VILNIUS UNIVERSITY

Arūnas Šilanskas

RESTRICTION ENDONUCLEASE-TRIPLEX FORMING OLIGONUCLEOTIDE CONJUGATES WITH CONTROLLABLE CATALYTIC ACTIVITY

Doctoral dissertation
Physical science, biochemistry (04 P)

Vilnius, 2012

The work presented in this doctoral dissertation has been carried out at the Institute of Biotechnology, Vilnius University during 2002-2010.

Scientific supervisor:

Prof. dr. **Virginijus Šikšnys** (Vilnius University, physical sciences, biochemistry - 04 P)

VILNIAUS UNIVERSITETAS

Arūnas Šilanskas

KONTROLIUOJAMO AKTYVUMO RESTRIKCIJOS ENDONUKLEAZIŲ-TRIPLEKSĄ FORMUOJANČIŲ OLIGONUKLEOTIDŲ KONJUGATAI

Daktaro disertacija
Fiziniai mokslai, biochemija (04 P)

Vilnius, 2012

Disertacija rengta 2002-2010 m. Vilniaus universiteto Biotechnologijos institute.

Mokslinis konsultantas

Prof. dr. **Virginijus Šikšnys** (Vilniaus universitetas, fiziniai mokslai, biochemija - 04 P)

CONTENTS

LIST OF ABBREVIATIONS

bp	base pair
BSA	bovine serum albumin
DNA	deoxyribonucleic acid
tRNA	transfer ribonucleic acid
DTT	1,4-dithiothreitol
EDTA	ethylenedinitrilotetraacetic acid
GdmCl	guanidinium monochloride
NBB	2-nitrobenzyl bromide
GMBS	N-[γ-maleimidobutyryloxy]succinimide ester
IPTG	isopropyl-β-D-thiogalactopyranoside
Kb	kilobase pair
nt	nucleotide
PAAG	polyacrylamide gel
PAGE	polyacrylamide gel electrophoresis
SDS	sodium dodecyl sulfate
PCR	polymerase chain reaction
REase	restriction endonuclease
HEase	homing endonuclease
ZFN	zinc finger nuclease
ZFP	zinc finger protein
TALE	transcription activator-like effector
TALEN	TALE nuclease
SDS	sodium dodecyl sulfate
TEMED	N,N,N',N'-tetramethylethylenediamine
MES	2–(N-morpholino)-ethanesulfonic acid

Tris	2-amino-2-hydroxymethyl-1,3-propanediol
Wt	wild type
GT	gene targeting
ds	double strand
TFO	triplex forming oligonucleotide
TFS	triplex forming sequence
AUC	analytical ultracentrifugation
CD	circular dichroism
DSB	double strand break
HR	homologous recombination
UV	ultra violet
aa	amino acid

INTRODUCTION

The human genome contains approximately 20,000-25,000 different protein-coding genes (Collins et al, 2004). Simple mutations within the coding region of critical genes can lead to the formation of abnormal proteins, resulting in various diseases (*e.g.* cancer), in failure of an embryo to develop, or premature death. Genetic diseases can only be truly cured *via* restoration of defective gene function. The goal of all gene therapy protocols is to repair the precise genetic defect without additional modification of the genome. One of the most promising strategies is based on homologous recombination, wherein the DNA sequence needed to correct the gene of interest is supplied *in trans* to stimulate site-specific recombination.

Homologous recombination is the ultimate tool for gene replacement, but is limited by low efficiency and reproducibility. Homologous recombination naturally occurs with a low frequency (1 in 10^6 transfected cells), however it is known that DNA double-strand breaks enhance the efficiency of homologous recombination by several orders of magnitude (up to 10,000-fold). However, for *in vivo* applications absolute site-specificity is required. Therefore, gene therapy *via* homologous recombination requires new molecular tools that should be highly specific and rigorously controllable. Engineered, highly specific DNA endonucleases (meganucleases) programmable according to the desired specificity are the key to a wider use of this technology in gene replacement. Meganucleases are by definition sequence-specific endonucleases with large (>12 bp) recognition sites. Because of their high specificity, these proteins are the perfect tools for genome engineering, for they are specific enough to bind and cut only one site in a chosen genome. In nature meganucleases are essentially represented by homing endonucleases (Stoddard, 2011).

Different experimental approaches are currently used for the development of meganucleases for gene therapy:

(i)   homing endonucleases (Arnould et al, 2006; Grizot et al, 2010; Paques & Duchateau, 2007)

(ii)   Zn-finger nucleases (Beumer et al, 2006; Bibikova et al, 2002; Kim et al, 1996; Miller et al, 2007)

(iii) TALE nucleases (Christian et al, 2010; Li et al, 2011)

(iv)   programmable restriction enzymes (Eisenschmidt et al, 2005).

In this work we have focused on the development of highly specific and rigorously controllable meganucleases using restriction enzyme-triple helix forming oligonucleotide (TFO) conjugates. In the restriction endonuclease (REase)-TFO conjugates TFO provides specificity for the extended recognition site through the triple helix formation (Fox, 2000) and addresses restriction enzyme to a particular target site where it introduces a double stranded break. A proof of principle for the development of restriction enzyme-TFO conjugates has been provided before using engineered single-chain PvuII variant (Eisenschmidt et al, 2005). Here we experimentally explored the possibility to extend the repertoire of available restriction enzymes and employ orthodox restriction enzymes for generation of TFO conjugates. To this end we used REases MunI and Bse634I, which were structurally and biochemically characterized before in our laboratory (Deibert et al, 1999; Grazulis et al, 2002; Sasnauskas et al, 1999; Stakenas et al, 1992; Zaremba et al, 2005; Zaremba et al, 2006). MunI is an orthodox type IIP restriction enzymes which recognizes and cleaves C/AATTG sequence ("/" designates a cleavage site) in both DNA strands to generate the double-strand break in presence of Mg^{2+} ions (Deibert et al, 1999; Stakenas et al, 1992). It is arranged as a dimer and each monomer contains a single active site, which act on one DNA strand. Bse634I restriction endonuclease is a type IIF restriction enzyme which recognizes and cleaves R/CCGGY sequence where R=A or G, Y=C or T (Repin et al, 1995). Bse634I is arranged as a tetramer (Grazulis et al, 2002) and requires two cognate DNA copies for optimal activity (Zaremba et al, 2005). However, it was demonstrated that the R226A replacement converts protein into a monomer

which is still able to form a functional Bse634I dimer at increased protein concentrations (Zaremba et al, 2006). This concentration depending monomer-dimer equilibrium provides an attractive possibility to generate controllable restriction endonuclease-TFO conjugates for gene editing.

The specific aims of this study were:

(i) To explore the possibility to regulate the catalytic activity of the MunI restriction enzyme by different internal and external triggers.

(ii) To generate the MunI-TFO and Bse634I-TFO conjugates with a controllable catalytic activity as potential tools for gene therapy experiments.

Scientific novelty. We provide the first demonstration that DNA cleavage activity of the caged MunI-TFO conjugate can be spatially and temporally regulated. This finding opens new possibilities for engineering controllable meganucleases for gene therapy. We also generated Bse634I-TFO conjugate and demonstrated that two monomers assemble into the active dimer on the DNA providing a possible alternative for the catalytic module of the zinc finger nuclease. Moreover, in contrast to the FokI non-specific catalytic domain in zinc finger nucleases, Bse634I retains specificity for the cognate site and therefore is less prone to the off-site cleavage.

Practical value. In the present study we provide proof-of-concept demonstrations of two alternative strategies to control the DNA cleavage activity of restriction endonuclease-TFO conjugates. In the first approach we successfully combined the restriction endonuclease photocaging and TFO-coupling to generate a photoswitchable MunI-TFO conjugate. The caged MunI-TFO conjugate has two important advantages in comparison to the scPvuII-TFO conjugate described in the study by Eisenschmidt and colleagues (Eisenschmidt et al, 2005). First, the preincubation of caged REase-TFO conjugate with target DNA can be performed in any buffer that is compatible with triple helix formation. In particular, the reaction buffer may contain Mg^{2+} ions present in cellular milieu that would otherwise induce DNA binding and

cleavage of unaddressed sites by the restriction endonuclease module of the REase-TFO conjugate. Second, it employs the native homodimeric form of an orthodox restriction endonuclease MunI. Given the wide choice of structurally characterized homodimeric REases, TFO-coupling combined with photocaging of essential functional regions of homodimeric REases (e.g., dimerization interface or DNA binding surface) has the potential of becoming a general approach to targeted DNA cleavage. The same strategy could be also successfully used to regulate other DNA interacting proteins in epigenetics.

The alternative strategy presented here makes use of attenuated dimerization of the monomeric mutant of restriction endonuclease Bse634I. Two nuclease monomers dimerize and cut DNA only when are brought into close proximity by the site-specific interactions of the DNA binding modules (TFO). Triggering of DNA cleavage activity by Bse634I upon triple helix formation should enable addressed DNA cleavage without a separate preincubation step given the reaction buffer is compatible with triplex formation and DNA cleavage.

The major findings presented for defence in this thesis:

- Catalytic activity of the MunI restriction endonuclease can be regulated by the red-ox switch introduced at the dimer interface.

- Catalytic activity of the MunI restriction endonuclease can be regulated by the photoremovable cage compound introduced at the dimer interface.

- Engineered caged MunI-TFO conjugate allows spatial and temporal control of the MunI cleavage activity.

- Bse634I-TFO conjugate can be targeted to a particular site and enable addressed DNA cleavage without a separate preincubation step given the reaction buffer is compatible with triplex formation and DNA cleavage.

1. LITERATURE OVERVIEW

1.1. Genetic diseases

Diseases which are caused by any genetic disorders are called genetic diseases. Genome changes leading to the genetic disorders can vary from a point mutation in a single gene to the addition or loss of an entire chromosome. These genetic traits can be inherited or acquired and are divided into monogenetic and polygenetic diseases.

Monogenetic diseases. A disease caused by a single gene defect is called a monogenetic disease. There are thousands of human monogenetic diseases, which are caused by mutations in the DNA sequence (Kelly, 2007). The illness emerges when a mutated gene produces either nonfunctional protein or a protein variant that can initiate an immune response. All monogenetic diseases are further classified into recessive disorders, dominant disorders and X-linked disorders (Kelly, 2007). Recessive disorder occurs only when a defective gene is inherited from both parents. The example of this class is a cystic fibrosis that occurs 1 in 2000-3000 of Caucasians newborns (Accurso, 2006). If just one gene is recessive another normal copy masks it to prevent a disease. Another class of monogenetic disease is a dominant disorder. It occurs when just one dominant gene is defective, like in the case of Huntington's disease (Walker, 2007). X-linked disorder contains gene defects on the X chromosome and mostly appears in males. Females carry two X chromosomes therefore a second normal gene can mask a defective one. Monogenetic diseases are suitable model systems for the development of gene therapy treatment (Yagi et al, 2011; Snyder et al, 1997; Southwell & Patterson, 2011), as genetic disorders have to be repaired just in a single gene.

Polygenetic diseases. Diseases that are influenced by genetic disorders of multiple genes and various environmental circumstances are called polygenetic diseases (Kelly, 2007; Kresina, 2001). Since this type of diseases is not caused

by a single gene mutation, these diseases are not strongly inherited down families, although a weak inheritance has been observed. Many of the well known diseases belong to this class of genetic disorders including cancer, heart disease, hypertension, diabetes, asthma, obesity, mental retardation, mood disorder or autoimmune diseases. Polygenetic diseases are caused by a complex of factors and they are difficult to study, however, identification of the genetic nature of polygenetic disease could facilitate their treatment.

Table 1. The most frequent monogenetic and polygenetic diseases.

Monogenetic diseases	Polygenetic diseases
Thalassaemia	Cancers
Haemophilia	Autoimmune diseases
Sickle cell anemia	Hypertension
Cystic Fibrosis	Diabetes
Fragile X syndrome	Obesity
Huntington's disease	Heart disease
Tay sachs disease	Atherosclerosis

1.2.   Gene therapy

The use of genetic manipulation for the therapy of human diseases is a new and rapidly evolving field of both the basic science and the clinical medicine (Kresina, 2001). To correct various genetic defects several approaches are developed: (i) gene insertion – a normal gene is inserted into the genome to replace a nonfunctional gene; (ii) gene modification – the defective gene is swapped with the normal gene using recombination; (iii) gene surgery – the defective gene is repaired through the reverse mutation to the normal gene; and (iv) gene regulation – alteration of the degree to which the gene may be turned on or off (Kelly, 2007).

1.2.1.   Types of the gene therapy

There are two types of the gene therapy to supplant the genes that are dysfunctional: somatic gene therapy and germ-line gene therapy (Kelly, 2007).

Somatic gene therapy is the most actively developing type of the gene therapy. To correct genetic disorders DNA is introduced to the tissue or to the cells therefore genetic displacements are not inherited. There are two types of somatic gene therapy: *ex vivo* and *in vivo* (Figure 1). *Ex vivo* gene delivery is performed by taking the cells outside the organism, followed by their modification *in vitro*, cultivation and putting them back into the organism. This method is more problematic than *in vivo* since it requires a mitotic cell population, a tissue culture method and cell transplantation technology; however, gene transfer efficiency is high. *In vivo* gene delivery could be performed directly to the target cells or tissues using various DNA delivery systems (vectors). Although this method is less developed and it could be more promising in the future if new better vectors would be created.
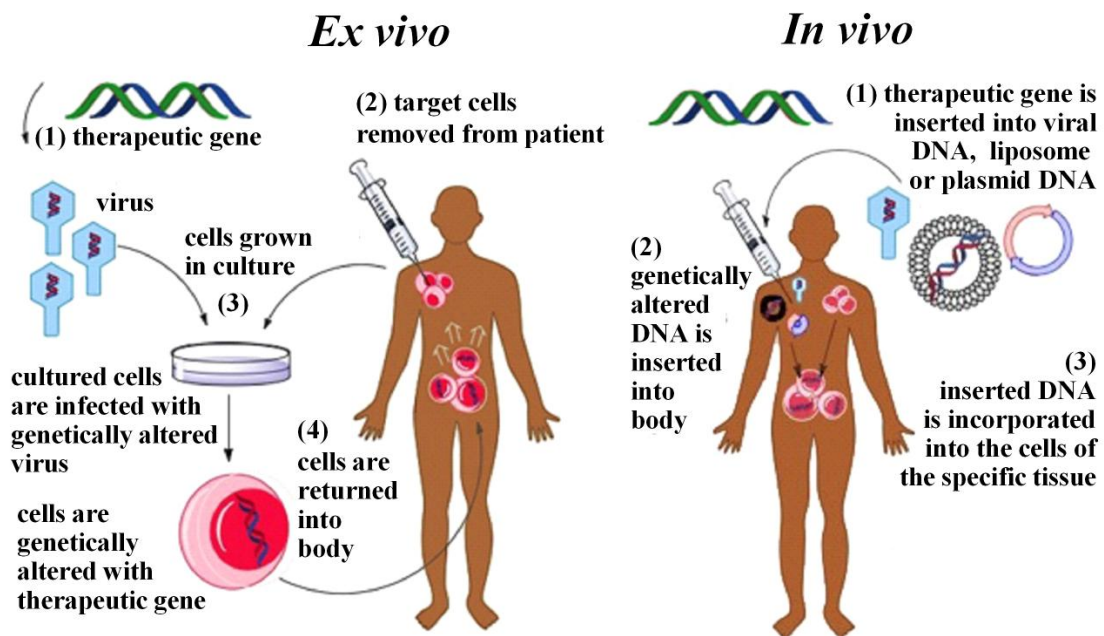
Figure 1. The scheme of *ex vivo* and *in vivo* gene therapy. The picture adopted from http://gene-therapy.yolasite.com/process.php.

Germ-line gene therapy includes insertion of the gene into the reproductive cells or into embryos. In theory, this method offers a possibility to treat many

genetic diseases and prevent their heredity in the subsequent generations. However, this type of therapy is ethically unacceptable because of the possibility to control human development.

1.2.2. Viral vectors used in gene therapy

To perform genetic manipulations often a large piece of genetic material is required to transfect the cells. To facilitate this process viruses are efficient carriers because of their ability to enter the cell and carry DNA. Naturally, viruses are cellular parasites therefore they should be adapted for gene therapy experiments by removing ability to replicate their own genome. Herewith, an inserted therapeutic gene must allow the formation of the capsid that should enter the cell. To perform the delivery of the corrected gene a number of different viruses are used as vectors.

Retroviruses. In contrast to the host's, whose genome is DNA, retroviruses contain RNA genome. After the injection of virus RNA into the cell, a reverse transcriptase synthesizes a double-strand DNA copy and an integrase enzyme incorporates it into the host's genome. The retroviruses work in the actively dividing cells however the cells of the body do not divide often. Therefore, this vector is used primarily *ex vivo*. The cells are taken from the body and virus vector with the correct gene can be inserted. Afterwards these cells are grown and replicated until the sufficient number is achieved and subsequently they are injected into the bloodstream (Kelly, 2007). Since the retroviruses can infect just actively dividing cells they cannot be used with brain or muscle cells. However, this feature could be used to fight against the actively dividing cancer cells, by introducing the "killing" gene using retroviral vector (Moolten, 1994).

Adenoviruses. These viruses contain double-stranded DNA genome and cause respiratory, intestinal, and eye infections in humans (Carr et al, 2011). Adenoviruses enter the cell by endocytosis and stay in the cytoplasm. Since they do not incorporate their DNA into the host genome they can infect all

17

cells including non-dividing ones however the expression time of the delivered gene is quite short.

Adeno-associated viruses (AAV). These viruses contain a single-stranded DNA genome and are capable to incorporate DNA at the specific site of the chromosome 19 (Grieger & Samulski, 2005). AAV vectors are effective since they cause no disease, no significant immune response and could infect dividing and non-dividing cells. However the production of AAV vector is complicated and the expression of the inserted gene is low.

Herpes simplex viruses. There are two types of the herpes simplex virus – HSV1 and HSV2 (Jenkins & Turner, 1996). The genome of this virus contains double-strand DNA. Advantage of the HSV vector is the ability to deliver of DNA up to ten times size of the above mentioned virus vectors. However, this vector like most other virus vectors causes an immune response.

1.2.3.   Nonviral vectors used in gene therapy

Delivery methods that do not include any viral particles are called nonviral methods. The simplest of them is a direct injection of the naked DNA into local tissues. However physical and chemical approaches have been also used to improve the efficiency of the gene transfer (Nishikawa & Huang, 2001).

Naked DNA injection. Injection of naked DNA plasmid is the simplest way to deliver DNA. It was shown that such transfection could be performed to the skeletal muscle (Wolff et al, 1990), liver (Hickman et al, 1994), thyroid (Sikes et al, 1994), heart muscle (Ardehali et al, 1995), brain (Schwartz et al, 1996) and urological organs (Yoo et al, 1999). It is possible to inject the unlimited amount of DNA; however, the transfer rate to the cell is low. The lifetime of the injected DNA is short because of enzymatic degradation, therefore weakly interacting polymers, such as polyvinylpyrrolidone can improve efficiency of the gene transfer (Mumper et al, 1996).

Gene gun. In this method gold particles coated with DNA are transferred (fired/shot) into the target tissues or cells by using high-pressure gas (Kelly,

18

2007). This method allows penetration of DNA directly through the cell membrane avoiding enzymatic degradation in lysosomes, however, the gene expression level is low and short.

Electroporation. Controlled electrical pulses could create permeability channels in the cell membrane to allow charged molecules enter inside the cell (Mir et al, 1988). This method has been used to introduce plasmid DNA into the skin (Titomirov et al, 1991), melanoma (Rols et al, 1998), liver (Heller et al, 1996) or muscle (Aihara & Miyazaki, 1998). Electroporation results in up to 1000-fold higher efficiency of the gene expression comparing with the injection of the naked DNA. The advantages of this method are simplicity and easy preparation of the DNA plasmid comparing with other vectors, a possibility to perform a local transfection in the proper place and no limitation of the gene size. However, the efficiency of the electroporation is lower than that of the viral vectors and the gene expression is short because the delivered DNA does not integrate into the genome.

Cationic lipid (liposome). Liposomes can separate hydrophilic or hydrophobic molecules from the solution. They are formed by self-assembled dissolved lipid molecules containing a hydrophilic head and hydrophobic tail. Mixing of the plasmid DNA or proteins with the cationic lipids decreases its negative charge and facilitates the transfer through the cell membrane (Geel et al, 2009; Nishikawa & Huang, 2001). Addition of the neutral lipids to the cationic lipid-DNA complex facilitates the release of the plasmid DNA from the endosome (Farhood et al, 1995). This method could be used for the plasmid DNA delivery to lung, brain, tumor or skin.

Cationic polymer. Cationic polymers such as poly-L-lysine (PLL), poly-L-ornithine, polyethylenimine (PEI) chitosan or starburst dendrimer condensate DNA more effectively than cationic lipids and could be also used for DNA delivery (Nishikawa & Huang, 2001). Similar to the cationic lipids, these polymers enhance the cellular uptake of the DNA plasmid by nonspecific adsorptive endocytosis (Mastrobattista & Hennink, 2012; Sun & Zhang, 2010).

19

1.3.   Nucleases for gene targeting

Gene targeting is a technique of the gene therapy that uses the homologous recombination (HR) to modify a target gene. This method is one of the most promising strategies for the genome editing in various types of eukaryotic cells. To achieve replacement of the deficient gene through the HR, the correct gene sequence should be supplied *in trans*. Alternatively, homologous recombination can also be used for the gene inactivation, insertion and deletion (Mombaerts et al, 1991; Smithies et al, 1985). The major limitation of this approach is extremely low frequency ($\sim 10^{-6}$) of natural recombination between an exogenous DNA and the homologous chromosomal target in mammalian cells. There are three strategies that could improve efficiency of HR: (i) increased homology stretches between two DNA molecules, (ii) overproduction of the active recombinase, (iii) an artificial double strand break (DSB) in both the exogenous and endogenous DNA (Lanzov, 1999). The last strategy looks the most promising. It is known that the frequency of HR can be greatly increased by five log-orders (up to $\sim 10^{-1}$) by introducing a recombinogenic DSB at or near the targeted gene (Szczepek et al, 2007). Therefore, site-specific nucleases capable of cleavage at unique sites in the eukaryotic genome hold great potential for the genetic studies, biotechnology and gene therapy.

Currently, several classes of natural and artificial nucleases are available for highly specific DNA cleavage (Silva et al, 2011). Naturally occurring HEases, which recognize long DNA sequences (up to 40 base pairs, Figure 2A), have been applied to the gene editing in numerous experimental designs and cell types (Arnould et al, 2006; Grizot et al, 2010; Paques & Duchateau, 2007). A number of artificial nucleases also have been developed for the purpose of the gene targeting. ZFN are created by fusing zinc finger motifs, which serve a DNA recognition module, to the nonspecific DNA cleavage domain of the FokI restriction endonuclease (Beumer et al, 2006; Bibikova et al, 2002; Kim et al, 1996; Miller et al, 2007) (Figure 2B). A similar strategy was recently

used to create novel chimeric nucleases consisting of the FokI catalytic domain fused to the TALE proteins (Christian et al, 2010; Li et al, 2011) (Figure 2C) or a cleavage deficient variant of I-SceI homing endonuclease (Lippow et al, 2009). The fourth group of the site-specific genome cleavage reagents is based on the fusion of the triplex forming oligonucleotides (TFO) as DNA binding module to various DNA damaging compounds (Arimondo et al, 2006; Majumdar et al, 2008; Simon et al, 2008) or type II REase (Figure 2D).

