. Nat Biotechnol 33:73–80
42. Ramakrishna S, Kwaku Dad AB, Beloor J, Gopalappa R, Lee SK, Kim H (2014) Gene disruption by cell-penetrating peptide-mediated delivery of Cas9 protein and guide RNA. Genome Res 24:1020–1027
43. Wang M, Zuris JA, Meng F, Rees H, Sun S, Deng P, Han Y, Gao X, Pouli D, Wu Q, Georgakoudi I, Liu DR, Xu Q (2016) Efficient delivery of genome-editing proteins using bioreducible lipid nanoparticles. Proc Natl Acad Sci U S A 113:2868–2873
44. Mout R, Ray M, Yesilbag Tonga G, Lee YW, Tay T, Sasaki K, Rotello VM (2017) Direct cytosolic delivery of CRISPR/Cas9-ribonucleoprotein for efficient gene editing. ACS Nano 11:2452–2458
45. Ha JS, Lee JS, Jeong J, Kim H, Byun J, Kim SA, Lee HJ, Chung HS, Lee JB, Ahn DR (2017) Poly-sgRNA/siRNA ribonucleoprotein nanoparticles for targeted gene disruption. J Control Release 250:27–35
46. Lee K, Conboy M, Park HM, Jiang F, Kim HJ, Dewitt MA, Mackley VA, Chang K, Rao A, Skinner C, Shobha T, Mehdipour M, Liu H, Huang W-C, Lan F, Bray NL, Li S, Corn JE, Kataoka K, Doudna JA, Conboy I, Murthy N (2017) Nanoparticle delivery of Cas9 ribonucleoprotein and donor DNA *in vivo* induces

homology-directed DNA repair. Nat Biomed Eng 1:889–901

47. D'Astolfo DS, Pagliero RJ, Pras A, Karthaus WR, Clevers H, Prasad V, Lebbink RJ, Rehmann H, Geijsen N (2015) Efficient intracellular delivery of native proteins. Cell 161:674–690
48. Dever DP, Bak RO, Reinisch A, Camarena J, Washington G, Nicolas CE, Pavel-Dinu M, Saxena N, Wilkens AB, Mantri S, Uchida N, Hendel A, Narla A, Majeti R, Weinberg KI, Porteus MH (2016) CRISPR/Cas9 beta-globin gene targeting in human haematopoietic stem cells. Nature 539:384–389
49. Gaj T, Staahl BT, Rodrigues GM, Limsirichai P, Ekman FK, Doudna JA, Schaffer DV (2017) Targeted gene knock-in by homology-directed genome editing using Cas9 ribonucleoprotein and AAV donor delivery. Nucleic Acids Res 45:e98
50. Rodrigues GMC, Gaj T, Adil MM, Wahba J, Rao AT, Lorbeer FK, Kulkarni RU, Diogo MM, Cabral JMS, Miller EW, Hockemeyer D, Schaffer DV (2017) Defined and scalable differentiation of human oligodendrocyte precursors from pluripotent stem cells in a 3D culture system. Stem Cell Rep 8:1770–1783
51. Gaj T, Schaffer DV (2016) Adeno-associated virus-mediated delivery of CRISPR-Cas systems for genome engineering in mammalian cells. Cold Spring Harb Protoc 2016:pdb prot086868
52. Guschin DY, Waite AJ, Katibah GE, Miller JC, Holmes MC, Rebar EJ (2010) A rapid and general assay for monitoring endogenous gene modification. Methods Mol Biol 649:247–256

INDEX

Jia Liu (ed.), *Zinc Finger Proteins: Methods and Protocols*, Methods in Molecular Biology, vol. 1867,
https://doi.org/10.1007/978-1-4939-8799-3,