(A) Homing endonuclease

(B) Zinc-finger nuclease

(C) TALE nuclease

(D) scPvuII-TFO

TFS

TFO
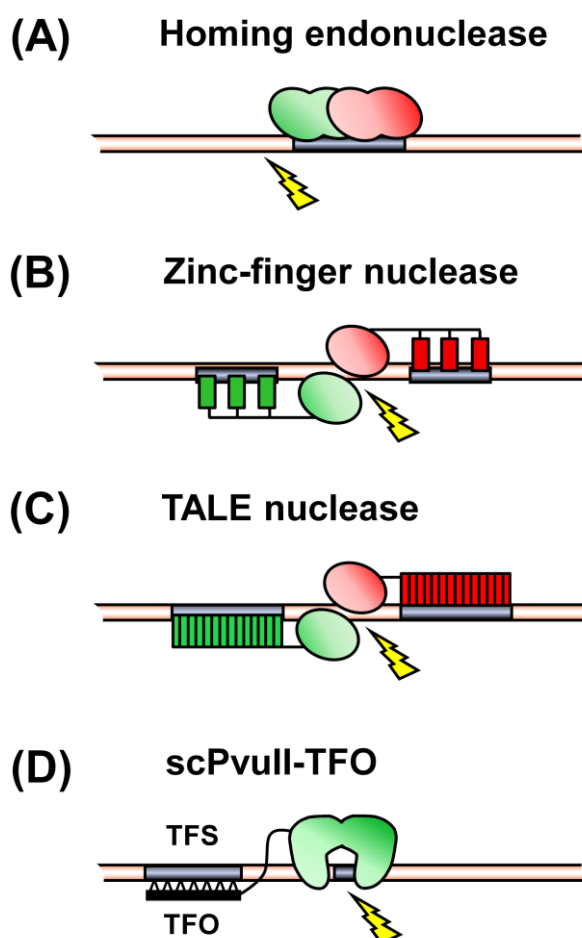
Figure 2. Endonucleases for genome engineering. Naturally occurring homing endonucleases (A) and zinc finger nucleases consisting of the nonspecific FokI DNA cleavage domain fused to DNA binding module composed of several zinc finger motifs (B) were successfully used to increase the efficiency of homologous recombination for the gene correction by cleaving the eukaryotic genome at specific sites. (C) TALE nucleases consisting of the nonspecific FokI DNA cleavage domain fused to DNA binding module composed of TALE protein also successfully increase the efficiency of HR. (D) The REase PvuII – triplex forming oligonucleotide (TFO) conjugate recognizes a bipartite DNA sequence consisting of the triplex forming sequence (TFS) and the restriction endonuclease site.

1.3.1.   Homing nucleases

"Homing" is the transfer of a self-splicing intervening genetic sequence from either an intron or an intein to a cognate allele that lacks that element (Belfort & Perlman, 1995; Belfort & Roberts, 1997; Chevalier & Stoddard, 2001; Dujon, 1989; Lambowitz & Belfort, 1993). The first observation of homing was reported about 40 years ago in the publication coming from the Pasteur Institute (Bolotin et al, 1971). It has been found the dominant inheritance of a genetic marker "Omega" during yeast mating experiments. Later studies showed that "Omega" is located in the intervening sequence known as self-splicing group I intron (Bos et al, 1978; Faye et al, 1979) and it's inheritance is determined by the site-specific homing endonuclease (I-SceI) encoded within the intron sequence (Jacquier & Dujon, 1985). This enzyme specifically binds to the intron-less allele and introduces a double strand break. Then cellular machinery perform repair of the DSB by using homologous recombination between the intron-less allele and the allele containing the intron (Figure 3A). This results in the generation of the second allele containing the intron.

DSB inducing HEases (or meganucleases) are the sequence specific nucleases recognizing long DNA target sites (14-40 bp) and are found in all branches of life (phage, eubacteria, archaea, and eukaryotes) (Stoddard, 2005). Based on the sequence and structure motifs these enzymes are divided into five families: LAGLIDADG, GIY-YIG, HNH, His-Cys box and PD-(D/E)XK (Silva et al, 2011; Stoddard, 2011) (Figure 4). The LAGLIDADG proteins are best characterized in the homing nuclease family.

LAGLIDADG endonucleases exist as either monomers or homodimers (Geese et al, 2003; Moure et al, 2008). Monomeric nucleases, as I-SceI, contain tandem repeat of two LAGLIDADG motifs and recognize fully asymmetric DNA target sites (Lucas et al, 2001).
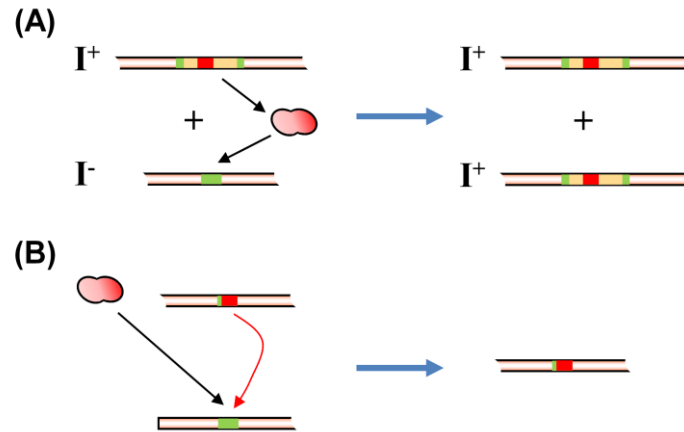
Figure 3. Mechanism of homing. (A) The homing endonuclease (red), encoded by a gene (red bar) within a self-splicing intron or intein (yellow bar), is expressed and cleaves a target site (green bar) of the host gene in the homologous intron-less allele. The resulting double strand break is repaired by the cellular machinery. If homologous recombination is performed by using allele with the intron as template, it results in duplication of the mobile element. (B). Ability to introduce double strand break could facilitate the homologous recombination process that is needed in the gene therapy experiments. Figure adapted from (Stoddard, 2011).
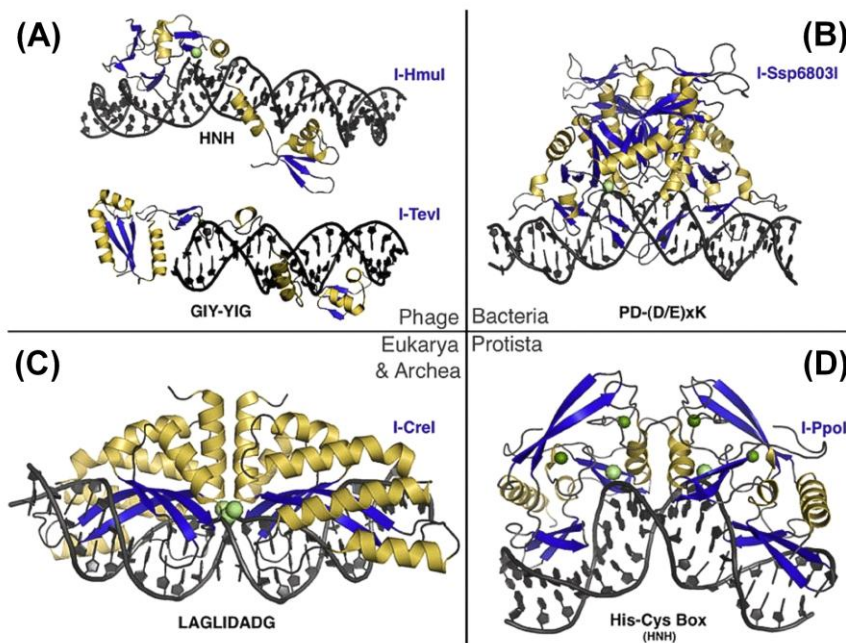
Figure 4. Structural families of the homing endonucleases. (A) The phage-encoded HNH endonuclease I-HmuI (Shen et al, 2004) and GIY-YIG endonuclease I-TevI (Van Roey et al, 2002; Van Roey et al, 2001). These nucleases are monomers consisting of an N-terminal catalytic domain and C-terminal DNA binding region with the C-terminal helix-turn-helix domain. (B) The cyanobacterial PD-(D/E)xK homing endonuclease I-Ssp6803I (Zhao et al, 2007). (C) The algal LAGLIDADG homodimeric homing endonuclease I-CreI (Jurica et al, 1998). (D) The His-Cys box homing endonuclease I-PpoI (Flick et al, 1998). Figure adapted from (Stoddard, 2011).

Homodimers, such as I-CreI, consist of two identical protein subunits, containing a single LAGLIDADG motif and recognize palindromic or pseudopalindromic substrates (Thompson et al, 1992; Wang et al, 1997). Both forms of these enzymes display two unique active sites, responsible for cleavage of each DNA strand. LAGLIDADG enzyme domains form an elongated protein fold that consists of a core fold with mixed α/β topology (Heath et al, 1997). In the domain interface two-helix bundle is strongly conserved between members of LADLIGADG enzymes, enabling swapping of domains from different enzymes to generate endonucleases with new specificities (Figure 5). The N-terminal acidic residues of these helices coordinate divalent cations that are required for catalytic activity. The DNA-binding interface is formed by an antiparallel β-sheet that is composed from the four β-strands of each domain and is well conserved (Bolduc et al, 2003). The amino acid residues from the first two strands of each β-sheet make contacts to the DNA bases in the major groove and the basic residues of the flanking loops make contacts to the phosphodiester backbone of the DNA.
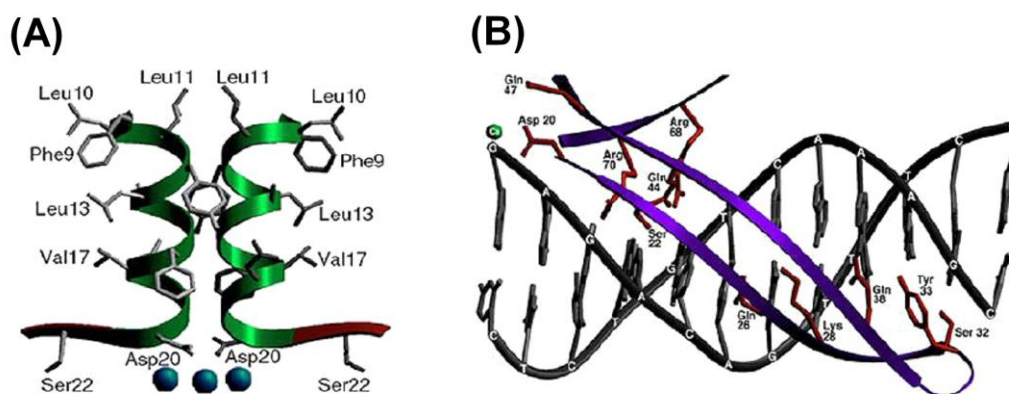
Figure 5. Structural features of LAGLIDADG homing endonucleases. (A) The LAGLIDADG motifs form the helices at the domain interface of the I-CreI structure and serve a similar role in all known LAGLIDADG enzymes. (B) Structure and orientation of the DNA-binding antiparallel β-sheets from I-CreI. Figure adapted from (Stoddard, 2005).

The ability to perform the precise cleavage of rare DNA sequences potentially could be used in the gene therapy experiments to induce DSB to facilitate HR process (Figure 3B). The structural properties of LAGLIDADG

homing endonucleases allow engineering of enzymes with new specificities for the gene targeting applications (Stoddard, 2011): (i) They represent the largest collection of known and characterized HEases with various biological host range (Dalgaard et al, 1997). (ii) The LAGLIDADG family enzymes are highly specific for the 19-22 bp target sites enabling cleavage at a single or few sites in the genome (Chevalier et al, 2003; Gimble et al, 2003; Scalley-Kim et al, 2007). (iii) Their small size and a relatively simple structure together with the long target sites provide a suitable scaffold for engineering of the nucleases with novel specificities. The protein-DNA interface of these enzymes can be engineered to create novel endonucleases with tailored specificities, thereby expanding the list of the available gene targets, yet preserving the tight coupling of the cleavage activity to the DNA recognition inherent to native enzymes (Silva et al, 2011). (iv) DNA cleavage activity of HEases is tightly coupled to the site-specific binding, significantly minimizing the off-site cleavage activity (Chevalier et al, 2004). Domain replacement between different but related enzymes represents another possibility to produce enzymes with novel specificities. In this way, DNA-binding domain of the PI-SceI was swapped with the homolog from *Candida tropicalis* to get a protein with altered specificity (Steuer et al, 2004). The redesigned cleavage specificity of the engineered enzyme confirmed that an exchanged DNA-binding module is responsible for the recognition of DNA sequence. Several studies demonstrated that domains from unrelated LAGLIDADG proteins can be fused to generate chimeric nucleases recognizing corresponding chimeric target sites (Chevalier et al, 2002; Epinat et al, 2003). In this way, domains of the homing endonucleases I-DmoI and I-CreI were fused to get a novel protein that was named H-DreI (formerly name E-DreI) (Figure 6) (Chevalier et al, 2002). Engineering of this enzyme was accomplished by combining computational redesign and an *in vivo* protein folding screen. E-DreI homing endonuclease recognizes a long chimeric target site and cleaves it with the same rate as its parents. Further work with the same homing endonucleases I-DmoI and I-CreI demonstrated the possibility to generate a single chain,

monomeric protein from the homodimer. In this case two copies of the I-CreI gene were connected *via* a linker present in I-DmoI (Epinat et al, 2003).
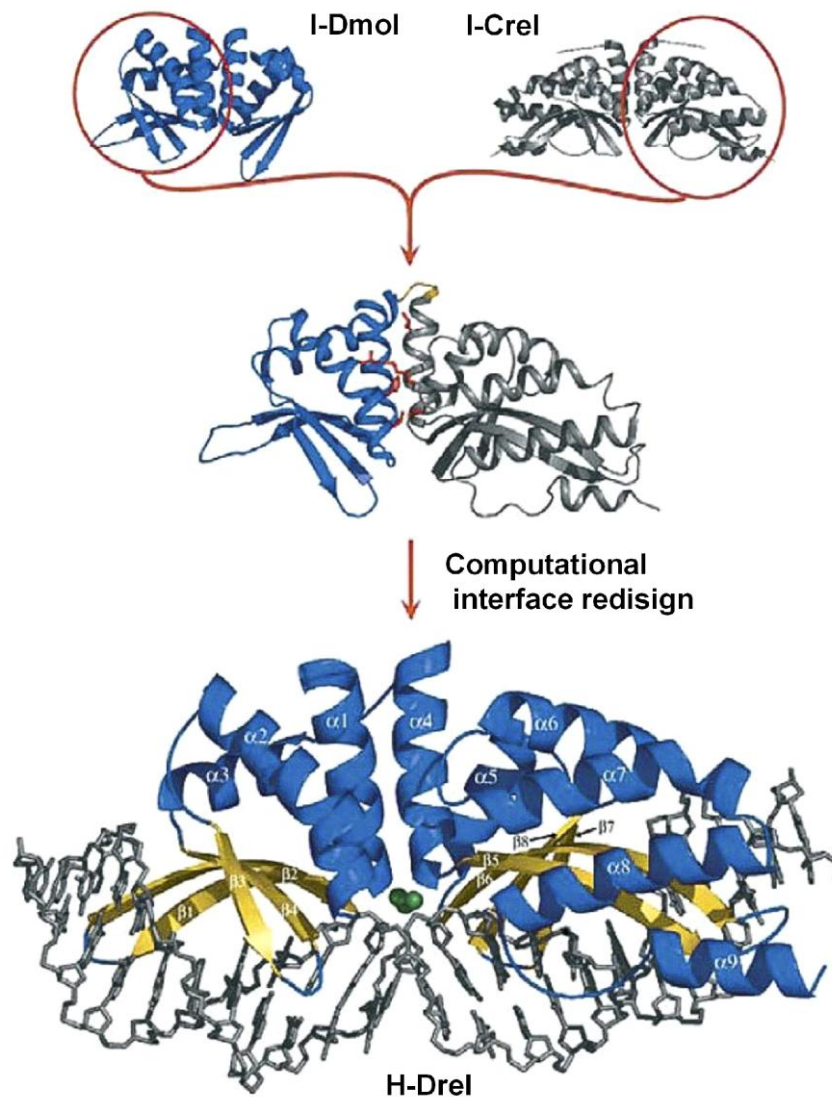
Figure 6. Generation of the artificial homing endonucleases by fusing individual domains of the I-DmoI and I-CreI enzymes. Figure adapted from (Stoddard, 2005).

Hybrid DmoCre enzyme showed the same DNA cleavage specificity as the native I-CreI. Subsequently, the experiments were performed to show interchangeability of the LAGLIDADG helix interaction (Silva & Belfort, 2004). The interface residues of the homodimeric I-CreI were grafted into the corresponding positions in the monomeric I-DmoI. The resulting mutant was endowed with novel nicking activities and oligomeric properties. Moreover, it was shown, that individual domains from the monomeric HEase can be

26

engineered to form stable and functional homodimers (Silva et al, 2006). I-DmoI domain A is able to form a homodimer (H-DmoA) that tightly binds a palindromic DNA sequence derived from the non-palindromic wt I-DmoI target site. Re-engineering of the LAGLIDADG dimerization interface of the H-DmoA resulted in a new homodimer H-DmoC with increased cleavage ability compared to that of H-DmoA. Finally, two copies of the H-DmoC were fused *via* a short peptide linker, yielding a novel active monomeric enzyme H-DmoC2. This example illustrates the modularity of LAGLIDADG HEase scaffolds.

Another technique to generate HEases with novel specificities is the mutation of the individual side chains interacting with the specific DNA target site (Gimble et al, 2003; Seligman et al, 2002; Sussman et al, 2004). The strategy relies on generation of a large number of mutants and performing selection and screening for high affinity DNA binding or/and efficient cleavage activity. Availability of the high resolution protein-DNA crystal structures of LAGLIDADG HEases allows computational redesign of the contact surface at the protein-DNA interface and facilitates mutational screening of protein libraries.

To generate the PI-SceI HEase mutant with a novel specificity a two-hybrid strategy was used to select variants from the randomized expression library that binds to different DNA substrates (Gimble et al, 2003). Selected variants with the altered binding specificity were characterized in the DNA cleavage reactions. Analysis of the cleavage activity revealed that the engineered endonucleases exhibited either a partially relaxed specificity or a marked shift in the target site recognition.

To isolate HEase derivatives with engineered specificities, two alternative bacterial selection strategies, based on the cleavage and elimination of a reporter gene, were used. The first strategy employed a system where the cleavage of the target site resulted in cell conversion from lac^+ to lac^- phenotype. This allowed for selection of desirable activities by screening of the

blue-white colonies (Rosen et al, 2006; Seligman et al, 2002; Sussman et al, 2004). Using this method, mutants with shifted specificities but reduced ability to discriminate between cognate and miscognate sites were obtained. A second bacterial selection strategy used a toxic gene, whose product resulted in a cell death. The cleavage of a homing site in this vector allowed the cell to survive (Doyon et al, 2006; Gruen et al, 2002).
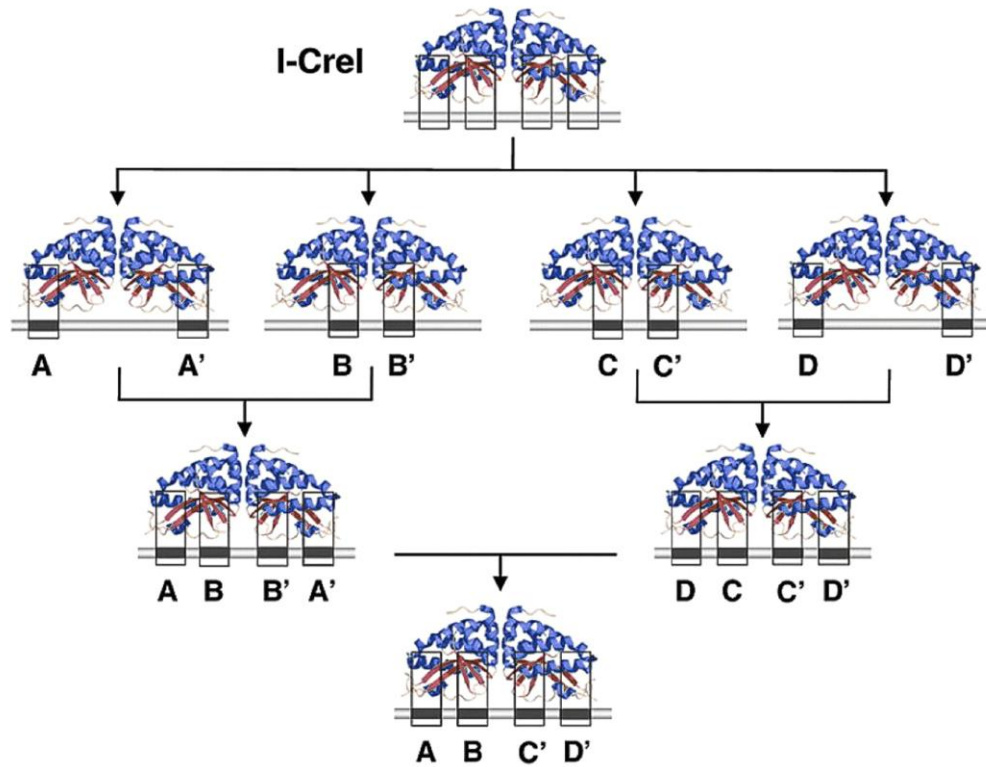
Figure 7. Strategy for the generation of HEases with new specificities. First, a collection of I-CreI mutants with locally altered specificity is generated. Then, combinatorial approach is used to generate homodimers, and after heterodimers with redesigned specificity. Figure adapted from (Smith et al, 2006).
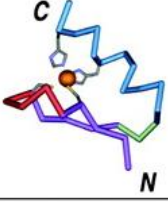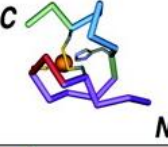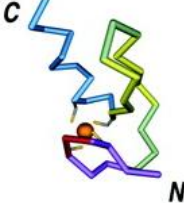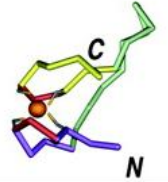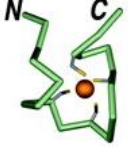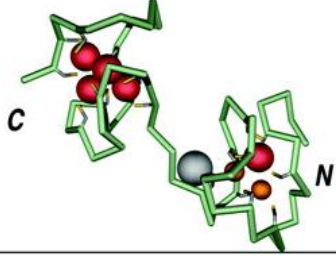
For the screening of mutants with the altered specificity a research company Cellectis developed a eukaryotic system that reports on the generation of DSB induced HR (Arnould et al, 2006; Chames et al, 2005). In this assay two different yeast strains are used. One of these strains contains a gene that determines growth phenotype. This gene is interrupted with an insert containing a desired HEase target site flanked by two direct repeats. Another yeast stain contains the HEase gene. HEase or a library of HEase variants is introduced to the target site by mating. This process is automated therefore the

same HEase library can be screened against multiple DNA target sites. Using this method hundreds of I-CreI mutants with redesigned specificity were identified, catalogued and archived (Arnould et al, 2006). On the other hand, it is enough to have smaller libraries that could be combined to create heterodimeric protein species thus greatly enhancing the number of potential targets (Smith et al, 2006) (Figure 7). Using this approach a few mutants of the HEase I-CreI were selected that were able specifically cleave at the site of mutations in human gene *RAG1* that determines severe combined immunodeficiency disease (SCID) phenotype (Smith et al, 2006) and XPC (xeroderma pigmentosum complementation group C) gene (Arnould et al, 2007).

1.3.2.   Zinc-finger nucleases

Zinc fingers (ZFs) are small protein domains that coordinate zinc ions by cysteine and histidine residues to stabilize the structural fold. The first zinc finger repeat motif was identified in *Xenopus laevis* transcription factor IIIA (Miller et al, 1985) where Zn ion is coordinated by C2H2 residues. Proteins containing zinc fingers interact with nucleic acids and other proteins and are involved in various cellular processes including replication, repair, transcription, translation, metabolism and apoptosis. Depending on the main chain conformation and secondary structure around the zinc-binding site zinc fingers were classified into eight fold groups (Table2) (Krishna et al, 2003). The classical C2H2-like finger domain is the most common structural motif in eukaryotes (Rubin et al, 2000). Tandem of two, three or more fingers usually interact with specific DNA sequences however some C2H2 zinc fingers bind RNA (Searles et al, 2000) or participate in the protein-protein interaction (Wolfe et al, 2000). Nuclear magnetic resonance (NMR) studies revealed the ββα fold of the C2H2 ZF motif (Lee et al, 1989).

Table 2. Structural classification of Zinc fingers (Krishna et al, 2003).

Fold group	Representative structure	Ligand placement	Members in the alignment
1. C2H2 like		Two ligands from a knuckle and two more from the C terminus of a helix	1ncsA, 1tf6A, 1tf6D, 1zfdA, 1ubdC, 2gliA, 1sp2A, 1rmdA, 1znfA, 2adrA, 1aayA, 1sp1A, 1bhiA, 1bboA, 2drpA, 1yuiA, 1ej6C, 1klrA, 1k2fA, 1fv5A, 1fu9A, 1g73C, 1jd5A, 1c9qA, 1e31A
2. Gag knuckle		Two ligands from a knuckle and two more from a short helix or loop	1altA, 1a6bB, 1dsvA, 1dsqA, 1fn9A, 1i3qA
3. Treble clef		Two ligands from a knuckle and two more from the N terminus of a helix	1chcA, 1borA,1jm7A, 1jm7B, 1rmdA, 1fbvA, 1g25A, 1ldjB, 1e4uA, 1dcqA, 1ptqA, 1faqA, 1kbeA, 1e53A, 1dvpA, 1vfyA, 1jocA, 1zbdB, 1fp0A, 1f62A, 1fjfN, 1jj2T, 1ee8A, 1k3wA, 1l2bA, 1ffyA, 1zfoA, 1xpaA, 4gatA, 2gatA, 1gnfA, 1b8tA, 1imlA, 1g47A, 1hcqA, 1kb6A, 1g2rA, 1en7A, 1bxiB, 1ql0A, 1a73A, 1mhdA, 1i3qJ, 1ef4A, 1lnrY, 1i3jA, 1hc7A, 1dgsA, 1lv3A
4. Zinc ribbon		Two ligands each from two knuckles	1jj2Z, 1jj2Y, 1d0qA, 1qypA, 1i50I, 1jj22, 1i5oD, 1qf8A, 1tfiA, 1i50B, 1pftA, 1i50A, 1aduA, 1yuaA, 1dfeA, 1gh9A, 1ileA, 1meaA, 1i50L, 1zinA, 1iciA, 1ma3A, 1a8hA, 1dx8A, 1irnA, 1dxgA, 2occF, 1freA, 1exkA,1b55A, 1f4lA, 1gaxA, 1lnrZ, 1lnr1, 1l1oF, 1e4vA, 1zakA, 1dl6A, 1b71A, 1j8fA, 1kjzA, 1ezvE, 1rfsA, 1g8kB, 1eg9A, 1fqtA
5. Zn2/Cys6		Two ligands from the N terminus of a helix and two more from a loop	1d66A, 1zmeC, 2hapC, 2alcA, 1co4A
6. TAZ2 domain like		Two ligands each from the termini of two helices.	1f81A, 1l8cA, 1wjbA, 1jr3A, 1jr3E
7. Zinc binding loops		Four ligands from a loop	A) 1hsoA, 1e3jA, 1i3qC, 1a5tA, 1cyqA1, 1gpcA B) 1enuA, 1iq8A, 1ia9A, 1cw0A, 1cyqA2, 1ldjB
8. Metallothionein		Cysteine rich metal binding loop	4mt2A

Two cysteines located in the antiparallel β-sheet and two histidines located in a α-helix coordinate a zinc ion thus holding the sheet and the helix together (Figure 8A). This fold is also stabilized by hydrophobic interactions. The first crystal structure of the Zif268 finger (Pavletich & Pabo, 1991) and a later refined structure (Elrod-Erickson et al, 1996) provide the molecular mechanism of DNA recognition. It was shown that adjacent ZFs are structurally independent and are connected by flexible linkers. In the case of Zif268 protein, the DNA-binding domain contains three zinc fingers (Christy et al, 1988) (Figure 8B). The α helix of each ZF contacts 3-4 bp in the major groove and most of the contacts are made to the bases located on one DNA strand. It has been demonstrated that the DNA binding specificity of individual ZFs can be altered by site-direct mutagenesis (Berg, 1988) or by phage display (Choo & Klug, 1994a; Choo & Klug, 1994b). As each of zinc fingers is specific for a DNA base triplet, tethering multiple ZF motifs in a row allows for the specific targeting of a wide variety of extended DNA sequences.
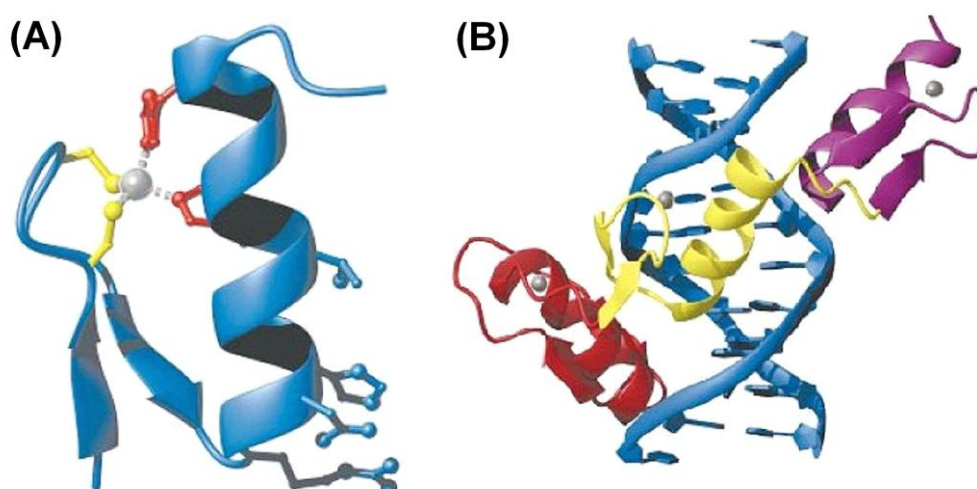
Figure 8. (A) The Cys2His2 zinc finger motif in finger 2 from Zif268. The zinc ion (silver sphere) is coordinated by two cysteine side chains (yellow) and two histidine side chains (red). (B) Zif268-DNAcomplex. Three zinc fingers (red, yellow and purple) are bound in the major grove of the DNA (blue). Coordinated zinc ions are represented as silver spheres. Figure adapted from (Pabo et al, 2001).

First application of the zinc finger protein (ZFP) to target DNA *in vitro* and *in vivo* was demonstrated in 1994. The three-finger protein was engineered to

31

block the expression of an oncogene transformed into a mouse cell line and switch on the gene by using ZF construct (Choo et al, 1994). Later more applications of the ZFP were reported including inhibition of HIV expression (Reynolds et al, 2003), disruption of the effective cycle of HSV infection (Papworth et al, 2003), activating the expression of the vascular endothelial growth factor-A in a human cell line (Liu et al, 2001) and in an animal cell line (Rebar et al, 2002) and regulation of the level of zinc finger expression by a small molecule (Beerli et al, 2000).

ZFP are highly specific DNA binding modules, therefore they are attractive candidates for engineering tools for the gene targeting experiments. The first ZFN were developed by Chandrasegaran and coworkers (Kim et al, 1996) by fusing ZFP to the nonspecific cleavage domain of the type IIS restriction enzyme FokI. Resulting "artificial" nucleases were active and cut DNA at a defined recognition site programmed by the zinc finger module. Type IIS restriction endonuclease FokI is a monomer consisting of two domains: a N-terminal domain for DNA recognition and a C-terminal (catalytic) domain that is a non-specific nuclease (Li et al, 1992). Since the catalytic domain of FokI has only one active site, two C-terminal domains have to associate to make a double-strand break in DNA (Bitinaite et al, 1998). However, FokI exhibits a very weak dimerization interface at the cleavage domain and therefore is unable to dimerize in free solution. In the case of ZFN-FokI fusion protein, a monomeric protein is inactive; however two FokI cleavage domains form a catalytically competent dimer when brought into a close proximity and in correct orientation by the ZF domain binding to the adjacent DNA sites (Figure 2B).

The same type of ZFN was used in the first *in vivo* experiments using *Xenopus laevis* oocytes (Bibikova et al, 2001). In this work two different ZFNs were constructed fusing a DNA binding module, consisting of three ZF, and a nonspecific cleavage domain of FokI. After the injection of constructed ZFNs and engineered DNA substrates into the nuclei of the *Xenopus laevis* oocyte,

DNA cleavage was observed yielding 19% of the subsequent homologous recombination. The dependence of the cleavage efficiency on the distance between the two DNA binding sites was investigated. The original ZF nuclease has an 18 amino acid linker between DNA binding and catalytic domains and cleaves DNA when the target sites are 6 to 18 bp apart, with the optimal distance being 8 bp. However, nuclease without a linker performs efficient cleavage only when binding sites are separated exactly by 6 bp (Bibikova et al, 2001).

Another set of gene targeting experiments using ZFNs were performed in *Drosophila* (Bibikova et al, 2003; Bibikova et al, 2002). A pair of ZFNs was designed that recognize unique site in the *yellow* (*y*) gene of *Drosophila*. The engineered nucleases were expressed in developing larvae resulting in 5.7% of somatic mutations in male specifically in the *y* gene. DNA sequencing revealed that all these mutations were small deletions or insertions located at the designed target and were caused by the cleavage and nonhomologous end joining (NHEJ). These results demonstrated a possibility to perform gene targeting in *Drosophila* when a homologous donor DNA is provided. More specifically, a mutant donor was constructed which carries 8 kb of *y* homology (Bibikova et al, 2003). The experiment was performed by providing a pair of ZFN's and the donor as integrated, circular extrachromosomal and linear extrachromosomal DNA in the developing larvae. The obtained adults were crossed and the frequency of the revealed germline *y* mutants was ~2% and ~0.5% in male and female, respectively, when linear extrachromosomal DNA was used as a donor. Most of these mutations were products of homologous recombination (63 and 73% in male and female, respectively). Later report has shown that higher frequency of mutants (up to 25%) in *Drosophila* could be achieved by increased induction of ZFN expression (Beumer et al, 2006).

Targeted mutagenesis could be applicable to all organisms, including important agricultural species, for improvement of their valuable traits. ZFNs encoding genes were introduced into the plant *Arabidopsis thaliana* and were

driven by a heat-shock promoter (Lloyd et al, 2005). Induction of ZFN expression by heat shock resulted in the mutations at the ZFN recognition sequence at frequencies of about 20%. Similar frequencies of mutations were observed in another experiment with plants, where ZFNs induced mutations in tobacco protoplasts (Wright et al, 2005).

All previous gene targeting experiments indicate that ZFNs is a powerful tool for making directed modifications in the experimental organisms and for creating models of human genetic diseases (Porteus & Carroll, 2005). To study the possibility of the gene modification in human cell line, green fluorescent protein (GFP) model system was used in HEK 293 cells (Porteus & Baltimore, 2003). In this work a zinc finger nuclease QQRL18-CN (Bibikova et al, 2001), that contained 18 amino acid linker between DNA binding and cleavage domains was modified and tested with QQR6 and QQR8 targets that contained 6 and 8 bp spacing between the binding sites, respectively. The gene targeting was stimulated 17-fold and 260-fold, accordingly to the background rate of gene targeting. Another ZFN QQRL0-CN construct lacking a spacer between the domains did not stimulate gene targeting (GT) on QQR8 however stimulated GT on QQR6 target by over 2000-fold. The next engineered nuclease, Zif-CN, contained the DNA binding domain from Zif268 and was shown to bind the Zif6 target site. Further experiments showed that transfection of QQRL0-CN or Zif-CN alone did not stimulate GT on target QQR6/Zif6, however co-transfection of both enzymes stimulates GT ~ 6000-fold (over 4000 GFP-positive cells per million transfected cells) (Porteus & Baltimore, 2003).

Next step in the development of zinc finger nucleases was to perform GT experiments addressing DNA sequences related to the known human genetic diseases. One of the first experiments was focused on the X-linked severe combined immune deficiency (SCID) mutation in the IL2R*g* gene. In order to induce DSB and stimulate HR process, ZFNs were designed to recognize a target site within exon 5 of the IL2RG gene, where a hotspot for the disease

causing mutations is located (Urnov et al, 2005). To recognize 24 bp target site two DNA-binding domains were constructed, each of them containing four zinc finger motifs. Moreover, the enzyme-DNA interface was optimized by single amino acid substitutions in the recognition α helices leading to a 5-fold improvement in the targeting. The experiments, performed in K562 cells, showed that ZFNs stimulated mono-allelic targeting in 11% and bi-allelic targeting in 6% of cells without selection. Reasonably high targeting frequencies were also observed in T cells. These data demonstrate that ZFNs could be useful tools for the gene therapy of genetic diseases.

1.3.3.   TALE nucleases

Transcription activator-like (TAL) effectors family consists of DNA binding proteins from plant pathogenic bacteria *Xanthomonas* and *Ralstonia solanacearum* (Boch et al, 2009; Moscou & Bogdanove, 2009; Scholze & Boch, 2011). These proteins enter plant cells *via* the type III secretion system and reprogram host cells by mimicking eukaryotic transcription factors thus minimizing their immune response (Gu et al, 2005). TAL effectors (TALE) consist of a central DNA-binding domain, a nuclear localization sequence (NLS) and an acidic transcriptional activation domain (AD) (Boch et al, 2009) (Figure 9A). TALE proteins are highly conserved and differ mainly in the composition of the central DNA binding repeat domain. Each domain consists of nearly identical 15.5-19.5 repeats that are about 34 amino acids long (Boch & Bonas, 2010). The last repeat always is only a half repeat. Amino acid sequences of these repeats differ just at the positions 12 and 13 (Figure 9A). These variable di-residues (RVD) participate in the recognition of a single nucleotide at the DNA target site. The number of the DNA binding domain repeats determines the length of the DNA target site recognized, e.g., AvrBs3 from *Xanthomonas campestris* pv. *vesicatoria* the repeat domain contains 17.5 repeats and recognizes 18 bp UPA-box (Boch et al, 2009). Depending on RVD at the positions 12 and 13, repeats are divided into the repeat types that recognize one or several DNA bp. The specificities of several repeat types

were determined by predicting targets of various TALEs or generating artificial effectors with novel specificities (Boch et al, 2009). It was shown that repeat types NI, HD, NG/IG and NK have a strong preference for A, C, T and G, respectively, NN repeat recognizes G and A, NS repeat recognizes A, C, G and T bp (Figure 9B). In addition, usually T is conserved at the 5' end of the TALE binding site and is recognized by the amino acid sequence preceding the first repeat (Boch et al, 2009) (Figure 9A). Despite the lack of the sequence conservation, the secondary structure prediction of this sequence (repeat 0) shows similarities with the repeat region. However, recent study revealed that it is possible to construct truncated variant of TALE AvrXa10 that can bind target sites preceded by A, C or G (Sun et al, 2012).

Until now TALE or similar proteins have been found only in the plant pathogenic bacteria *Xanthomonas* spp. and *Ralstonia solanacearum*. Different strains of *Xanthomonas* usually carry none or only few *TAL* genes however some of them contains up to 28 *TAL* genes per strain (Table 3). Many strains of *R. solanacearum* carry homologous of *Xanthomonas TAL* genes (Heuer et al, 2007; Salanoubat et al, 2002) however they have low similarity at the N-terminal and C-terminal regions.
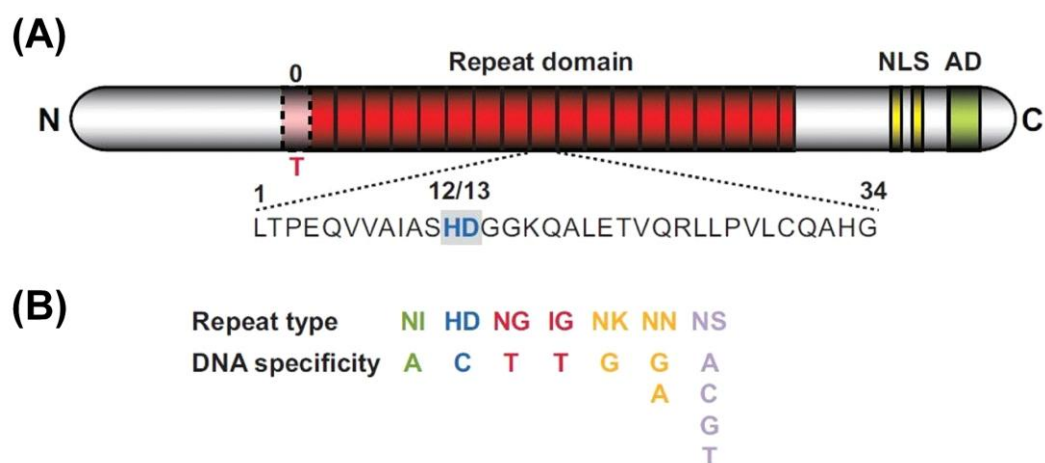
Figure 9. Sequence organization and DNA target specificity of TAL effectors. (A) TAL effectors contain central tandem repeats (red), nuclear localization sequence (NLS) (yellow) and an acidic transcriptional activation domain (AD) (green). Composition of one repeat is shown, hypervariable (RVD) amino acids 12 and 13 are shaded in gray. (B). Repeat types recognize a single or several different DNA bases. Figure adapted from (Scholze & Boch, 2011).

Table3. Number of TAL effectors per strain. Table adapted from (Scholze & Boch, 2011).

	Host	TALs
***Xanthomonas* spp.**		
X. campestris pv. *musacearum*	Banana	0
X. campestris pv. *vasculorum*	Sugarcane	0
X. campestris pv. *vesicatoria*	Pepper and tomato	0-1
X. gardneri	Pepper and tomato	0-1
X. campestris pv. *campestris*	*Brassicaceae*	0-1
X. axonopodis pv. *glycines*	Soybean	1
X. axonopodis pv. *manihotis*	Cassava	≥1
X. campestris pv. *armoraciae*	*Brassicaceae*	0-3
X. axonopodis pv. *citri*	Citrus	1-4
X. fuscans subsp. *aurantifolii*	Citrus	2-4
X. transluens pvs.	Cereals and grasses	0-6
X. campestris pv. *Malvacearum*	Cotton	6-10
X. oryzae pv. *Oryzae*	Rice	7-16
X. oryzae pv. *Oryzicola*	Rice	12-28
***Ralstonia* spp.**		
Ralstonia solanacearum	Broad host range	0-1

Usually all TAL effectors are highly conserved and differ mainly in the composition of the central domain that consists of nearly identical 15.5-19.5 repeats. However, in some members the number of repeats varies between 1.5 and 28.5, therefore the influence of the repeat number to the target gene expression was investigated and experiments have shown that a minimum of 6.5 repeats is necessary for the gene induction and 10.5 or more repeats are sufficient for full induction (Boch et al, 2009). These data demonstrate that TALE proteins with a few repeats are inactive and a certain minimal number of repeats is required to recognize the DNA binding site.

Recently crystal structures of TAL effectors PthXo1 in DNA–bound state and dHax3 in DNA–free and DNA–bound state (Deng et al, 2012; Mak et al, 2012) have been published. Both structures revealed that each repeat comprises two helices connected by a short loop containing variable residues (RVD) and

forms a left-handed bundle. The repeats self-associate to form a right-handed super-helical structure around the DNA major groove (Figure 10). The first RVD residue participates in the stabilization of the RVD loop, while the second makes a base-specific contact.

Figure 10. Structure of the TALE PthXo1 in complex with its DNA target site. Picture adopted from (Mak et al, 2012).

Experiments with redesigned TAL effectors demonstrate that engineered protein can be directed to the chosen DNA target site (Boch et al, 2009; Morbitzer et al, 2010). The programmable DNA binding specificity of TALEs have a potential for application in gene targeting, similar to ZFNs. ZF array specificity is not completely predictable because of the possible overlapping interactions between the neighboring ZFs. In contrast, the DNA binding specificity of TALEs is fully predictable and neighbor-independent (Boch et al, 2009; Moscou & Bogdanove, 2009). One more disadvantage of the ZFNs is their low efficiency and reproducibility (Ramirez et al, 2008). In this respect, TALEs provide a powerful alternative (Mahfouz et al, 2011; Sun et al, 2012). Since TALEs contain a large number of repeat domains, engineering of new proteins recognizing target sites on demand looks difficult. To assembly TALE

38

RVDs in particular number and order different protocols have been derived using mainly the Golden Gate cloning (Cermak et al, 2011; Geissler et al, 2011; Li et al, 2011; Morbitzer et al, 2010; Weber et al, 2011; Zhang et al, 2011), which uses restriction ligation technology to assembly multiple DNA fragments in an ordered fashion in a single reaction (Engler et al, 2009; Engler et al, 2008). To engineer highly specific DNA recognizing modules for various applications TALEs can be coupled to a variety of functional domains such as nucleases, activation and repression domains and theoretically to methyltransferases and integrase domains (Engler et al, 2009).

Similar to ZF, TALE proteins were fused with the C-terminal nuclease domain of the type IIs restriction enzyme FokI (Christian et al, 2010; Li et al, 2011; Mahfouz et al, 2011) (Figure 11A). Resulting TALE nuclease (TALEN) functions as a homodimer or heterodimer and cleaves DNA within the region between two binding sites. Two TALEN monomers bind DNA in a tail-to-tail orientation to allow dimerization of the FokI cleavage domains (Figure 11B). TALENs induced DSB can be repaired by NHEJ resulting in a small deletion or insertion. On the other hand, gene correction or gene addition can occur *via* HR in the presence of the donor DNA.
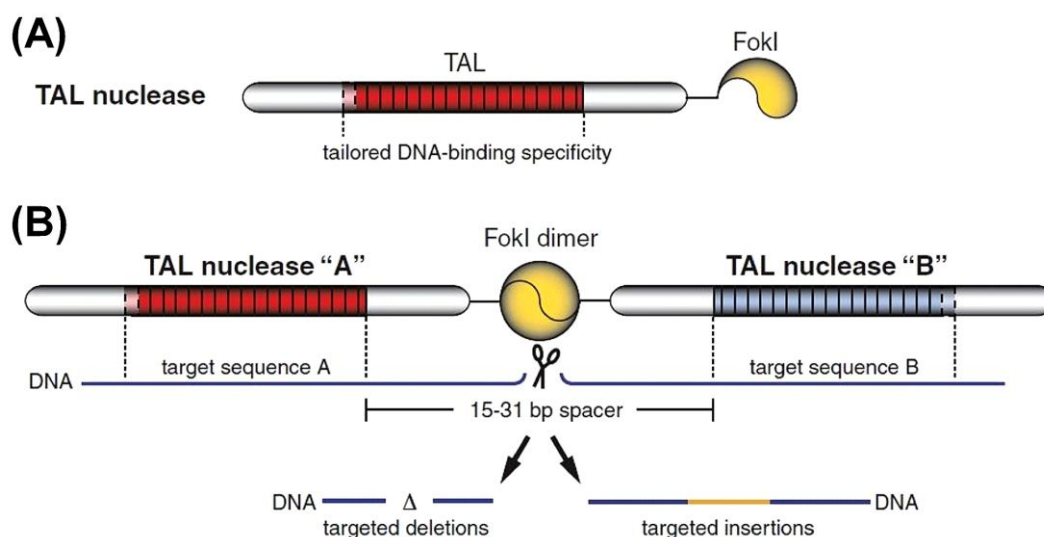

Figure 11. The scheme of the TALEN. (A) TALE nuclease is a fusion between TAL effector and the FokI cleavage domain. (B) Two TALENs binds neighboring DNA target sites tail-to tail, both FokI domains dimerize and cleave DNA to induce homologous recombination. Figure adapted from (Scholze & Boch, 2011).

To perform the gene modification generating DSBs TALENs can be expressed in any cell type. To demonstrate the possibility to use TALENs in plant engineering, Hax3-based hybrid TALE nuclease with a user-defined DNA binding specificity was created (Mahfouz et al, 2011). TALEN dHax3-N contains 11.5 repeats and was designed to recognize 12 bp sequence in the *RD29A* promoter. First, it was demonstrated that a new constructed TALE binds to its target sequence *in vitro* and the homodimeric TALEN can induce DSB *in vitro* when target sites have a proper spacing and orientation. DSB was induced *in vivo* when dHax3-N TALEN was expressed in *Nicotiana benthamiana* leaves (Mahfouz et al, 2011). TALEN induced DSB also was demonstrated in *Arabidopsis* protoplast, where 4-15 bp deletions were observed after the NHEJ repair process (Cermak et al, 2011).

To demonstrate the possibility TALENs application in mammalian system, 13 repeat TALE (TALE13) from *Xanthomonas axonopodis* was chosen as a platform for generation of a new protein with a desired specificity (Miller et al, 2011). To create a TALE that could regulate an endogeneous mammalian *NTF3* gene, a repeat domain of TALE13 was replaced with a new 18 repeat array, specific for 18 bp site in *NTF3* promoter. The generated NT-L TALE and a truncation variant NT-L+95 (retained 95 residues of a C-terminal flanking region) were conjugated with the same VP16 activation domain, expressed in human HEK293 cells, and tested for activation of the endogenous *NTF3* locus. More than a 20-fold induction of the *NTF3* transcript and protein product was observed. Then heterodimer partner for the NT-L was designed (NT-R) and both proteins were fused with the cleavage domain of FokI. Following expression of new TALENs in K562 cells the expected gene modification was observed. Subsequent repair by NHEJ also was observed when new TALEs were designed to modify human *CCR5* gene. Moreover, in the presence of the donor DNA fragment gene editing occurred through the homology directed repair. PCR analysis revealed up to 16% alleles possessing the inserted sequence (Miller et al, 2011).

Full-length TALE protein contains additional fragments besides the DNA binding repeat domain (Figure 9). Several groups reported that minimal central TALE repeat domain lacks activity (Christian et al, 2010; Miller et al, 2011; Sun et al, 2012). Miller *et al.* (Miller et al, 2011) demonstrated that truncation of the first 152 N-terminal residues from TALE protein does not influence its specificity. TALEN variant that retains 28 or 63 of the original 278 C-terminal amino acids (a.a.) exhibits the highest DNA cleavage efficiency (Miller et al, 2011). Similar results were obtained by Sun *et al.* (Sun et al, 2012). They tested TALEN AvrXa10 variants bearing different N- and C-terminal extensions against the target sites with various spacer lengths. This test demonstrated that TALEN with a short C-terminal extension (31 a.a.) requires at least 125 a.a. N-terminal fragment for efficient DNA cleavage. Longer C-terminal extension enables DNA cleavage when the target sites are separated with longer spacers (Sun et al, 2012).

In summary, TALENs recognize long DNA sequences and cut the target site specifically. They exhibit robust nuclease activity and low cytotoxicity and therefore are perspective tools in gene therapy.

1.3.4.   TFO-linked nucleases

Triplex forming oligonucleotide (TFO) binds to the polypurine sequences of the ds DNA with high affinity forming triple-helical DNA. Triple-helical nucleic acids were first described in 1957, when mixing of the polyuridylic acid (polyU) and polyadenylic acid (polyA) in the ratio of 2:1 resulted in the formation of a stable complex (Felsenfeld & Rich, 1957). In 1986 the formation of the first stable specific DNA triple helix was demonstrated (Dervan, 1986). Due to its ability to bind ds DNA in a sequence-specific fashion TFO could be used as an addressable DNA binding module in a variety of biological applications. DNA triplex is a complex where the third strand (TFO) binds in the major groove of the target DNA duplex containing polypurine stretch (Fox, 2000) (Figure 12). In contrast to classical Watson-

Crick base pairing in ds DNA, the specificity and stability of the triple helix DNA structure is maintained *via* Hoogsteen hydrogen bonds.
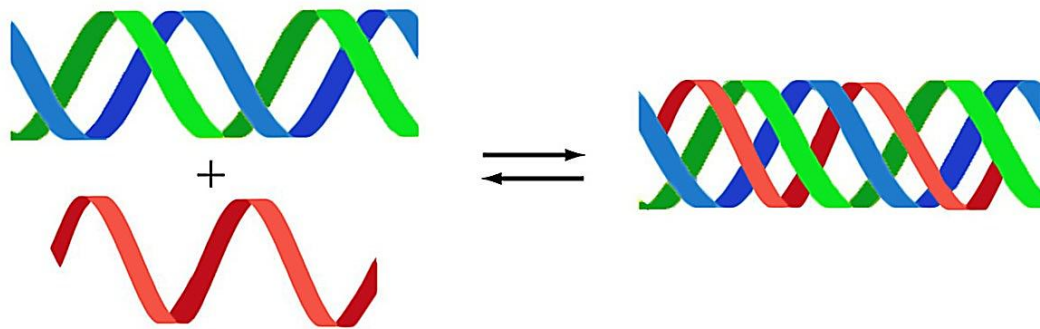

Figure 12. Schematic representation of the DNA triple helix formation. Duplex DNA shown in blue and green, TFO, that binds in the major groove is shown in red. Figure adapted from (Jain et al, 2008).

There are several types of triplexes depending on the orientation of the third strand. Pyrimidine motif TFO consisting of C and T nucleotides, form $C^+*G{:}C$ and T*A:T triplexes and bind polypurine sequences in parallel orientation *via* Hoogsteen hydrogen bonds (Figure 13A) (Radhakrishnan & Patel, 1994). This type of DNA triple helix requires lower pH (<6.0) to protonate cytosines in the third strand (Lee et al, 1984). In contrast, the purine motif TFO consist of A and G nucleotides, form G*G:C and A*A:T triplexes and bind polypurine sequences in antiparallel orientation *via* reverse-Hoogsteen hydrogen bonds at the neutral pH (Figure 13B). Mixed TFO's (G/T) also exist that can bind polypurine sequences in parallel orientation forming Hoogsteen hydrogen bonds or in antiparallel orientation forming reverse-Hoogsteen hydrogen bonds (Duca et al, 2008). To obtain a stable DNA triple helix the sequence of TFO should be fully complementary to the polypurine target site on the ds DNA (Gowers et al, 1999). However, this interaction thermodynamically is weaker than that of the two strands of the duplex (Shafer, 1998). Moreover, under physiological conditions there is a negative charge repulsion between the phosphates of the three DNA strands. Therefore, various triplex stabilizers reducing a negative charge, such as Mg^{2+} (Blume et al, 1999; Floris et al, 1999) or polyamines (Thomas & Thomas, 1993; Thomas et al, 1996) are often used.
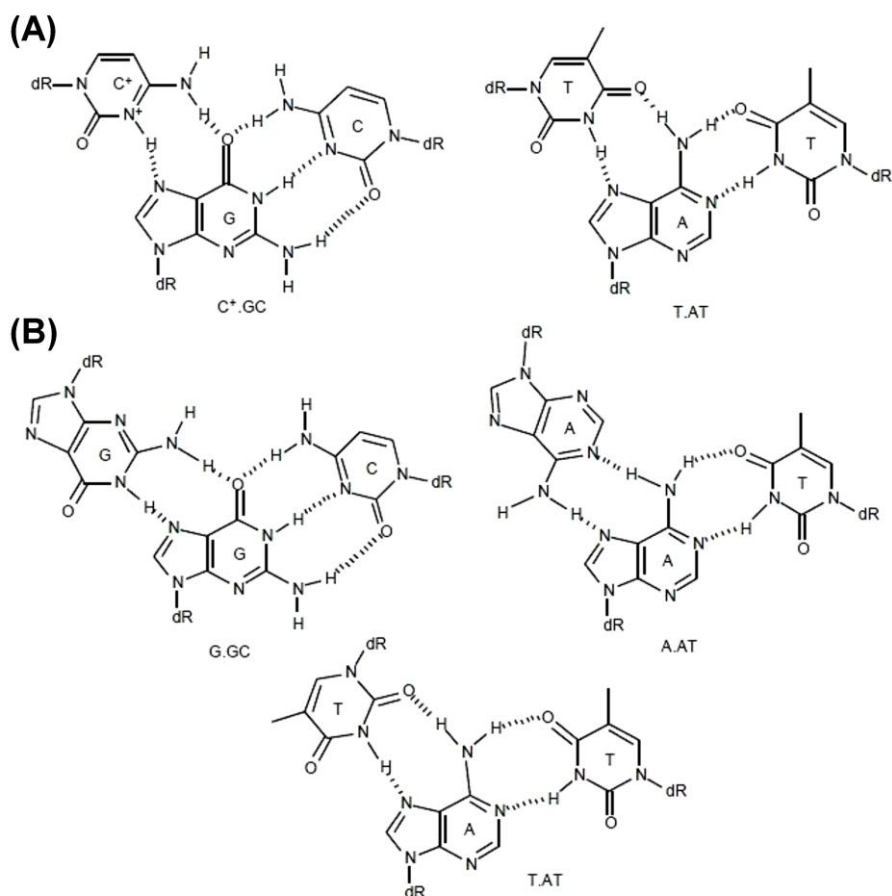
Figure 13. (A) Structures of the parallel triplets C+*G:C and T*A:T. (B) Structures of the antiparallel triplets G*G:C, A*A:T and T*A:T. Figure adapted from (Fox, 2000).

To overcome the limitations, that disturb the stability of the DNA triple helix, a number of modifications are used in the bases, backbone or sugar of the TFO (Figure 14).
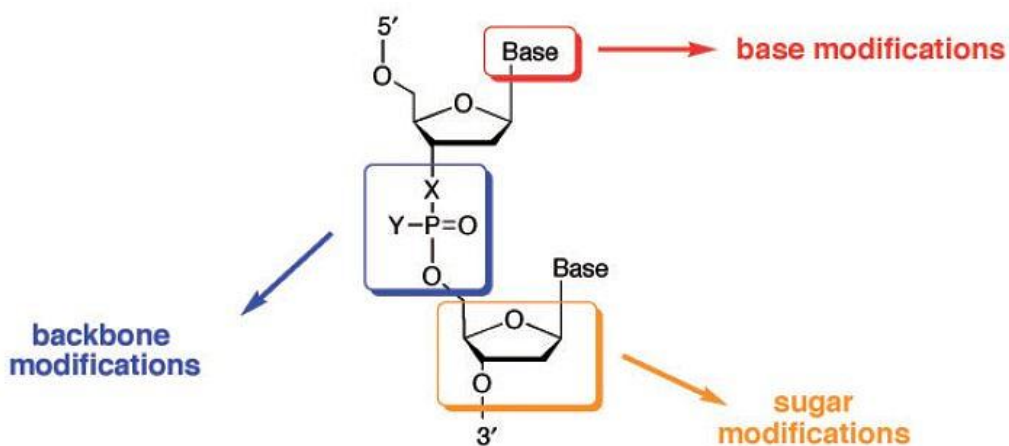
Figure 14. Chemical modifications introduced in TFOs. Figure adapted from (Duca et al, 2008).

Nucleotide analogs in TFO. To form a stable triple helix using pyrimidine TFO, a protonation of cytosine is required. To overcome this problem, nucleotide analogs have been used that reduce the pH dependence of the triplex formation. One of the first analogs used was 5-methylcytosine, which has slightly higher pKa than cytosine (Lee et al, 1984; Xodo et al, 1991). Additional methyl group in the pyrimidine ring increases the triplex stability by improving the base stacking in the major groove. Triplexes with this analog are more stable than with natural cytosine, however, not enough under the physiological conditions. Another cytosine substitution is pseudoisocytosine that is 2'-O-methyl derivative (Fox, 2000). This base contains a hydrogen at position N3 and forms Hoogsteen hydrogen bonds with a guanine base similar to a protonated C. Other pyrimidine analogs are 6-oxocytosine, 2-aminopyridine or N6-methyl-8-oxo-2-deoxyadenosine (Fox, 2000; Krawczyk et al, 1992; Xodo et al, 1991). Substitution of a thymine by deoxyuridine or 5-propynyldeoxyuridine promotes triplex formation at neutral pH (Mills et al, 2002). In the antiparallel purine triplex motif guanine could be substituted by a 6-thioguanine (Olivas & Maher, 1995; Rao et al, 1995), or adenine by 7-deazaxanthine (Milligan et al, 1993). Finally, synthetic nucleotides have been created capable to recognize all four bases by the triple helix formation at physiological pH (Rusling et al, 2005).

Modifications in the sugar backbone of TFO which helps to stabilize triple helix have been developed. It was observed that a number of ribose analogues could stabilize a triple helix. It was found that 2'-methoxylation (2'-OMe) (Escude et al, 1992; Shimizu et al, 1992) stabilizes the C3'-endo conformation of the sugar resulting in the smallest distortion of the duplex DNA (Asensio et al, 1999). Furthermore, TFO containing 2'-OMe modification is more resistant to RNA nucleases. Another modification, 2'-aminoethylation, adds a positive charge to the C3'-endo conformation at physiological pH (Seidman et al, 2005). Conformational restriction could help to stabilize the complex because of an entropic advantage (Duca et al, 2008). It

was shown that, blocking the sugar using locked nucleic acid (LNA) monomers, (Petersen & Wengel, 2003) significantly enhances the triplex stability (Brunet et al, 2005; Torigoe et al, 2001). However, just one or several nucleotides in TFO should be locked in the C3'-endo conformation, since TFO composed entirely of LNA monomers (O2',O4'-methylene-linked nucleic acid) does not form DNA triple helixes (Sun et al, 2004). One more locked conformation is O2',O4'-ethylene-linked nucleic acid (ENA). Fully modified ENA-TFOs form stable triple helices at pH 7.2. Also a new LNA-type TFO (called 2'-4'-BNA) was described that exhibits a higher affinity to duplex DNA than LNA and ENA (Rahman et al, 2008).

Another method to increase the stability of the triple helix is the TFO backbone modification by introducing a phosphorothioate or N3'-P5' phosphoramidate. These modifications change the electrostatic properties of the negatively charged phosphodiester backbone and help to stabilize triple helix (Michel et al, 2005; Michel et al, 2003).

Stabilizing compounds. One more possibility to stabilize the DNA triple helix is to use compounds which bind selectively to the DNA triplexes. The first such known ligand was DNA duplex intercalator ethidium ion (Scaria & Shafer, 1991). Later, polycyclic ligands like the benzo[e]pyridoindole (BePI) (Mergny et al, 1992) were developed which selectively bind to the DNA triplex, for example, by intercalation (Pilch et al, 1993). BePI selectively stabilizes T*A:T triplets, however C+*G:C triplet contains a positive charge and prevents binding of the cationic ligand. Triplex stabilization also was obtained with the BePI derivatives benzo[g]pyridoindole (BgPI), benzo[f]pyridoquinoxaline (Escude et al, 1998), benzo[f]quinoquinoxaline (Marchand et al, 1996) and dibenzophenanthriline.

One more way to stabilize the DNA triple helix is to use intercalating or cross-linking agents which could be attached to the 5'-end or 3'-end of the triplex forming oligonucleotide. Several groups showed that using TFO with the acridine compound on the 5'-end or 3'-end significantly increases the

triplex stability (Birg et al, 1990; Grigoriev et al, 1992; Stonehouse & Fox, 1994; Sun et al, 1989; Sun et al, 1991). Acridine intercalates at the triplex-duplex junction and a higher stability is obtained when the attachment is at the 5'-end of the TFO (Sun et al, 1989). Attachment of an acridine to the TFO increases the triplex stability at least 100-fold (Fox, 2000). DNA triple helix can also be stabilized by a chemical cross-linking agent such as azidoproflavin (Le Doan et al, 1987) or azidophenacyl (Praseuth et al, 1988) coupled to one or other end of the TFO. Long wavelength irradiation of attached psoralen covalently conjugated TFO to its duplex target site (Giovannangeli et al, 1992; Takasugi et al, 1991).

Specific DNA duplex recognition by the triplex forming oligonucleotides offers a new approach for the gene modification. Triplex formation can help in modulating gene expression, DNA damage, mutagenesis or enhance the homologous recombination (Maurisse et al, 2002). Various compounds that could damage DNA could be attached to TFO (Duca et al, 2008): photoactivatable agents (Perrouault et al, 1990), metal complexes, such as Fe–EDTA (Strobel & Dervan, 1990), orthophenantroline (Francois et al, 1989a; Francois et al, 1989b), metalloporphirines (Bigey et al, 1995) and enzymes such as nucleases (Landgraf et al, 1994; Pei et al, 1990). The most widely used agents, conjugated to TFO, are psoralens (Faruqi et al, 1998; Rogers et al, 2004; Rogers et al, 2002; Vasquez et al, 1996). Conjugation of photoactivable psoralen to TFO allows controlling the timing of introduced damage to reduce the non-specific alteration of genome (Jain et al, 2008). More attractive than psoralens are nucleases, that could be conjugated with TFO. Targeting of individual genes in complex genomes requires high specificity, therefore nuclease-TFO conjugate could be a very useful tool to address nuclease to a particular site in the genome. The first known nuclease conjugated with TFO is staphylococcal nuclease (Pei et al, 1990). This enzyme is relatively non-specific and hydrolyzes both single-stranded and double-stranded DNAs. Staphylococcal nuclease-TFO conjugate was able to bind the polypurine

sequence *via* triple helix formation and hydrolyze both DNA strands in the target duplex when the Ca^{2+} ions were added. However, the staphylococcal nuclease is a non-specific nuclease and much higher specificity could be obtained by conjugation TFO with REase that cleaves DNA at a defined phosphodiester bond.

Type II restriction endonucleases are sequence specific nucleases that recognize 4-8 bp DNA and cleave it at defined sites (Pingoud et al, 2005). To generate TFO-REase conjugate single-chain variant of the naturally homodimeric REase PvuII (scPvuII) was used (Simoncsits et al, 2001). TFO used to conjugate with scPvuII was composed of nucleotide analogs. To ensure a triple helix formation at physiological pH, a cytosine was replaced with 5-methylcytosine (Xodo et al, 1991). Replacing a thymine with 5-propynyldeoxyuridin also helps to stabilize DNA triple helix (Phipps et al, 1998). After the coupling of the scPvuII with TFO, the resulting construct recognizes two DNA sites: one is the REase PvuII target site 5'-CAGCTG, another is the polypurine track that is complementary to the TFO sequence (Eisenschmidt et al, 2005) (Figure 2D). The DNA cleavage activity of scPvuII-TFO was studied using a linearized 5.5 kbp plasmid with five PvuII recognition sites, one of which (the "addressed" site) is separated by 9 bp from the triplex forming sequence. First, the conjugate and DNA were mixed and incubated overnight to promote the formation of the DNA triple helix. To this mixture 1 mM of spermine was added, that helps to alleviate charge repulsion between the double stranded DNA and the third strand (Singleton & Dervan, 1993). Cleavage reaction was initiated by adding 2.5 mM of $MgCl_2$ and the cleavage was observed only at the addressed PvuII site. This work demonstrates the ability to address restriction endonuclease to a specific locus and introduce cleavage at the defined target site.

1.4. Off-target cleavage problem

1.4.1. Improving zinc finger nucleases

Different engineered proteins are promising tools for genome manipulations including gene therapy; however off-target cleavage problem still limits their application (Cornu et al, 2008). It was shown that a number of ZFNs induce significant cytotoxicity in human cells by cleavage at off-target sites (Alwin et al, 2005; Miller et al, 2007; Porteus, 2006; Porteus & Baltimore, 2003; Szczepek et al, 2007).

Dimerization. ZFNs are most often created by fusing zinc-finger motifs, which serve as DNA recognition modules, to the nonspecific DNA cleavage domain of the FokI restriction endonuclease (Beumer et al, 2006; Bibikova et al, 2002; Kim et al, 1996; Miller et al, 2007). The activity control of ZFNs is achieved due to weak dimerization of the FokI cleavage domain. Being unable to dimerize in free solution, two FokI cleavage domains form a catalytically competent dimer only when they are brought into close proximity and in correct orientation by the zinc-finger domain binding to the adjacent DNA sites (Figure 2B). However, at very high ZFN concentrations (Beumer et al, 2006; Bibikova et al, 2002; Kim et al, 1996; Miller et al, 2007)dimerization may occur in the absence of correct DNA binding, resulting in the nuclease activation and cleavage at off-target sites (Szczepek et al, 2007). One of the possibilities to reduce such toxicity is the modification of enzyme's specificity through redesign of the dimer interface (Bolon et al, 2005; Sims et al, 2006). Computational analysis of FokI crystal structures (Wah et al, 1998; Wah et al, 1997) indentified residues that affect dimer formation. Based on the structure, modifications at the dimer interface were engineered that reduced homodimerization but allowed heterodimerization retaining the endonuclease activity (Miller et al, 2007; Szczepek et al, 2007). These specific mutations within the dimer interface of ZFNs significantly reduced their toxicity keeping an ability to induce robust homologous recombination.

DNA binding. In some cases the reason of ZFNs toxicity could be due to an unexpected DNA binding by the engineered zinc fingers. The best studied zinc finger protein is three-finger transcription factor Zif268 containing Cys2His2 zinc finger motif (Elrod-Erickson & Pabo, 1999; Pavletich & Pabo, 1991). Each finger recognizes 3-4 bp and binds DNA independently of the neighboring fingers, therefore new ZFNs were engineered by modular assembly (Bibikova et al, 2003; Bibikova et al, 2001; Porteus, 2006; Porteus & Baltimore, 2003). However, the specificity of the assembled zinc fingers is not always predictable (Alwin et al, 2005). Therefore ZFNs made using modular-assembly were compared with ZFNs made using a bacterial two-hybrid (B2H) selection-based method by targeting the same target site (Pruett-Miller et al, 2008). It was found that ZFNs made by B2H strategy were more efficient and less toxic than made by a modular-assembly.

Another work investigated the relationship between DNA binding affinity and specificity of ZF domains and their toxicity as ZFNs in human cells (Cornu et al, 2008). To address this problem, a set of ZFNs specific for the same target site was engineered, and it was shown that they bind the same target site with different affinities and specificities. These results demonstrated that DNA binding affinity and the zinc finger specificity influence the activity and toxicity of ZFN.

Half-life of a ZF protein. One possibility to reduce cytotoxic effects of many ZFNs is shortening their half-lives and regulating protein levels in a cell with small molecules. It was demonstrated that ZFNs could be destabilized by coupling an ubiquitin moiety to the N-terminus of the protein and regulation of their level could be achieved by using a proteosome inhibitor (Pruett-Miller et al, 2009). Also, ZFNs can be linked to a modified destabilizing FKBP 12 domain at the N-terminus and regulation of ZFN level could be performed by a small molecule that blocks the destabilization effect of the N-terminal domain. This work demonstrated that regulation of ZFNs level ensures a high activity while reducing toxicity (Pruett-Miller et al, 2009).

49

Caging. Recently, a caged ZFN that is regulated by light to avoid toxicity, has been generated (Chou & Deiters, 2011). The Tyr471 residue, that is located close to the FokI active center, was specifically substituted with a photocaged amino acid ortho-nitrobenzyltyrosine (ONBY). The introduced bulky group prevented the protein-DNA interaction thus inactivating the ZFN. However, irradiation with UV light removed the caging group from Tyr generating an active enzyme. This work demonstrated that light-activated, sequence-specific gene editing can be achieved with the site-specifically caged ZFNs in mammalian cells. ZFN control by UV irradiation not only helps to solve the toxicity problem but also allows to control ZFN mediated gene editing (Chou & Deiters, 2011).

1.4.2. Light activation of the DNA binding protein

The ability to regulate biological processes by an external signal (e.g. light) seems to be very attractive and useful for many *in vivo* or *in vitro* applications. This could be achieved by caging of biomolecules with photoremovable compounds. The principle of this strategy is simple: the caged biomolecule is inactive, while irradiation with light removes the cage resulting in reactivation of the biomolecule. When the biological process one wishes to regulate is an enzymatic reaction each participant in this reaction (substrate, inhibitor, cofactor or enzyme) could be targeted for caging (Goeldner & Givens, 2005). **Caging strategies.** Several caging strategies are available when the target is an enzyme/protein (Goeldner & Givens, 2005). Various functional groups of a protein (carboxylates, amines, thiols) can be caged by chemical modification or introduction of unnatural (caged) amino acids by *in vitro* or *in vivo* translation. Furthermore, different functionally important regions (active site, substrate binding site, oligomerization interface) of the enzymes can be selected for the modification to abolish their activity. Chemical caging of the active site cysteines of the SssI DNA (cytosine-C5)-methyltransferase with 4,5-dimethoxy-2-nitrobenzyl bromide (DMNBB) decreases the activity by up to 95%, while irradiation of the caged methyltransferase restored 60% of its

50

activity (Rathert et al, 2007). The *Taq* polymerase was inactivated by sterically blocking the space reserved for incoming dNTPs (Chou et al, 2009). In this case a tyrosine residue containing the 2-nitrobenzyl group was introduced at the position 671 using an *in vivo* translation system. Irradiation with UV light regenerated the wt *Taq* polymerase and thus restored DNA polymerization activity. When an enzyme is active in a particular oligomeric state there is a possibility to disrupt protein-protein interaction or damage communication between the monomers by introducing a bulky photoremovable group (see below as an example the restriction endonuclease BamHI).

DNA cutting enzymes as target for caging. Photoremovable caging groups can be used to regulate the activity of the DNA cutting enzymes, which could have potential applications in the gene targeting and gene therapy (Pingoud et al, 2011). HEases (see 1.3.1.) were successfully used as scaffolds for generating molecular scissors for a variety of genomic applications (Arnould et al, 2007; Grizot et al, 2009; Redondo et al, 2008; Stoddard, 2005). The Type II REases can be adapted for the gene targeting purposes by fusing them to a DNA binding module recognizing long DNA sequences. The non-specific DNA cleavage domain of the Type IIS restriction endonuclease FokI was extensively used for a fusion with various DNA binding modules such as ZF motifs (Chandrasegaran & Smith, 1999; Kim et al, 1996; Miller et al, 2007; Szczepek et al, 2007; Urnov et al, 2010), TALE repeats (Blume et al, ; Christian et al, 2010; Schierling et al, 2010), a non-cleaving mutant of the homing endonuclease I-SceI (Lippow et al, 2009) and other DNA binding domains (Horner & DiMaio, 2007; Kim & Chandrasegaran, 1994; Kim et al, 1997; Kim et al, 1998; Lariguet et al, 1999; Ruminy et al, 2001) (see 1.3.2.-1.3.3.). TFO may also serve as DNA binding modules as demonstrated for the PvuII restriction endonuclease (Eisenschmidt et al, 2005) (see 1.3.4.). However, despite the advances made in the construction of the specific DNA scissors, cleavage of off-targets remains a major limitation for their application in gene therapy (*in vivo*) (Beumer et al, 2006; Miller et al, 2007; Szczepek et

al, 2007), which could be minimized by controlling the activity of the nucleases. For example, formation of the DNA triple helix is quite slow; therefore, the active enzyme in the REase-TFO fusion could start to cleave at non-specific DNA sites that are not addressed by the triple helix formation. Caging with the photoremovable compound could help to regulate the activity of these designed enzymes and avoid non-specific activity.

Caged REases. To date most of the work on caging of DNA cleaving enzymes with photosensitive compounds was done with the Type IIP restriction endonuclease BamHI. In the case of BamHI, which is active as a dimer, the dimerization interface was chosen as a target for the caging to control the enzyme activity (Chevalier et al, 2004; Endo et al, 2004; Nakayama et al, 2005). Unnatural amino acids, a lysine derivative with a photoremovable group (6-nitroveratryloxycarbonyl) and alanine derivatives such as azophenylalanine or 4'-carboxyphenylazophenylalanine, which undergo *cis-trans* isomerization upon illumination were incorporated at the dimer interface of BamHI substituting amino acid Lys132 using an *in vitro* translation system with special codons and the corresponding tRNA (Chevalier et al, 2004; Endo et al, 2004; Nakayama et al, 2005). These light sensitive compounds destroy the intricate salt-bridge network at the dimer interface thereby inactivating the enzyme. Irradiation of the caged enzymes with UV light restores their catalytic activity.

The work of the Eisenschmidt et al. demonstrated the possibility to perform the addressing of the restriction enzyme to the particular target site on DNA (Eisenschmidt et al, 2005). The single chain PvuII (scPvuII) enzyme conjugated with TFO was able to bind and cleave only at the PvuII target site "addressed" with the adjacent triplex forming site (TFS) and no cleavage was observed at the second, "non-addressed" site. The major limitation of scPvuII-TFO application in the *in vivo* experiments is the requirement to perform preincubation of the conjugate with the substrate DNA in the absence of Mg^{2+} ions to ensure the formation of the DNA triple helix and prevent cleavage of

the "non-addressed" target sites. However, in cells it is not possible to remove divalent cations therefore photosensitive derivative was used to regulate activity of scPvuII (Schierling et al, 2010). Azobenzene derivative can reversibly izomerize between the extended *trans-* and the more compact *cis-*configuration depending on the presence of the UV or blue-light (Renner & Moroder, 2006) (Figure 15).
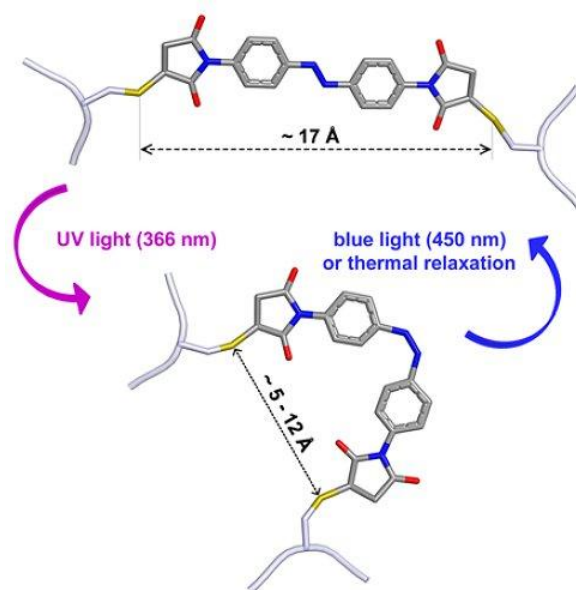

Figure 15. A scheme of the trans → cis isomerization of 4,4′-bis (maleimido)azobenzene attached to cysteine residues induced by illumination with UV light and cis → trans isomerization with blue light. Figure adapted from (Schierling et al, 2010).

In order to generate a photoswitchable enzyme more than 30 scPvuII variants were produced containing two or four cysteine residues required to attach a bifunctional azobenzene cross-linker (Schierling et al, 2010). The DNA cleavage activity of the generated cross-linked variants showed that introducing a single azobenzene results in only a small photoswitchable effects, however large effects were observed with the multiple cross-links. The best photoswitchable effect (up to approximately 16-fold) was obtained with the scPvuII variant that had two azobenzene cross-links near the active site in each domain and additional amino acid substitutions.

The successful example of REases activity regulation by light shows the possibility to apply REases-TFO conjugates in the *in vivo* experiments.

2. MATERIALS AND METHODS

## 2.1.	Materials

All chemicals used in this study were of the highest quality available.

### 2.1.1.	Enzymes

T4 polynucleotide kinase, DNA polymerases TaqI and PfuI, calf intestine alkaline phosphatase, T4 DNA ligase, bovine serum albumin and all REases used in this study were obtained from Fermentas UAB (Vilnius, Lithuania). All these products were used according to the manufacturer's instructions.

### 2.1.2.	*E. coli* strains

E.coli ER2267**:**

Δ(argF-lac)U169recA1/F$^{'}$proAB lacI$^q\Delta$(lacZ)M15zzf::mini-Tn 10(Kan^r).

E. coli ER2566:

F–λ–fhuA2[lon]ompT lacZ::T7 gene1 gal sulA11.(mcrC-mrr)114::IS10 R(mcr-73::miniTn10-TetS)2 R(zgb-210::Tn10)(TetS) endA1 [dcm].

### 2.1.3.	DNA

λ dam$^-$dcm$^-$ DNA was obtained from Fermentas UAB (Vilnius, Lithuania), pUCGK-4 plasmid provided by G. Kruckas (Fermentas UAB, Vilnius, Lithuania), pUCAT1, pUHE25-2-MunIR (Ap^r) wt R.MunI expression plasmid and pMunIM 6.2 (Tc^r, Cm^r) M. MunI expression plasmid provided by A. Lagunavičius ("Thermo Fisher Scientific"), pUHE25-2-MunI-H6G4C (Ap^r) His-tagged R. MunI expression plasmid, pRIZ′-scPvuII-H6G4C scPvuII expression plasmid and pECFP-PvuII DNA substrate plasmid were provided by K. Eisenschmidt (Justus Liebig University, Giessen, Germany), pBse634IR3.7 (Ap^r) Bse634I expression plasmid provided by M. Zaremba (Vilnius University, Institute of Biotechnology, Vilnius, Lithuania),

pBse634IR-H6G4C (Ap^r) His-tagged Bse634I expression plasmid, pHpaIIM (Cm^r) M.HpaII expression plasmid.

2.1.4.   Oligonucleotides

All non-modified oligonucleotides were purchased from Metabion (Martinsried, Germany). Oligonucleotides, containing modified nucleotides were purchased from Eurogentec (Liege, Belgium).

All DNA triple helix forming oligonucleotides and double-strand oligonucleotides used in this study and their short descriptions are given in Table 4.   Oligoduplexes   were   made   by   annealing   two oligodeoxyribonucleotides with complementary sequences.

Table 4. Oligomucleotides and oligoduplexes used in this study*

Oligonu-cleotide(s)	Sequence	Specification
TFO	5'-NH$_2$-(CH$_2$)$_{12}$-CTCTCTCTCTTTTTT-3'	DNA triple helix forming oligonucleotide for coupling with caged MunI.
TFO1	5'-NH$_2$-(CH$_2$)$_{12}$-MPMPMPMPMPPPPP-3'	DNA triple helix forming oligonucleotide for coupling with His-tagged Bse634I.
TFO2	5'-NH$_2$-(CH$_2$)$_{12}$-PMPPMPPMPPPMPPP-3'	DNA triple helix forming oligonucleotide for coupling with His-tagged Bse634I.
24/24 oligoduplex	5'-GTGATACT**CAATTG**GAATCCGTCA-3' 3'-CACTATGA**GTTAAC**CTTAGGCAGT-5'	24/24 cognate oligoduplex for MunI subunits exchange analysis.
23/23 oligoduplex	5'-CGATGTGCTGAAGTTTAGACCTG-3' 3'-GCTACACGACTTCAAATCTGGAC-5'	23/23 non-cognate oligoduplex for MunI subunits exchange analysis.

Oligonu-cleotide(s)	Sequence	Specification
22/19 oligoduplex	5′–ACCCGCGG**CAATTG**GGCCCTTT–3' 3'–TGGGCGCC**GTTAAC**CCGGG–5'	22/19 bp cognate oligoduplex for wt cross-linked and caged MunI gel shift analysis.
22/22 oligoduplex	5'–AATAGGTCCTATAGGCGAATGG–3' 3'–TTATCCAGGATATCCGCTTACC–5'	22/22 bp non-cognate oligoduplex for wt and caged MunI gel shift analysis.

* MunI recognition site is in boldface and underlined. M is cytidine analog 5-methyl-2′-deoxycytidine and P is thymine analog 5-[1-propynyl]-2′-deoxyuridine.

Oligonucleotides used in DNA binding studies were labeled on 5'-end. Radioactive labels were introduced into the 5'-ends of individual DNA strands prior to the annealing with unlabelled strands. 5'-labelling of the top strand was performed with [γ^{33}-P] ATP (Hartmann Analytic, Braunschweig, Germany) and DNA labeling kit (Fermentas UAB, Vilnius, Lithuania). Labeling reactions were performed according to the manufacturer's instructions.

2.1.5. Buffers

Electrophoresis buffer: 30 mM MES-His (pH 6.5 at 25 °C), 0.01 mM EDTA.

Binding buffer: 30 mM MES-His (pH 6.5 at 25 °C), 0.02 mM EDTA, 0.2 mg/ml BSA 10 % (v/v) glycerol.

Equilibration buffer I: 10 mM K-phosphate (pH 7.4 at 25 °C), 100 mM NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA.

Equilibration buffer II: 20 mM Na-phosphate (pH 7.4 at 25 °C), 0.5 M NaCl, 1 mM 2-mercaptoethanol.

Equilibration buffer III: 10 mM Na-phosphate (pH 7.4 at 25 °C), 0.1 M NaCl.

Equilibration buffer IV: 10mM Na-phosphate (pH 6.0 at 25 °C), 0.1 M NaCl.

Storage buffer I: 10 mM Tris-HCl (pH 7.4 at 25 °C), 100 mM KCl, 1 mM EDTA, 1 mM DTT and 50 % v/v glycerol.

Storage buffer II: 10 mM K-phosphate (pH 7.4 at 25 °C), 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 50 % (v/v) glycerol.

Subunit exchange buffer: 10 mM Bis-Tris-propane (pH 6.5 or pH 8.3 at 25 °C), 100 mM NaCl, 10 mM DTT.

Gel filtration buffer: 20 mM Tris-HCl (pH 7.5 at 25 °C), 0.2 M KCl.

Titration buffer: 10 mM Tris-HCl (pH 7.5 at 25 ºC), 200 mM KCl.

Denaturation buffer: 10 mM Tris-HCl (pH 7.5 at 25 °C), 200 mM KCl and 1 mM DTT.

Decaging buffer: 100 mM Na-acetate (pH 5.8 at 25 °C), 0.2 M KCl and 20 mM DTT.

Cross-linking buffer: 100 mM K-phosphate (pH 7.0 at 25 °C), 100 mM KCl.

Reaction buffer I: 10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM $MgCl_2$, 50 mM NaCl, 0.1 mg/ml BSA.

Reaction buffer II: 10 mM Tris-HCl (pH 8.5 at 25 °C), 100 mM KCl, 10 mM $MgCl_2$, and 0.1 mg/ml BSA.

Reaction buffer III: 80 mM Tris-phosphate (pH 7.1 at 25 °C), 2 mM spermine, 2 mM $MgCl_2$ and 10 mM DTT.

Reaction buffer IV: 80mM Tris-phosphate (pH 7.1 at 25 °C), 4 mM Mg^{2+}, 1mM DTT.

Triplex buffer I: 80 mM Tris-phosphate (pH 6.1 at 25 °C), 1.0 M NaCl and 4 mM Mg^{2+}, 1mM DTT.

Triplex buffer II: 80 mM Tris-phosphate (pH 6.1 at 37 °C), 2 mM spermine, 2 mM $MgCl_2$ and 10 mM DTT.

Loading dye solution I: 75 mM EDTA, pH 9.0, 0.01 % bromphenol blue and 50 % (v/v) glycerol.

Loading dye solution II: 50 mM EDTA (pH 8.0 at 25 °C), 0.1% SDS, 50% v/v glycerol, 0.01% bromphenol blue.

2.2. Methods

2.2.1. Electrophoresis

2.2.1.1. Non-denaturing electrophoresis through agarose

Separation of supercoiled, open circular and linear forms of plasmid DNA was performed in 1 % agarose gels in the electrophoretic buffer containing 100 mM H_3BO_3-NaOH, 15 mM sodium acetate, 2 mM EDTA (pH 8.2 at room temperature) and 0.5 µg/ml ethidium bromide. DNA samples were mixed with 1/3 volume of Loading dye solution I (75 mM EDTA, pH 9.0, 0.01 % bromphenol blue and 50 % (v/v) glycerol) and electrophoresed at 3 V/cm until the bromphenol blue dye migrated for 3 cm. Digital images of the gels were taken with the Biometra BioDocAnalyze gel documenting system. The amounts of supercoiled, open-circular and linear forms of the plasmid were quantified with 1-D Main software (Advanced American Biotechnology).

DNA fragments required for genetic engineering procedures were separated in 1.0-1.5 % agarose gels in the ethidium bromide-free electrophoretic buffer containing 40 mM Tris-acetate and 0.1 mM EDTA (pH 8.3 at room temperature). The gel slices containing required DNA fragments were excised according to the ethidium-bromide stained markers. DNA was recovered by phenol extraction: the gel slices were chopped up and an equal volume of phenol saturated with pH 8.0 buffer was added. After rigorous mixing, the samples were kept for 30 min at -70 °C and centrifuged at room temperature for 15 min at 13000 g. The DNA-containing aqueous phase was removed and extracted twice with $CHCl_3$. DNA was precipitated by adding 1/10 volume of 3 M sodium acetate (pH 7.0) and equal volume of 2-propanol. Then DNA was washed using 70 % (v/v) ethanol solution, dried and dissolved in water.

2.2.1.2. Denaturing (SDS) polyacrylamide gel electrophoresis of proteins

Denaturing SDS-PAGE of proteins was employed to verify homogeneity of protein preparations. The electrophoretic buffer consisted of 25 mM Tris, 190 mM glycine (pH 8.3 at room temperature) and 0.1 % SDS (w/v). The stacking and separating gels consisted of 4 % (in 125 mM Tris-HCl (pH 6.8 at room temperature) and 0.1 % SDS (w/v)) and 12 % (in 375 mM Tris-HCl (pH 8.8 at room temperature) and 0.1 % SDS (w/v)) acrylamide/N,N'-methylenebisacrylamide (37.5:1 (w/w)), respectively (Sambrook, 1989). Protein samples were mixed 1:1 (v/v) with sample buffer (100 mM Tris-HCl (pH 6.8 at room temperature), 4 % SDS (w/v), 200 mM DTT, 20 % (v/v) glycerol) and placed for 3 min into a bath with 95 °C water before loading. Electrophoresis was run at room temperature for 1-1.5 hours at ~30 V/cm. Gels were stained with Coomasie Blue. Digital images of the gels were taken with Biometra BioDocAnalyze gel documenting system. Amounts of protein were quantified by densitometric analysis using 1-D Main software (Advanced American Biotechnology).

2.2.1.3. Non-denaturing polyacrylamide gel electrophoresis

Non-denaturing polyacrylamide gel electrophoresis (PAGE) was employed in gel-shift experiments. The Electrophoresis buffer I consisted of 30 mM MES-His (pH 6.5 at 25 °C), 0.02 mM EDTA. The gels consisted of 6 % acrylamide/N,N'-methylenebisacrylamide (29:1 (w/w)) in the Electrophoresis buffer I, polymerisation was initiated by adding TEMED and ammonium persulfate. The gels were 1 mm thick and ~20 cm in length. Prior to gel casting, one of the glass plates was processed with "bind silane" (3-methacryloxypropyltrimethoxysilane) and the other with "repeal silane" (5 % (v/v) dichlorodimethylsilane in CHCl$_3$).

The Binding buffer consisted of 30 mM MES-His (pH 6.5 at 25 °C), 0.02 mM EDTA, 0.2 mg/ml BSA 10 % (v/v) glycerol. Samples of radiolabelled

DNA, Binding buffer and protein were mixed in the wells of ELISA plate, left for 10 min at room temperature and then loaded on the gel. Electrophoresis was run at room temperature for 2 hours at ~6 V/cm. After electrophoresis the glass plate with "repeal silane" was removed and the gel was dried on the glass plate with "bind silane" under a hot air flow. Radiolabelled DNA was detected in the dried gels using BAS-MS image plates (FujiFilm) and Cyclone[TM] phosphorimager (Perkin Elmer).

2.2.2. Expression and purification of MunI

2.2.2.1. MunI plasmids and strains

The plasmids pUHE25-2-MunIR (Ap[r]) and pUHE25-2-MunI-H6G4C (Ap[r]) encoding the wt MunI and MunI variant with the C-terminal His-tag (His-tagged MunI) were expressed in *E. coli* ER2267 cells containing a compatible plasmid pMunIM 6.2 (Tc[r], Cm[r]) harboring the methyltransferase gene munIM (Siksnys et al, 1994).

2.2.2.2. Construction of the His-tagged MunI

The plasmid pUHE25-2-MunI-H6G4C encoding the MunI variant with the C-terminal His-tag (His-tagged MunI) was derived from the pUHE25-2-MunIR plasmid (Silanskas et al, 2011) by inserting a DNA fragment encoding the GGSHHHHHHGGGGC amino acid sequence immediately downstream of the C-terminal amino acid of native MunI. The DNA fragment encoding the His-tag was derived by PCR using the pRIZ′-scPvuII-H6G4C plasmid (Eisenschmidt et al, 2005) as a template.

2.2.2.3. MunI purification

ER2267 cells carrying compatible plasmids pUHE25-2-MunIR and pMunIM 6.2 (Tc[r], Cm[r]) were grown at 37 °C to late logarithmic phase in Luria broth (Sambrook, 1989) medium containing 50 mg/l ampicillin and 30 mg/l chloramphenicol. After induction with IPTG (0.2 mM, 4 h, at 37 °C cells were

harvested by centrifugation and resuspended in Equilibration buffer I (10 mM K-phosphate (pH 7.4 at 25 °C), 100 mM NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA). Crude cell extract was obtained by sonication, and cell debris was separated by centrifugation. The resulting supernatant was applied to a heparin–Sepharose (GE Healthcare, Uppsala, Sweden) column and eluted using a NaCl gradient. The fractions containing active endonuclease were pooled and dialysed against the Equilibration buffer I. Further protein purification was achieved by subsequent chromatography on blue-sepharose (GE Healthcare, Uppsala, Sweden) and phosphocellulose (Whatman) columns. Final fractions containing purified enzyme were pooled and dialysed against Storage buffer I (10 mM Tris-HCl (pH 7.4 at 25 °C), 100 mM KCl, 1 mM EDTA, 1 mM DTT and 50 % v/v glycerol) and stored at –20 °C. The enzyme was >95% homogeneous as shown by SDS-PAGE and Coomassie-Blue analysis. The protein concentration was determined spectrophotometrically at 280 nm using the extinction coefficient of 45720 M^{-1} cm^{-1} for monomer.

2.2.2.4.   His-tagged MunI purification

ER2267 cells carrying compatible plasmids pUHE25-2-MunI-H6G4C and pMunIM 6.2 (Tc^r, Cm^r) were grown at 37 °C to late logarithmic phase in Luria broth (Sambrook, 1989) medium containing 50 mg/l ampicillin and 30 mg/l chloramphenicol.. After induction with IPTG (0.2 mM, 4 h, at 37 °C cells were harvested by centrifugation and resuspended in equilibration buffer II (20 mM Na-phosphate (pH 7.4 at 25 °C), 0.5 M NaCl, 1 mM 2-mercaptoethanol). Crude cell extract was obtained by sonication, and cell debris was separated by centrifugation. The resulting supernatant was applied to a 1 ml Ni^{2+} HisTrap HP column and eluted using an imidazole gradient. The fractions containing active endonuclease were pooled and dialysed against the Equilibration buffer I (10 mM K-phosphate (pH 7.4 at 25 °C), 100 mM NaCl, 7 mM 2-mercaptoethanol, 1 mM EDTA). When protein was loaded on 1 ml HiTrap Heparin HP columns (GE Healthcare, Uppsala, Sweden) and eluted using a

NaCl gradient. Final fractions containing purified enzyme were pooled and dialysed against Storage buffer I (10 mM Tris-HCl (pH 7.4 at 25 °C), 100 mM KCl, 1 mM EDTA, 1 mM DTT and 50 % v/v glycerol) and stored at –20 °C. The enzyme was >95% homogeneous as shown by SDS-PAGE and Coomassie Blue analysis. The protein concentration was determined spectrophotometrically at 280 nm using the extinction coefficient of 45720 M^{-1} cm^{-1} for monomer.

2.2.3. Expression and purification of His-tagged Bse634I

2.2.3.1. Bse634I plasmids and strains

The plasmid pBse634IR-H6G4C (Ap[r]) encoding the Bse634I variant with the C-terminal His-tag (His-tagged Bse634I) was expressed in *E. coli* ER2267 cells containing a compatible plasmid pHpaIIM (Cm[r]) harboring the methyltransferase gene HpaIIM (Grazulis et al, 2002).

2.2.3.2. Construction of the His-tagged Bse634I

The plasmid pBse634IR-H6G4C (Ap[r]) encoding the Bse634I mutant R226A+C10S+C186S with the C-terminal His-tag (His-tagged Bse634I) was generated using the pBse634IR3.7 plasmid (Grazulis et al, 2002) as a template using PCR-based site-directed mutagenesis. The DNA fragment encoding the His-tag (GGSHHHHHHGGGGC) was derived by PCR using the pRIZ′-scPvuII-H6G4C plasmid (Eisenschmidt et al, 2005) as a template.

2.2.3.3. His-tagged Bse634I purification

ER2267 cells carrying compatible plasmids pBse634IR-H6G4C (Ap[r]) and pHpaIIM (Cm[r]) were grown at 37 °C to late logarithmic phase in Luria broth (Sambrook, 1989) medium containing 50 mg/l ampicillin and 30 mg/l chloramphenicol.. After induction with IPTG (0.2 mM, 4 h, at 37 °C cells were harvested by centrifugation and resuspended in Equilibration buffer II (20 mM Na-phosphate (pH 7.4 at 25 °C), 0.5 M NaCl, 1 mM 2-mercaptoethanol).

Crude cell extract was obtained by sonication, and cell debris was separated by centrifugation. The resulting supernatant was applied to a 1 ml Ni^{2+} HisTrap HP column (GE Healthcare, Uppsala, Sweden) and eluted using an imidazole gradient. The fractions containing active endonuclease were pooled and dialysed against the Equilibration buffer I. When protein was loaded on 1 ml HiTrap Heparin HP columns (GE Healthcare, Uppsala, Sweden) and eluted using a NaCl gradient. Final fractions containing purified enzyme were pooled and dialysed against Storage buffer II (10 mM K-phosphate (pH 7.4 at 25 °C), 100 mM KCl, 1 mM DTT, 0.1 mM EDTA, 50 % (v/v) glycerol) and stored at − 20 °C. The enzyme was >95% homogeneous as shown by SDS-PAGE and Coomassie-Blue analysis. The protein concentration was determined spectrophotometrically at 280 nm using the extinction coefficient of 35410 M^{-1} cm^{-1} for monomer.

2.2.4. MunI subunit exchange

Equal amounts of the wt and His-tagged MunI were mixed with or without oligo duplex or Ca^{2+} ions until final protein concentration of 4 μM of a dimer in Subunit exchange buffer consist of 10 mM Bis-Tris-propane (pH 6.5 or pH 8.3), 100 mM NaCl, 10 mM DTT. In this experiment were used 20 μM of specific 24/24 bp oligoduplex containing MunI recognition site and non-specific 23/23 bp oligoduplex that lack MunI target site. The final $CaCl_2$ used concentration was 10mM. After the mixing the proteins solutions were shared in to the Eppendorf tubes with 300 or 500 μl of volume, which were incubated at 25ºC and 37ºC for varies times. For analysis samples were loaded on the 1ml HiTrap Heparin column (GE Healthcare, Uppsala, Sweden) and the liner (0.1-0.8 M) NaCl gradient was performed with 60 ml of Equilibration buffer IV 10mM Na-phosphate (pH 6.0), 0.1 M NaCl, 7mM 2-mercaptoethanol (ÄKTA FPLC system, GE Healthcare, Uppsala, Sweden). Fractionation of resulted wt MunI and dimers containing one and two histidine tails were based on different their affinity to heparin sepharose. The subunits exchange rate of MunI was

monitored integrating areas of respective elution profiles and calculating the growing fraction of heterodimer (Figure 17).

2.2.5.   Gel filtration

Molecular masses of wt and caged MunI were carried out at room temperature on an ÄKTA FPLC system using a Superdex 200 HR column (GE Healthcare, Uppsala, Sweden) pre-equilibrated with Gel filtration buffer (20 mM Tris-HCl (pH 7.5 at 25 °C), 0.2 M KCl). The apparent molecular mass values of wt and caged MunI were calculated by interpolation from the standard curve obtained by using a set of proteins of known molecular mass from GE Healthcare (Uppsala, Sweden).

2.2.6.   Titration of the cysteine residues of wt MunI with DTNB

The MunI protein was dialysed against Titration buffer (10 mM Tris-HCl (pH 7.5 at 25 ºC), 200 mM KCl) to remove DTT. Titration of MunI cysteine residues (20 μM monomer) was performed by incubation with 200 μM 5,5'-dithiobis (2-nitrobenzoic acid) (DTNB) in the absence and presence of 1.5 M guanidinium monochloride (GdmCl) in Titration buffer. The reaction product TNB^{2-} was quantified spectrophotometrically using the extinction coefficient of 13600 cm^{-1} M^{-1} at 412 nm.

2.2.7.   GdmCl-induced protein unfolding of MunI

Protein denaturation experiments were carried out by mixing 1 μM (for fluorescence analysis), 8 μM (for CD analysis)and 5.4 μM (for AUC analysis) of MunI with various concentrations of GdmCl in Denaturation buffer (10 mM Tris-HCl (pH 7.5 at 25 °C), 200 mM KCl and 1 mM DTT). Samples were incubated overnight at 25 °C.

2.2.7.1. Intrinsic fluorescence measurement

Fluorescence spectra were recorded at 25 °C in the thermostatically controlled cell holder of a Perkin-Elmer LS 50 luminescence spectrophotometer. The excitation wavelength was 295 nm and the emission was recorded from 310 to 380 nm using a 5 nm band width. Differences in fluorescence intensity between native and denatured protein were measured at a fixed emission wavelength of 335 nm. Data reported here represent the average of three spectra.

2.2.7.2. Circular dichroism measurements

CD spectra were recorded at 25°C in the thermostatically controlled cell holder of a Jasco J-710 spectropolarimeter, using a cylindrical cuvette with a 0.05 cm light path. CD spectra were recorded between 185 and 250 nm. To monitor the denaturation of MunI, 10 readings of the CD signal at 220 nm were taken (with a 20 seconds averaging time) and the data averaged.

2.2.7.3. Analytical ultracentrifugation

GdmCl induced protein denaturation was monitored with an analytical ultracentrifuge by determining apparent molar masses from sedimentation equilibrium experiments. Samples contained 5.4 μM MunI mixed with different concentrations of GdmCl. A Beckman XLA analytical ultracentrifuge equipped with UV absorption optics scanning the samples at 280 nm, an 8-hole analytical rotor and 6 channel charcoal filled epon centerpieces was used. Samples were spun at 16000 rpm and 20 °C and scanned continuously until no change in the concentration profile could be detected over a period of 12 hours. Scans collected over these 12 hours were averaged and evaluated assuming a single sedimenting species and thus yielding apparent molar masses (Siksnys et al, 1999).

2.2.8. Cross-linking of MunI subunits

To perform cross-linking of the C123 residues of the MunI subunits copper phenantroline solution (0.2 mM $CuSO_4$ and 0.8 mM phenantroline in Equilibration buffer III) was used. Since C123 residues are located at the dimer interface 1 M of GdmCl was added to disrupt protein into monomers and get accessibility of Cys residues. Reaction was performed 1 hour at room temperature and stopped by addition 100 mM of EDTA in Equilibration buffer III. After this, reaction mixture was dialyzes against Equilibration buffer III to remove GdmCl.

2.2.8.1. Purification of the cross-linked MunI

Separation of the cross-linked MunI from unreacted wt protein was performed under non-reducing conditions with the Equilibration buffer III (10 mM Na-phosphate (pH 7.4 at 25 °C), 0.1 M NaCl) on the 1ml HiTrap Heparin column (GE Healthcare, Uppsala, Sweden). After the running of NaCl gradient cross-linked MunI eluted previously than wt MunI. SDS-PAGE analysis of reaction products indicated the crosslinked MunI as a dimer. The yield of resulted cross-link was more than 95%. Fractions containing cross-linked MunI were pooled, dialysed against Storage buffer I without DTT (10 mM Tris-HCl (pH 7.4 at 25 °C), 100 mM KCl, 1 mM EDTA and 50 % v/v glycerol) and stored at -20 °C.

2.2.9. Caging of MunI

Stock solution of the 2-nitrobenzyl bromide (NBB) (Sigma-Aldrich) was prepared in dimethylsulfoxide (DMSO). To remove DTT from the storage buffer first MunI protein was dialysed against Equilibration buffer III (10 mM Na-phosphate (pH 7.4 at 25 °C), 0.1 M NaCl). Caging of MunI was performed by overnight incubation of 60 μM protein (monomer concentration) with 500 μM NBB in the presence of 1 M GdmCl at 4 °C. 1M of GdmCl was added to denaturate MunI until monomers and let to be accessible for cysteines at the

dimmer interface. After the reaction, unreacted NBB was quenched with 1 mM DTT. To remove unreacted NBB from the solution the reaction mixture was dialysed against Equilibration buffer III.

2.2.10. Purification of the caged MunI

Caged MunI (monomer) was separated from the unreacted wt MunI (dimmer) by gel filtration using a Superdex 200 HR column (GE Healthcare, Uppsala, Sweden) equilibrated with Gel filtration buffer (20 mM Tris-HCl (pH 7.5 at 25 °C), 0.2 M KCl). The yield of resulted monomer was more than 95%. Fractions containing caged MunI were pooled, dialysed against Storage buffer I without DTT (10 mM Tris-HCl (pH 7.4 at 25 °C), 100 mM KCl, 1 mM EDTA and 50 % v/v glycerol) and stored at -20 °C.

2.2.11. Decaging of MunI

Decaging of caged MunI (4 μM monomer) was performed by irradiation with UV light of different wavelengths for 1 hour in Decaging buffer (100 mM Na-acetate (pH 5.8 at 25 °C), 0.2 M KCl and 20 mM DTT) at 4 °C. The Fluoromax-3 (Jobin Yvon, Stanmore, UK) spectrofluorimeter equipped with a Xe lamp was used as UV source. After irradiation, the protein sample was analysed by gel filtration with a Superdex 200 HR column (GE Healthcare, Uppsala, Sweden) equilibrated with Gel filtration buffer. The best efficiency of decaging was obtained by irradiation with a wavelength of 330 nm (29 nm band width) (**Figure 25**).

2.2.12. Generation and purification of REase-TFO cross-links

2.2.12.1. Generation of caged MunI-TFO

Each monomer of the His-tagged MunI contains two cysteines: one Cys residue at the dimer interface what should be caged and a second Cys residue at the C-term what is needed for protein coupling with TFO. To perform both Cys modifications in the correct way three steps of procedures were performed.

In the first step C-terminal Cys residues were protected from thiol-reactive reagents by forming intersubunit disulfide bonds. In order to do this His-tagged MunI was loaded on the 1 ml HiTrap Heparin HP column (GE Healthcare, Uppsala, Sweden) and washed with Equilibration buffer III (10 mM sodium phosphate (pH 7.4 at 25 °C), 100 mM NaCl) to remove DTT. Next, copper phenantroline solution (0.2 mM $CuSO_4$ and 0.8 mM phenantroline in Equilibration buffer III) was loaded on the column to stimulate formation of intersubunit disulfide bonds between the C-terminal Cys residues (Figure 31A). In this step Cys residues at the dimmer interface are located in the safe distance and do not form disulfide bonds. After 10 min incubation, copper phenantroline was washed from the column with 100 mM EDTA in Equilibration buffer III and protein was eluted using 0.1-1.0 M NaCl gradient. In the second step the resultant protein was partially denatured with 1 M GdmCl and the exposed dimerization interface cysteines (Cys123) were modified by caging with NBB as described previously (see 2.2.9.) (Figure31B). The reaction mixture was dialyzed against Equilibration buffer III supplemented with 50 mM DTT to quench unreacted NBB and reduce disulfide bonds between C-terminal cysteines. Resultant monomeric caged MunI was separated from the uncaged homodimer on the Heparin HP column (GE Healthcare, Uppsala, Sweden) using a 0.1-1.0 M NaCl gradient in Equilibration buffer IV with 0.2 mM tris(2-carboxyethyl)phosphine (TCEP). Finally, in the third step, the regenerated thiol groups of caged MunI were used for coupling with TFO according to the method described Eisenschmidt et al. (Eisenschmidt et al, 2005) (Figure31C). To activate TFO 90 µM of nucleotide was incubated with 20 mM bifunctional cross-linker N-(γ-maleimidobutyryloxy) succinimide ester (GMBS, Sigma-Aldrich) in 900 µL 100 mM Na-phosphate (pH 7.4 at 25 °C) for 1 hour at 25 °C. The stock solution of the GMBS was prepared in dimethylsulfoxide (DMSO). After the reaction unreacted GMBS was removed by passage through NAP-5 and subsequently NAP-10 desalting columns (GE Healthcare, Uppsala, Sweden).

Desalted and activated TFO was dried and dissolved in 500 µL of Cross-linking buffer (100 mM K-phosphate (pH 7.0 at 25 °C), 100 mM KCl). To perform coupling of oligonucleotide to protein 50 µM of activated TFO was incubated with 16 µM caged MunI in 900 µL of Cross-linking buffer for 1 hour at 25 °C.

2.2.12.2. Generation of Bse634I-TFO

His-tagged Bse634I used for the TFO1 and TFO2 coupling reactions was dialysis against 10 mM sodium phosphate (pH 7.0 at 25 °C), 100 mM NaCl and 0.2 mM TCEP to remove DTT. After, coupling with TFO1 and TFO2 was performed (see 2.2.12.1.).

2.2.12.3. Purification of REase-TFO cross-links

The resultant caged MunI-TFO and Bse634I-TFO cross-links were purified on a MonoQ 5/50 GL column (GE Healthcare, Uppsala, Sweden) using 20 mM Tris-HCl (pH 7.0 at 25 °C) and a linear gradient of 0.1-1.0 M NaCl. Unreacted proteins eluted in the flow through whereas REase-TFO conjugates and excess of unreacted TFO eluted as separate peaks. Fractions containing MunI-TFO and Bse634I-TFO conjugates were dialyzed against the MunI and Bse634I storage buffers mentioned above and stored at –20 °C. Conjugate concentrations were determined by densitometric analysis of SDS-PAGE gels containing samples of REase-TFO conjugates along with known amounts of His-tagged proteins.

2.2.13. DNA cleavage activity by cross-linked MunI

DNA cleavage activity of cross-linked MunI was tested using the pUCGK-4 plasmid (provided by G. Kruckas) containing a single copy of the MunI recognition sequence. DNA cleavage was performed by incubation of 0.75 nM of wt and cross-linked MunI with 1.5 nM of plasmid DNA in Reaction buffer I (10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM MgCl$_2$, 50 mM NaCl, 0.1 mg/ml

BSA) at 25 ℃. The reaction was quenched with Loading dye solution II containing 50 mM EDTA (pH 8.0 at 25 °C), 0.1% SDS, 50% v/v glycerol, 0.01% bromphenol blue and analysed by agarose gel electrophoresis.

2.2.14. DNA cleavage activity by caged and decaged MunI

DNA cleavage activity of wt, caged and decaged MunI was tested using bacteriophage λ DNA as a substrate. 400 nM (dimer concentration) of protein was incubated in 50 μL with 1μg of λ DNA in Reaction buffer I (10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM $MgCl_2$, 50 mM NaCl, 0.1 mg/ml BSA) for 15 min at 37 ℃.

The pUCGK-4 plasmid (provided by G. Kruckas) containing a single copy of the MunI recognition sequence was used in DNA cleavage experiments. DNA cleavage was performed by incubation of 0.75 nM wt MunI and 1.5 nM of caged and decaged MunI, respectively, with 1.5 nM of plasmid DNA in Reaction buffer I (10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM $MgCl_2$, 50 mM NaCl, 0.1 mg/ml BSA) at 25 ℃.

All samples were collected at given time intervals, quenched with the Loading dye solution II (50 mM EDTA (pH 8.0 at 25 °C), 0.1% SDS, 50% v/v glycerol, 0.01% bromphenol blue) and analysed by agarose gel electrophoresis. The amount of supercoiled (SC), open-circular (OC), and linear (FLL) DNA forms was evaluated by densitometric analysis of ethidium bromide-stained gels.

2.2.15. Addressing and DNA cleavage activity by REase-TFO cross-links

DNA cleavage activity of caged MunI-TFO conjugate was tested on an equimolar mixture of the "addressed" DNA substrate (1880 bp DNA fragment containing a single MunI site separated by 9 bp from two symmetrically positioned triplex forming sequences 5′-GAGAGAGAGAGAAAAAA-3′) and the "non-addressed" substrate (3675 bp DNA fragment containing a single MunI

site and no TFS). "Addressed" and "non-addressed" substrates were obtained by cleavage pECFP-MunI plasmid (derived from pECFP-PvuII) with VspI and NsbI. 100 nM of both DNA fragments were preincubated overnight with 400 nM of caged MunI-TFO at 4 °C in Triplex buffer I (80 mM Tris-phosphate (pH 6.1 at 25 °C), 1.0 M NaCl and 4 mM Mg^{2+}, 1 mM DTT). To activate the MunI-TFO conjugate, the reaction mixture was supplemented with 20 mM of DTT, placed into a quartz cuvette and irradiated with 330 nm UV light for 1 hour at 4 °C. A Fluoromax-3 (Horiba Jobin Yvon, Stanmore, UK) spectrofluorimeter equipped with a Xe lamp was used as the UV source. DNA cleavage was initiated by 10-fold dilution with the Reaction buffer II (80mM Tris-phosphate (pH 7.1 at 25 °C), 4 mM Mg^{2+}, 1 mM DTT) and incubated at 20 °C. Samples were collected at given time intervals, quenched with Loading dye solution II (50 mM EDTA (pH 8.0 at 25 °C), 0.1% SDS, 50% (v/v) glycerol, 0.01% bromphenol blue) and analyzed by agarose electrophoresis.

DNA cleavage activity of the Bse634I-TFO cross-links was studied on a linearized 5.5 kbp pECFP-Bse634I plasmid bearing seven Bse634I recognition sites, one of which (the "addressed" site) is separated by 9 bp from the triplex forming sequences 5′-TTTTTTCTCTCTCTC-3′ (TFS1) and 5′-AGAAGAAGAAAGAAA-3′ (TFS2). To analyze "addressed" cleavage, 50 nM of linear DNA was incubated with equimolar mixture of Bse634I-TFO1 and Bse634I-TFO2 conjugates (total concentration 200 nM monomer) at 37 °C in Triplex buffer II (80 mM Tris-phosphate (pH 6.1 at 37 °C), 2 mM spermine, 2 mM $MgCl_2$ and 10 mM DTT. The reactions were quenched and analyzed as described above.

2.2.16.   DNA binding studies

DNA binding by wt, cross-linked, caged and decaged MunI was analyzed by a gel mobility-shift assay using ^{33}P-labelled cognate 22/19 oligoduplex, containing single recognition sequence of MunI, and non-cognate 22/22 oligoduplex that lack MunI recognition sequence. Different amounts of protein

were incubated with 2 nM of cognate or non-cognate oligonucleotide duplex for ten minutes at room temperature in 20 μL of the Binding buffer (30 mM MES-His (pH 6.5 at 25 ℃), 0.02 mM EDTA, 10 % v/v glycerol, 0.2 mg/ml BSA). Samples were loaded onto 6% (w/v) acrylamide gels (29:1 (w/w) acrylamide/bisacrylamide) and run in Electrophoresis buffer I (30 mM MES-His (pH 6.5 at 25 ℃), 0.01 mM EDTA) for two hours at ~6 V/cm. After electrophoresis, all gels were analysed using a Cyclone Storage Phosphor System with OptiQuant Image Analysis Software, version 3.00 (Perkin-Elmer).

3. Results

3.1.   MunI restriction enzyme as an alternative for generation of REase-TFO conjugates

Restriction endonucleases seems to be ideally suited for the introduction of DSB's near the defective locus in the gene to increase the frequency of HR, however their application is limited due to the short target sequences. In a pioneering study Eisenschmidt et al has demonstrated that a single-chain variant of the homodimeric REase PvuII scPvuII-TFO conjugate can be specifically addressed to the defined PvuII site in DNA (Eisenschmidt et al, 2005) (see 1.3.4.). In this work authors used a single-chain variant of the homodimeric REase since monomeric protein allows introduction of a unique single Cys for the coupling of the TFO. Conjugation of the scPvuII with the TFO extended their specificity allowing the recognition of the bipartite target site consisting of the PvuII target site and TFS. However, scPvuII has a limitation since not every REase could be redesigned into a single chain variant. Therefore, we have looked for other restriction enzymes as alternatives for PvuII.

Restriction endonuclease MunI was identified in *Mycoplasma unidentified* strain (Stakenas et al, 1992). It recognizes the hexanucleotide sequence C/AATTG and cleaves it in the presence of Mg^{2+} ions as indicated by "/" (Deibert et al, 1999; Stakenas et al, 1992). MunI is arranged as a dimer and each subunit contains a single active site (Deibert et al, 1999). The structure of MunI-DNA complex has been solved by X-ray crystallography (Figure 16A) (Deibert et al, 1999). This makes MunI an attractive model system to develop restriction enzymes into gene targeting tools.
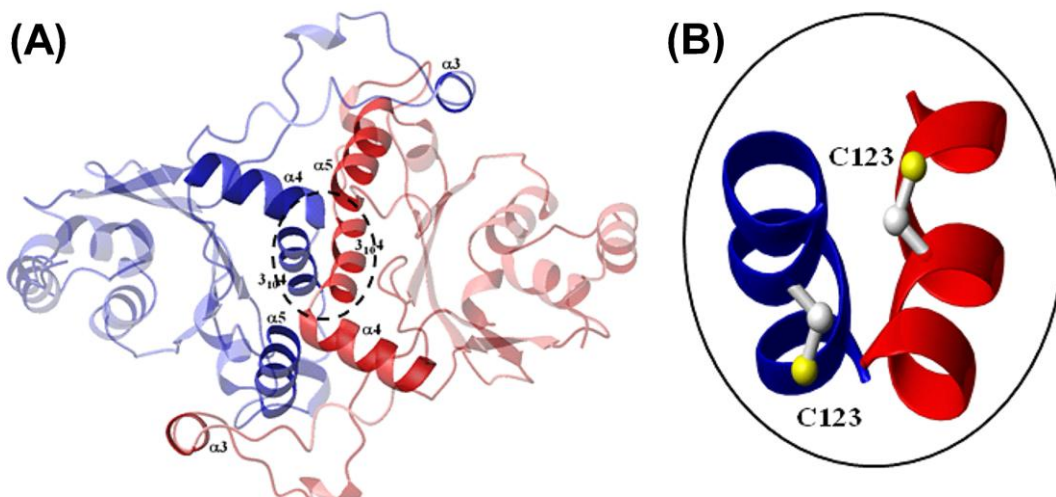
Figure 16. Dimer interface of MunI. (A) Crystal structure of MunI restriction endonuclease. Ribbon representation of the two monomers shown in blue and red, respectively. Helices involved in protein dimerization are shown in more intense color. The region blown-up in (B) is circled. (B) Close-up view of a region of the MunI dimerization interface showing the cysteines residues (C123) subjected to caging.

To probe whether an orthodox dimeric restriction enzyme can be engineered into a gene targeting tool we used MunI as a model system. In theory, a single TFO can be attached to the dimeric restriction enzyme if the subunit exchange within the dimer is slow. Therefore, first we investigated the subunit exchange rate of MunI.

3.1.1. MunI subunit exchange

Different experimental approaches are used to monitor subunit exchange in oligomeric proteins, including Förster resonance energy transfer (FRET) between the subunits mapped with fluorescent donor and acceptor (Merickel et al, 2002), a real-time nanoelectrospray mass spectrometry (Sobott et al, 2002), ion exchange chromatography (Schneider et al, 2001) or size exclusion chromatography when subunits exchange between two proteins having different molecular mass (Park et al, 2004).

To monitor subunit exchange of MunI restriction endonuclease we applied a heparin affinity chromatography. Equal amounts of the wt and the C-terminal His-tagged MunI were mixed together and incubated for 24 hours in the

Subunit exchange buffer at pH 6.5 or pH 8.3 at 25 ºC or 37ºC. His-tagged and wt MunI variants were separated on a heparin sepharose column using a buffer with pH 6.0 (Figure 17A). At this pH value, His-tagged MunI variants showed much stronger interaction with heparin sepharose than wt MunI, most likely due to the increased positive charge of the protein due to the protonization of histidines in the His-tag. After the elution with NaCl gradient three peaks were recovered (Figure 17B). The SDS-PAGE analysis of the fraction corresponding to the middle peak revealed the presence of wt and His-tagged proteins in equimolar amounts, suggesting the formation of the heterodimer (Figure 17C).

To monitor subunit exchange rate of wt and His-tagged MunI, protein mix was incubated at different temperatures for various times. After the incubation samples were loaded on the 1 ml HiTrap Heparin HP column (GE Healthcare, Uppsala, Sweden) and eluted using NaCl gradient to separate different protein forms. The subunit exchange rate of MunI was determined by integrating respective peak areas at different time intervals (Figure 17C).

Effect of temperature on subunit exchange. In contrast to EcoRV restriction enzyme, which does not exchange subunits at 37ºC (Stahl et al, 1996), we detected subunit exchange in MunI dimer. After 20 hours incubation of wt and His-tagged MunI at 37ºC full subunit exchange was reached (Figure 17D). MunI subunit exchange rate showed no significant dependence on the pH however it decreased dramatically at lower temperature. When the mixture of wt and His-tagged MunI was incubated at 25ºC, after 24 hours the fraction of heterodimer was just about 1%.
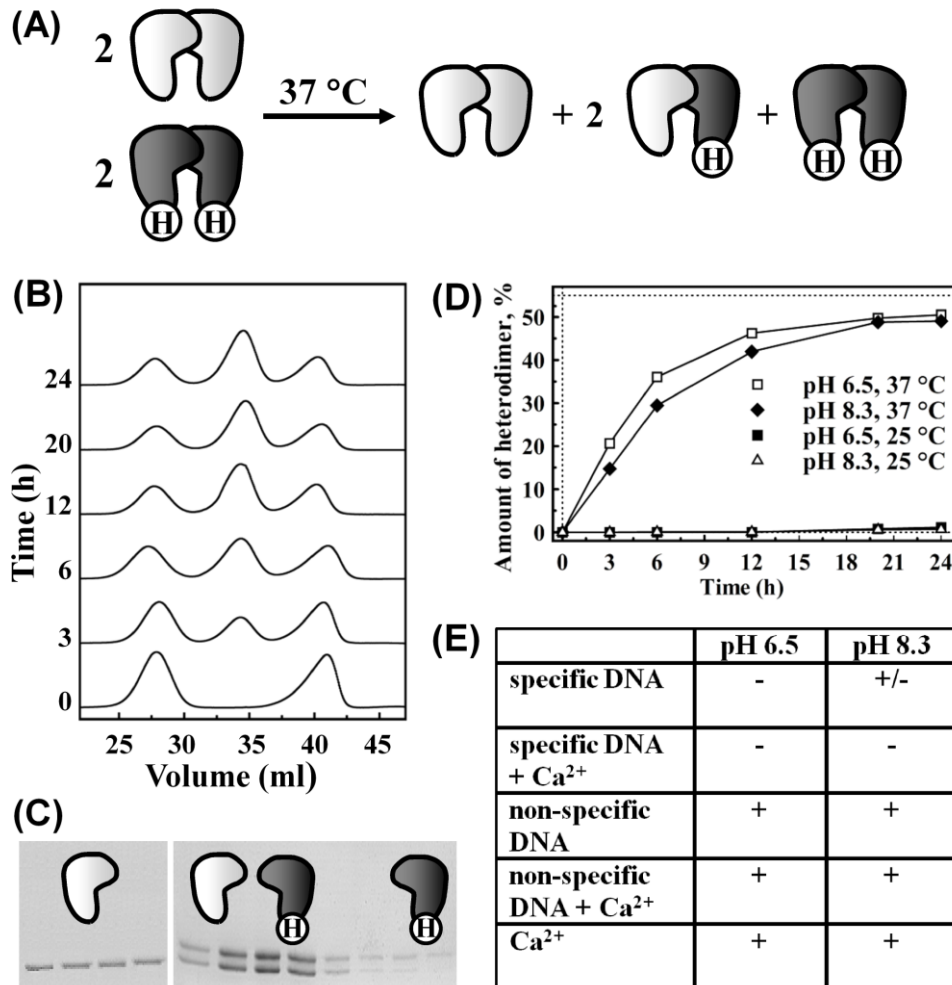
Figure 17. Subunit exchange of MunI. (A) The scheme of the subunit exchange. (B) The elution profiles of the wt and His-tagged MunI mixture at 37°C monitored on 1 ml HiTrap Heparin column (GE Healthcare, Uppsala, Sweden). (C) The SDS-PAGE analysis of eluted fraction after MunI 24 h subunit exchange at 37°C. (D) Subunit exchange dependence on temperature at different pH of the buffer. (E) Subunit exchange dependence on the presence of DNA and Ca^{2+} at different pH of the buffer: (+) – subunits exchange, (-) – subunits do not exchange.

Effect of DNA on subunit exchange. To probe whether DNA binding affects the MunI dimer stability, the subunit exchange of MunI was analysed at 37°C in the presence of 24 bp specific oligoduplex containing target site CAATTG or 23 bp non-specific oligoduplex (Figure 17E). No MunI subunit exchange was detected in the presence of the specific DNA at pH 6.5, however MunI incubation with non-specific DNA did not prevent subunit exchange and heterodimer formation. It is worth to note that at pH 8.3 at 37°C specific oligoduplex did not fully prevent subunit exchange since ~20% of heterodimer was formed under these conditions. Electrophoretic mobility shift assay

76

(EMSA) experiments demonstrated that restriction endonuclease MunI binds specific DNA in pH 6.5 buffer with high affinity whereas at pH 8.3 binding is much weaker (Haq et al, 2001; Lagunavicius et al, 1997). It is likely that at pH 6.5 tight binding to specific DNA stabilizes MunI dimer and prevents subunit exchange, therefore no heterodimer is formed under these conditions. Due to the weaker specific DNA binding at pH 8.3, nearly 1000-fold increase of MunI and DNA concentrations is required to slow down subunit exchange.

Effect of Ca^{2+} ions on subunit exchange. It has been shown that at pH 8.3 Ca^{2+} ions simulate MunI binding to the cognate DNA but have no effect on the binding affinity at pH 6.5 (Lagunavicius et al, 1997). Therefore, MunI subunit exchanged has been analysed in the presence or absence of specific and nonspecific DNA at pH 6.5 and pH 8.3 in the presence of Ca^{2+} ions (Figure 17E). We have found that at pH 8.3 Ca^{2+} ions in the presence of cognate DNA effectively prevent MunI subunit exchange, however Ca^{2+} ions have no effect on the subunit exchange in the presence of non-cognate oligoduplex at pH 6.5 or 8.3.

Taken together, our data indicate that subunit exchange in the MunI dimer occurs at higher temperature (37ºC), however decrease of the temperature (<25ºC) or addition of the cognate DNA stabilize the dimer and prevent subunit exchange.

3.1.2. Investigation of the MunI dimerization interface

In order to develop the MunI variant with a controllable catalytic activity we made use of the crystal structure which revealed that single Cys residue of MunI resides in the dimer interface (Figure 16B). We proposed two different possibilities to regulate the activity of MunI: to restrict their subunits by cross-linking them to each other during Cys residues at the dimer interface (Figure 16B) or to disrupt the dimer into inactive monomers modifying the same Cys residues with photoremovable caged compound. Generated new REase-TFO conjugates with the possible activity regulation would be one step closer to the

practical application of the restriction enzymes in gene editing. Before starting any works with MunI activity regulation it is useful to study the accessibility of Cys at the dimer interface and investigate in detail the stability of this protein.

3.1.2.1.   Accessibility of Cys at the dimer interface of MunI to chemical compounds

Ellman's reagent (5,5'-dithiobis-(2-nitrobenzoic acid) or DTNB) is a chemical used to quantify the number or concentration of thiol groups in a sample (Ellman, 1959). Titration of Cys with DTNB results in yellow reaction product TNB^{2-}, which could be monitored spectrophotometrically at 412 nm (Figure 18A) (Siksnys & Pleckaityte, 1992). Accessibility of MunI Cys residues to DTNB has been analysed by mixing of 20 μM MunI (in terms of monomer) with 200 μM DTNB in Titration buffer. The experiment showed that in the wt protein Cys are not accessible at the dimer interface (Figure 18B.).
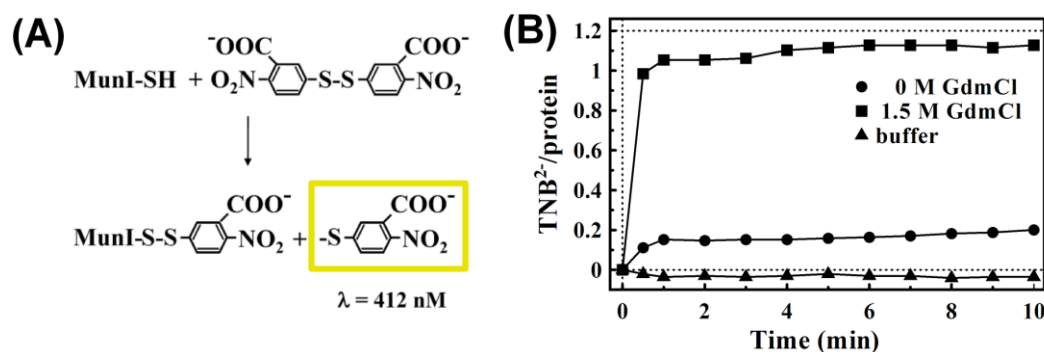

Figure 18. Accessibility of Cys at the MunI dimer interface to a chemical reagent. (A) Reaction of MunI Cys residue with DTNB was monitored spectrophotometrically at λ= 412 nm. (B) Determination of the number of modified Cys per monomer of MunI in the presence and in the absence of GdmCl.

We assumed that dissociation of the MunI dimer into monomers by denaturing agents should increase Cys accessibility to DTNB. Therefore, we performed the Cys titration by DTNB in the presence of 1.5 M of GdmCl. In this case Cys was fully accessible to DTNB. To determine whether MunI denaturation induced by GdmCl proceeds through MunI dissociation into monomers, we studied MunI denaturation pathway using different methods.

3.1.2.2.   GdmCl-induced unfolding of MunI

MunI unfolding induced by GdmCl was monitored by the intrinsic protein fluorescence, circular dichroism (CD) and analytical ultracentrifugation (AUC) (Figure 19). MunI contains six tryptophan residues that could be exited at 295 nm to provide a strong fluorescence emission signal at 335 nm. In the unfolded state tryptophan residues residing in the hydrophobic protein interior are exposed into hydrophilic environment resulting in a decreased emission signal. To monitor the effect of GdmCl on the intrinsic MunI fluorescence, we measured emission changes at 335 nm where the difference between the native and fully denatured forms of the protein is most pronounced. The MunI fluorescence signal decreased in two discrete stages: the first occurred between 0 to 1 M GdmCl; the second - between 1.5 to 2.5 M GdmCl (Figure 19A).

To investigate the influence of GdmCl on the secondary structure of MunI, the unfolding transitions were also monitored by CD. Far-UV CD spectra at 222 nm, where the difference between the native and fully denaturated α-helices is most pronounced, revealed that the molar ellipticity of MunI changed in the range from 1.5 to 3 M GdmCl whereas from 0 to 1.5 M GdmCl there were no significant changes (Figure 19B). The changes in the CD signal indicate that GdmCl induced MunI unfolding is accompanied by the loss of α-helical structural elements. This finding suggests that the fluorescence decrease from 1.5 to 2.5 M GdmCl corresponds to MunI unfolding (Figure 19A and 19B).

Next, we performed equilibrium sedimentation study of MunI at different GdmCl concentrations. AUC data showed that the molecular mass of MunI dropped from 51 kDa to 26 kDa between 0 M to 1 M GdmCl and further increase of GdmCl concentration up to 3.3 M did not significantly change the molecular weight of protein (Figure 19C). According to these data, the first stage of the fluorescence decrease in Figure 19A corresponds to the MunI dimer dissociation. Additional studies revealed that the GdmCl-induced denaturation of MunI is completely reversible and that the refolded MunI has

the same catalytic and DNA binding properties as the native protein (data not shown).

In the case of an oligomeric protein the dissociation into monomers might occur prior to unfolding of the polypeptide chain (Miller & Schildbach, 2003), or both processes can occur concertedly (Barry & Matthews, 1999). GdmCl-induced denaturation studies show that MunI first dissociates into monomers followed by monomer unfolding at increased GdmCl concentrations (Figure 19D). A similar denaturation pathway has been reported for the tetrameric restriction endonuclease Bse634I (Zaremba et al, 2005). On the other hand, dissociation and unfolding of the dimeric restriction endonucleases BfiI and PvuII occur simultaneously (Dupureur & Dominguez, 2001; Zaremba et al, 2004).
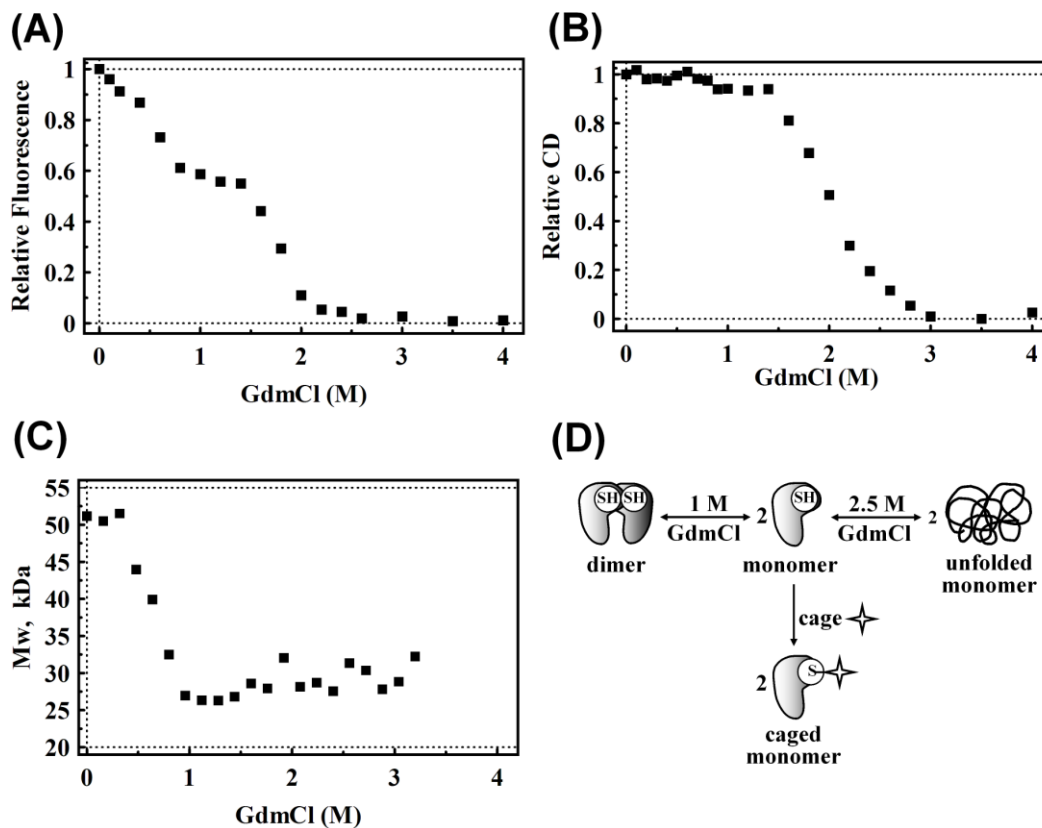

Figure 19. Unfolding of MunI induced by GdmCl. Unfolding was monitored by (A) fluorescence emission at 335 nm (λ_{exc}=295 nm), (B) far-UV CD at 222 nm and (C) analytical ultracentrifugation. (D) Schematic representation of the MunI denaturation pathway in the presence of GdmCl.

According to the denaturation pathway of MunI (Figure 19D) the Cys123 residues located at the dimer interface should be accessible for chemical modification in the presence of 1.0-1.5 M of GdmCl when MunI dimer is dissociated into monomers. Indeed, titration of MunI with DTNB in the presence of 1.5 M GdmCl showed that the Cys123 residues are accessible and reactive (Figure 18B) allowing their modification for activity regulation purpose.

3.2.   Regulation of the MunI activity

The pioneering proof-of-concept demonstration of REase-TFO conjugate was presented by Eisenschmidt and colleagues (Eisenschmidt et al, 2005) (Figure 2D). A single-chain variant of PvuII REase was fused to TFO to produce a conjugate that cuts a specific DNA site *in vitro*. Kinetics of triple helix formation is much slower than DNA binding and cleavage by REases. As Mg^{2+} ions, that are usually required for REase activity, normally are present in cellular milieu, they would trigger DNA cleavage by the restriction enzyme before the TFO-TFS interaction has occurred. For *in vivo* purposes, it would be desirable to generate REase-TFO conjugates with the switchable endonuclease activity in such a way that it could be activated only after stable DNA triplex has been formed. Here, we demonstrate the ability to regulate the activity of restriction enzyme MunI i) by intersubunit cross-linking through the Cys residues at the dimer interface and ii) by caging Cys residue with a photoremovable cage compound.

3.2.1.   Cross-linking of MunI subunits

To introduce the disulfide bond between Cys123 residues in the neighbouring MunI subunits (Figure 16B) first it is necessary to disrupt the dimer into monomers to provide Cys accessibility for a chemical reagent. Disruption of MunI dimer was performed in the presence of 1 M GdmCl (Figure 19D) and copper phenantroline was used to facilitate oxidation of Cys

residues to yield a disulfide bond. This reaction results in efficient formation (the yield is close to 90% according SDS-PAGE analysis, Figure 20B, lane R) of the intersubunit disulfide bond.
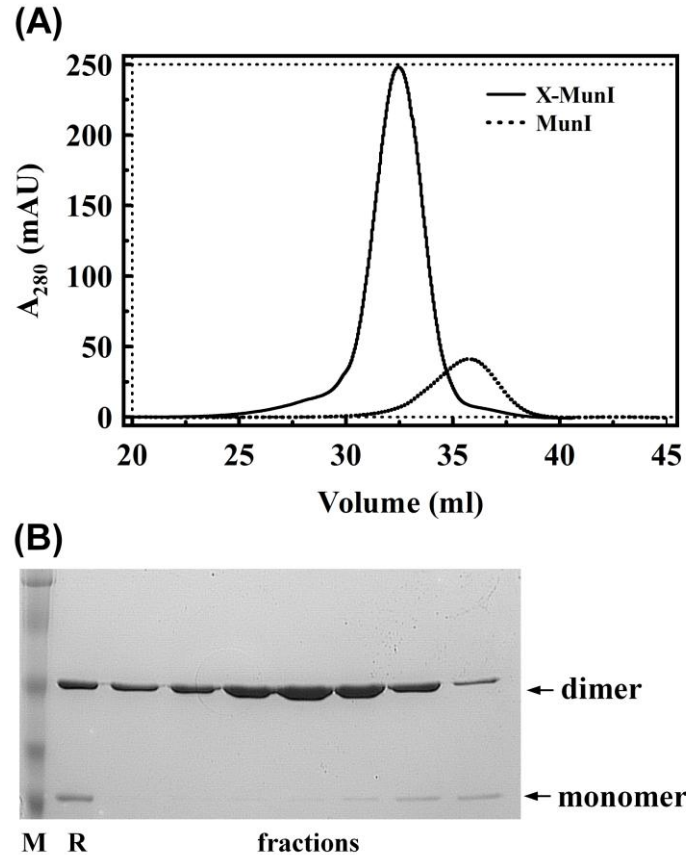

Figure 20. Intersubunit cross-linking of MunI. (A) Elution profile of the cross-linked MunI on HiTrap Heparin column (GE Healthcare, Uppsala, Sweden). Solid line corresponds to the cross-linked MunI, dashed line indicates the profile of the wt MunI. (B) SDS-PAGE analysis of the cross-linked MunI: lane M – marker, lane R – reaction mixture and fraction eluted from the heparin sepharose.

The cross-linked MunI (X-MunI) dimer was separated from the wt MunI using heparin-sepharose affinity chromatography. We found that X-MunI shows weaker interaction with heparin sepharose and elutes from the column at lower NaCl concentration before wt MunI (Figure 20A). Purification of X-MunI was performed under non-reducing conditions and the SDS-PAGE analysis of the reaction products showed that under denaturing conditions wt protein is a monomer while the cross-linked protein is a dimer (Figure 20B).

3.2.1.1.   DNA cleavage by X-MunI

To determine the catalytic activity of the X-MunI, DNA cleavage experiments were performed using supercoiled plasmid DNA that contains a single copy of the MunI recognition sequence.
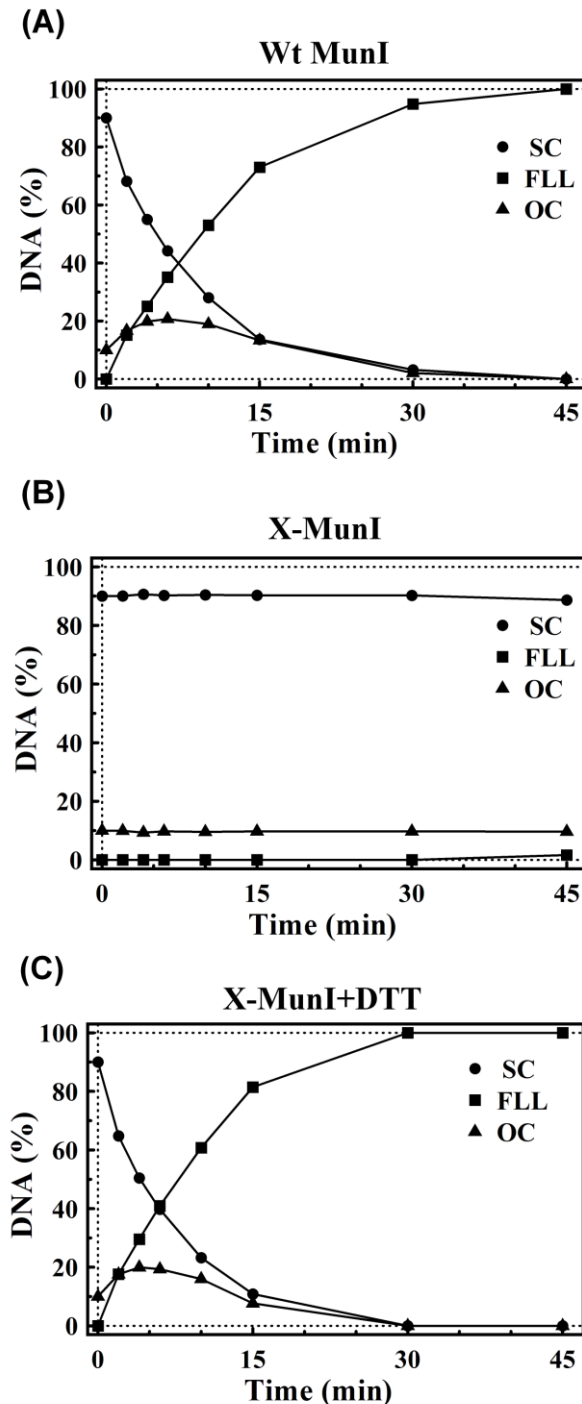
(A)

(B)

(C)

Figure 21. **Catalytic activity of X-MunI.** Panels (A), (B) and (C) show cleavage patterns of plasmid DNA containing a single MunI recognition site by the wt, X-MunI and X-MunI+DTT, respectively. Kinetic experiments were performed by using 0.75 nM of the proteins and 1.5 nM of plasmid DNA. Following forms of the DNA are monitored during the reaction: supercoiled DNA (SC); open circular DNA (OC); full-length linear DNA (FLL).

The X-MunI and plasmid DNA at 1:2 ratio were incubated in the Reaction buffer I at 25 ºC, samples were withdrawn at different time intervals and

83

analysed in the agarose gel. Analysis of the reaction products revealed that in contrast to the wt protein (Figure 21A) the X-MunI exhibited no cleavage activity (Figure 21B). However, addition of DTT to the reaction mixture fully restored MunI catalytic activity (Figure 21C).

3.2.1.2. DNA binding by the X-MunI

To determine the effect of cross-linking on MunI binding to specific DNA, we performed gel mobility-shift assay in polyacrylamide gel (PAAG). Two ^{33}P-labelled oligoduplexes (22/19 and 22/22), containing and lacking recognition sequence of MunI, respectively, were used in the binding experiment. In order to monitor the formation of the enzyme-DNA complexes, increasing concentrations of MunI were preincubated with 2 nM of cognate or non-cognate DNA, loaded on a PAAG and subjected for electrophoresis. Gel shift analysis revealed that in contrast to the wt MunI that binds cognate DNA with the $K_d \sim 10$ nM (Figure 22A) and does not specifically bind non-cognate oligoduplex (Figure 22B), cross-linked enzyme shows very low binding to specific DNA with the $K_d > 2500$ nM (Figure 22C). Binding studies suggest that X-MunI lost the catalytic activity because of dramatically reduced binding to specific DNA. Indeed, addition of the DTT reduced the disulfide bond and DNA binding was restored ($K_d \sim 25$ nM (Figure 22D).

The loss the DNA cleavage activity of X-MunI presumably results from the conformational restrictions introduced at the dimer interface by the disulphide bond. In the crystal structure of MunI-DNA complex (Deibert et al, 1999) the sulphur atoms of the –SH groups of both Cys are located at ~ 8 Å distance and point to the opposite directions (Figure 16B).
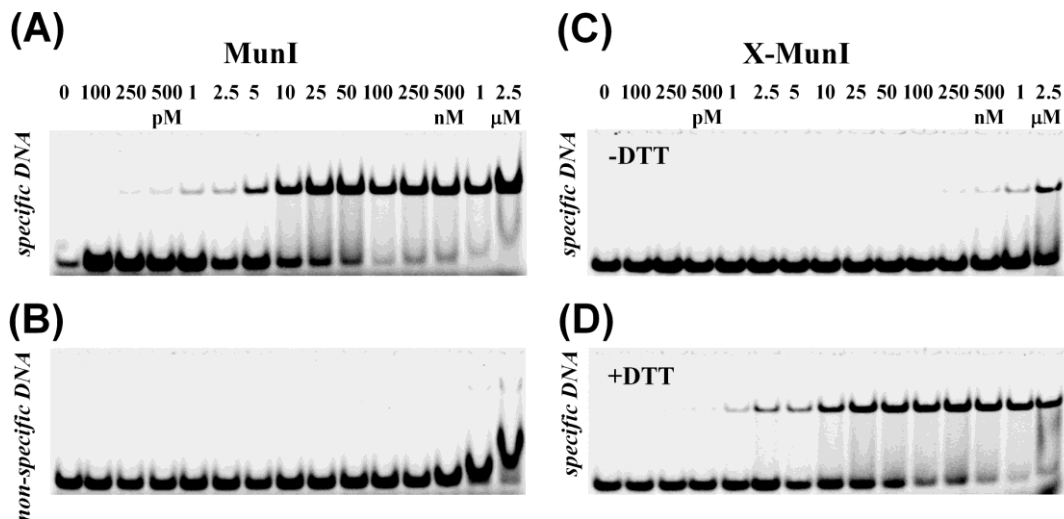
Figure 22. Binding of wt and X-MunI to specific and non-specific DNA. The labelled oligonucleotides with and without recognition sequence of MunI (at 2 nM final concentration) were incubated with increasing concentrations of wt, X-MunI and X-MunI+DTT and analyzed by gel mobility-shift assay. After electrophoresis, the gels were dried and subjected to autoradiography and analysed using Cyclone Storage Phosphor System and OptiQuant Image Analysis Software, Version 03.00 (Packard Instrument Co).

The length of disulphide bond is ~2.1 Å, therefore oxidation of Cys residues at the dimer interface to yield a disulphide bond requires significant conformational changes that may disrupt dimer interface and compromise MunI binding and cleavage activities. These changes, however, are reversible, since DTT treatment restores MunI functional activity. Thus, disulphide linkage at the MunI dimer interface acts as a red-ox switch which regulates MunI catalytic activity.

3.2.2. Caging of MunI

Regulation of proteins by light is a new and promising strategy for the external control of biological processes. This could be achieved by caging of biomolecules with photoremovable compounds. In this way it is probable that introduction of a bulky photoremovable cage compound at the MunI dimer interface would disrupt the functionally important dimer by forming inactive monomers (Figure 23A, B). As a target for caging experiments were chosen the same C123 residues at the dimer interface that were used for intersubunits

cross-linking of MunI. These Cys are located in the dimerization helices close to each other and look attractive for chemical modification. For caging was used NBB that conjugates to Cys residues and dissociates from the protein after irradiation with UV light leaving unmodified Cys residues.
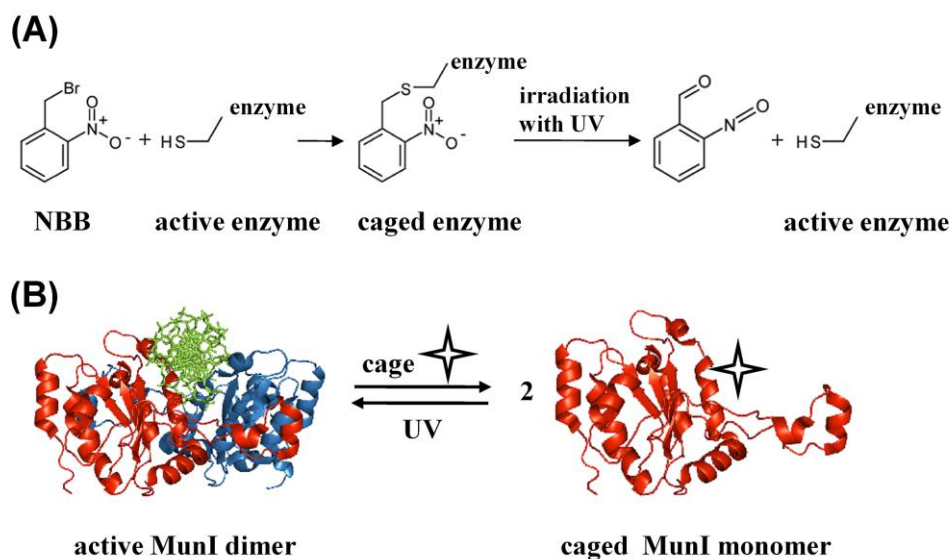

Figure 23. Generation of photoswitchable restriction endonucleases. (A) The scheme of caging of the enzyme with NBB and reactivation by irradiation with UV light. (B) Generation of photoswitchable MunI. A Cys residue located at the dimer interface of wt of MunI (residue 123) is modified using NBB. The resulting caged enzyme is an inactive monomer that is unable to bind to DNA. Irradiation at 330 nm leads to the release of the caging group and formation of the active MunI dimer.

Caging of MunI was performed by overnight incubation with NBB in the presence of 1 M GdmCl at 4°C. GdmCl in this mixture was used to disrupt MunI dimer into monomers to get accessibility of Cys residues to the NBB (Figure 19D). After the reaction, mixture was dialysed to remove GdmCl and purification and analysis of reaction products was performed by gel filtration (Figure 24A). In contrast to wt MunI, which is a dimer in solution (~46kDa), the caged MunI mainly elutes as a monomer with a molecular mass of ~26kDa (Figure 24A, C). After caging only a small fraction (~10%) of MunI remains dimeric presumably due to incomplete chemical modification (Figure 24A). Altogether, gel filtration analysis showed that the caging efficiency of MunI with NBB is high, yielding ~90% of the caged monomer.

3.2.2.1.   Decaging of caged MunI

To reactivate caged MunI, it was irradiated with near UV light using different wavelengths (312-350 nm) for 1 hour at 4 °C. The decaging efficiency was determined by gel filtration calculating the fraction of the MunI dimer formed after irradiation (Figure 24B). Irradiation at 330 nm wavelength yielded the highest amount (~70%) of MunI dimer (Figure 25).
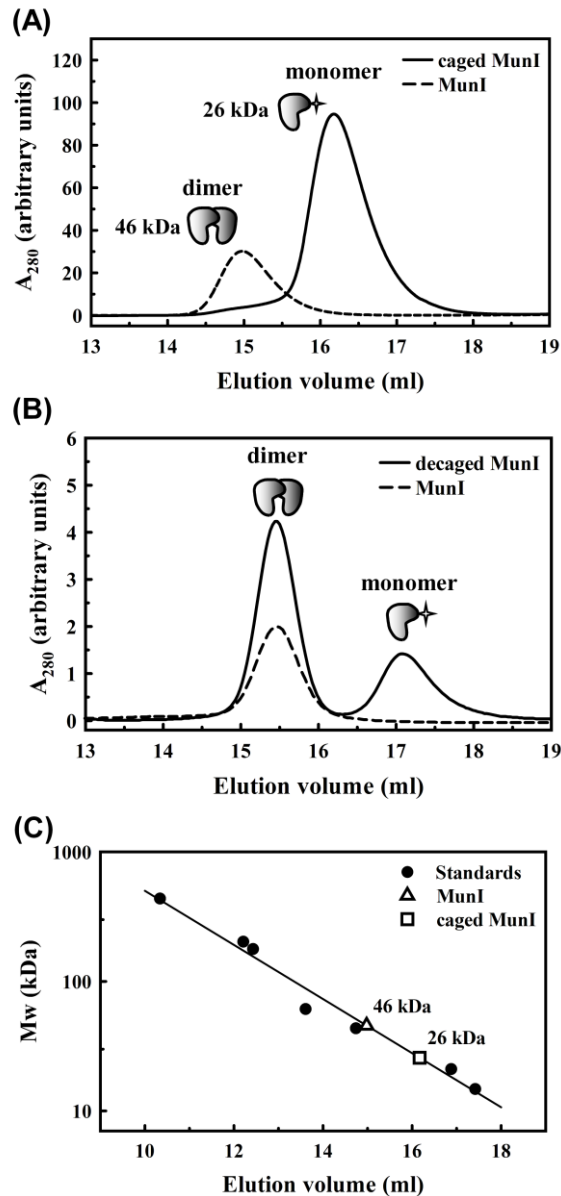

Figure 24. Caging of MunI with NBB (A) and decaging with UV light (B) monitored by gelfiltration. (A) Caging of MunI with NBB at the dimer interface generates a monomer in solution. Solid line indicates reaction products after caging of MunI; dashed line indicates the profile of wt MunI (B) Irradiation of caged monomeric MunI with UV light promotes dimer formation. Solid line indicates reaction products after decaging; dashed line indicates the profile of the wt MunI. (C) The apparent molecular mass values of wt and caged MunI were calculated by interpolation from the standard curve (closed circle) obtained by using a set of proteins of known molecular mass.
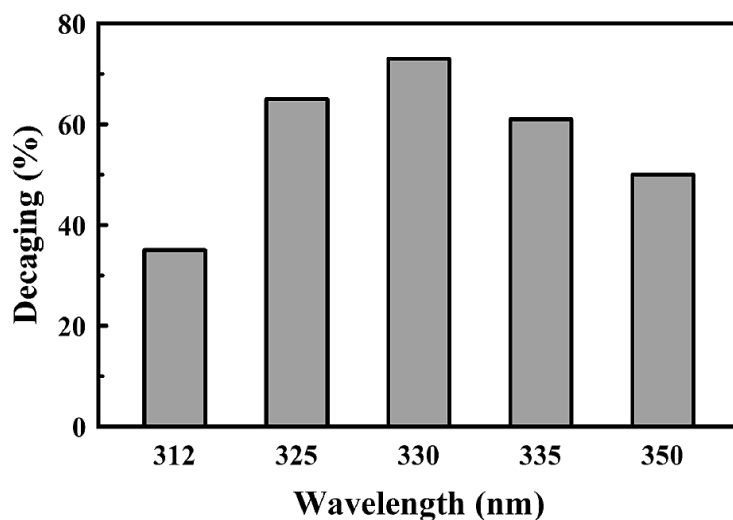
Figure 25. **Efficiency of MunI decaging at different wavelengths.** Decaging was performed at different wavelengths. Efficiency of decaging was monitored by gelfiltration calculating the fraction of resulted dimer.

3.2.2.2. λ DNA cleavage by caged and decaged MunI

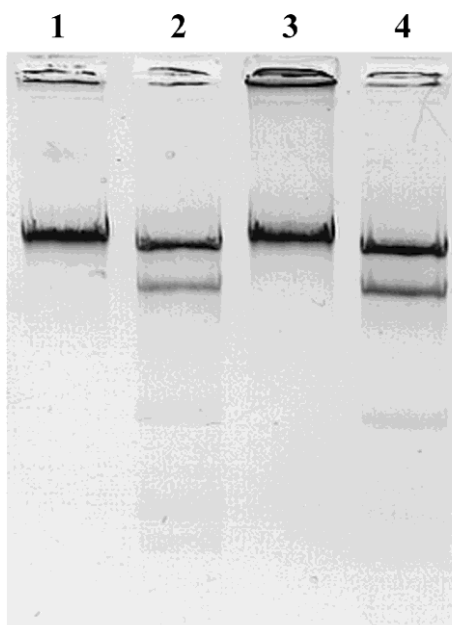To test DNA cleavage activity of the caged and decaged MunI the phage λ DNA was used as a substrate.

Figure 26. **Cleavage of λ DNA with caged and decaged MunI.** Lane 1, substrate DNA only; lane 2, cleavage with wt MunI; lane 3, cleavage with caged MunI; lane 4, cleavage with decaged MunI.

400 nM of wt, caged and decaged MunI was incubated with 1µg of λ DNA in a Reaction buffer I (10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM MgCl$_2$, 50 mM NaCl, 0.1 mg/ml BSA) for 15 min at 37ºC. Analysis of DNA cleavage products by agarose gel electrophoresis showed that the caged monomeric MunI exhibit no cleavage activity (Figure 26). However, after the irradiation with UV light the caged enzyme restores catalytic activity with the same specificity as the wt MunI (Figure 26).

3.2.2.3.   Plasmid DNA cleavage by caged and decaged MunI

To determine more precisely the catalytic activity of caged and decaged MunI, DNA cleavage experiments were performed using supercoiled plasmid DNA that contains a single copy of the MunI recognition sequence. 0.75 nM of wt MunI and 1.5 nM of caged and decaged MunI were incubated with 1.5 nM of plasmid DNA in Reaction buffer I (10 mM Tris-HCl (pH 7.5 at 25 °C), 10 mM MgCl$_2$, 50 mM NaCl, 0.1 mg/ml BSA) at 25 ºC. At equimolar concentrations of protein and DNA caged MunI showed no catalytic activity, not even nicking (Figure 27B). However, in the case of decaged MunI the profile of plasmid DNA cleavage was similar to that of wt MunI with the exception that a ~2-fold higher concentration of the decaged protein had to be used (Figure 27A, C). This result indicates that photoreactivation of MunI reaches ~50%.
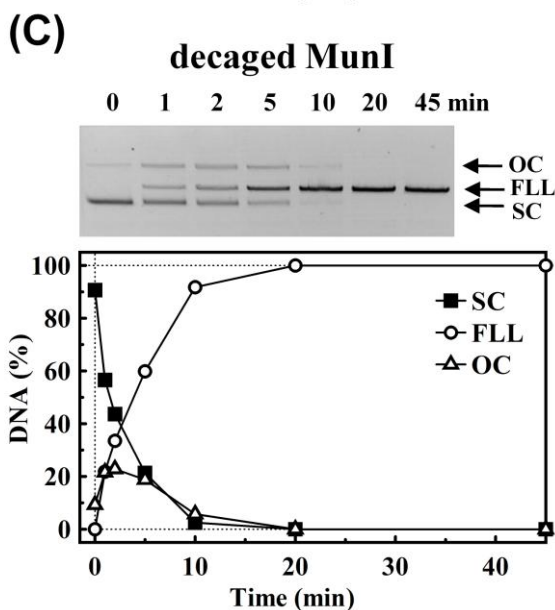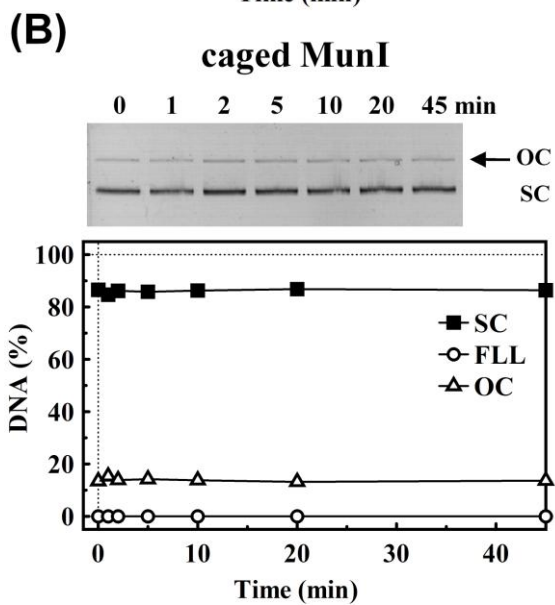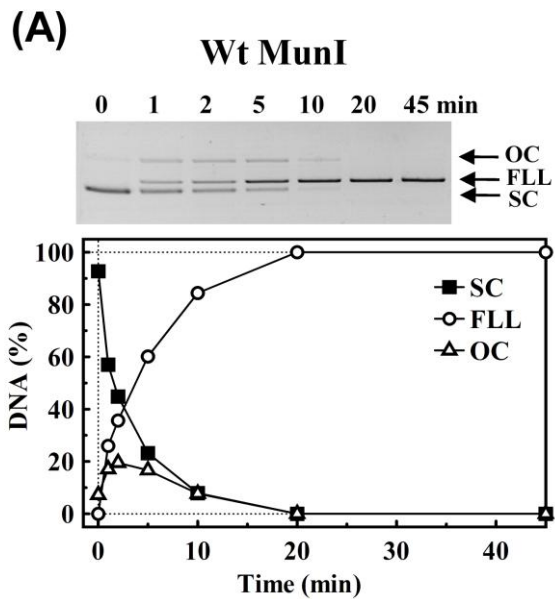
(A)

Wt MunI

(B)

caged MunI

(C)

decaged MunI

Figure 27. **Photoactivation of MunI.** (A), (B) and (C) represent cleavage of the plasmid DNA containing a single MunI recognition site by the wt, caged and decaged MunI, respectively. Kinetic experiments were performed by using 0.75 nM of the wt MunI and 1.5 nM of the caged and decaged MunI with 1.5 nM of plasmid DNA. Following forms of the DNA are observed during the reaction: supercoiled DNA (SC); opencircular DNA (OC); full-length linear DNA (FLL).

3.2.2.4. DNA binding with caged MunI

We performed gel mobility-shift assay experiments to determine DNA binding abilities of caged MunI to specific DNA.
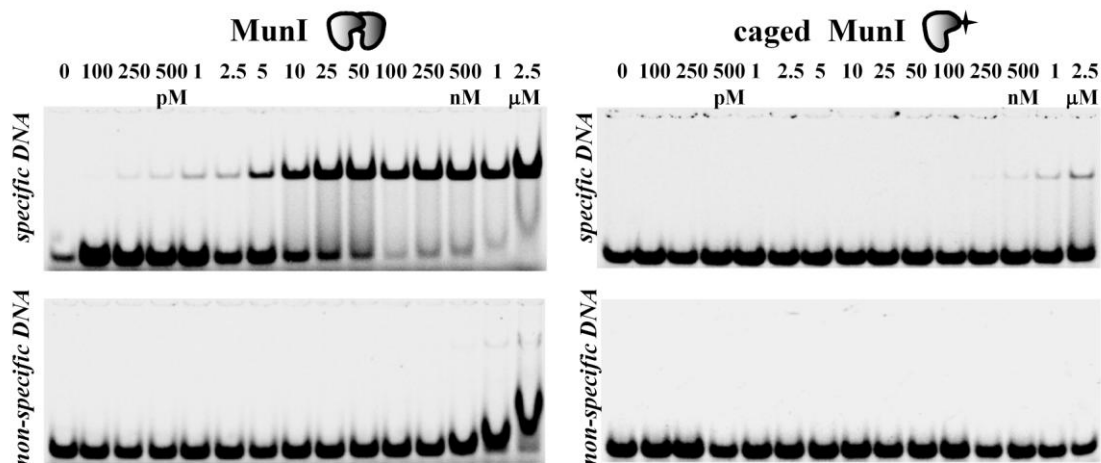

Figure 28. **Binding of wt and caged MunI to specific and non-specific DNA.** The labelled oligonucleotides with and without recognition sequence of MunI (at 2 nM final concentration) were incubated with increasing concentration of wt and caged MunI and analyzed by gel mobility-shift assay. After electrophoresis, the gels were dried and subjected to autoradiography and analysed using Cyclone Storage Phosphor System and OptiQuant Image Analysis Software, Version 03.00 (Packard Instrument Co).

For these experiments as substrate we used ^{33}P-labelled 22/19 oligoduplex, containing single recognition sequence of MunI and 22/22 oligoduplex that lack MunI recognition sequence. In order to monitor the formation of enzyme-DNA complexes, different amounts of protein were preincubated with 2 nM of cognate or non-cognate DNA, loaded on a polyacrylamide gel (PAAG) and electrophoresis was run. Gel shift analysis displayed that in contrast to wt MunI that binds cognate DNA with the $K_d \sim 10$ nM, cross-linked enzyme has very reduced binding to specific DNA with the $K_d > 2500$ nM (Figure 28).

3.2.2.5. MunI caging: summary and practical applications

Kinetic experiments together with gel filtration data show that the chemical modification of the Cys residues in MunI by NBB at the dimer interface (the yield about 90%) inactivates the enzyme preventing association of the caged monomers into the active dimer (Figure 24A and 27B). Gel shift experiments

in PAAG demonstrated that in contrast to the wt enzyme the caged monomeric MunI is unable to bind a cognate DNA (Figure 28). Activation of the caged enzyme was performed with near UV light irradiation resulting in removal of the cage compound and formation of the active MunI dimer. Decaging efficiency up to 70% and regeneration of cleavage activity up to 50% were achieved under our experimental conditions.

The strategies used in this work could be used to regulate activity of other enzymes. Native Cys residues or Cys residues introduced by mutagenesis at the oligomerization interface or at the substrate binding site could be selectively caged with NBB resulting in an inactive enzyme, whose catalytic activity could be restored by irradiation with near UV light. The procedures that we have developed are sufficiently general that they can be used for many other enzymes, for which it is desirable to turn them "on" by light.

3.3. Generation of the controllable restriction endonucleases

Targeting of individual genes in complex genomes requires endonucleases of extremely high specificity and restriction endonucleases are one of the potential classes of nucleases that could be used. Restriction enzymes exhibit strong specificity to cognate DNA and are relative small proteins therefore they could be easily delivered to cells. Example with scPvuII-TFO conjugate demonstrate the possibility to address restriction enzyme to defined DNA target site (Eisenschmidt et al, 2005), however, for application *in vivo* regulation of RE activity is required.
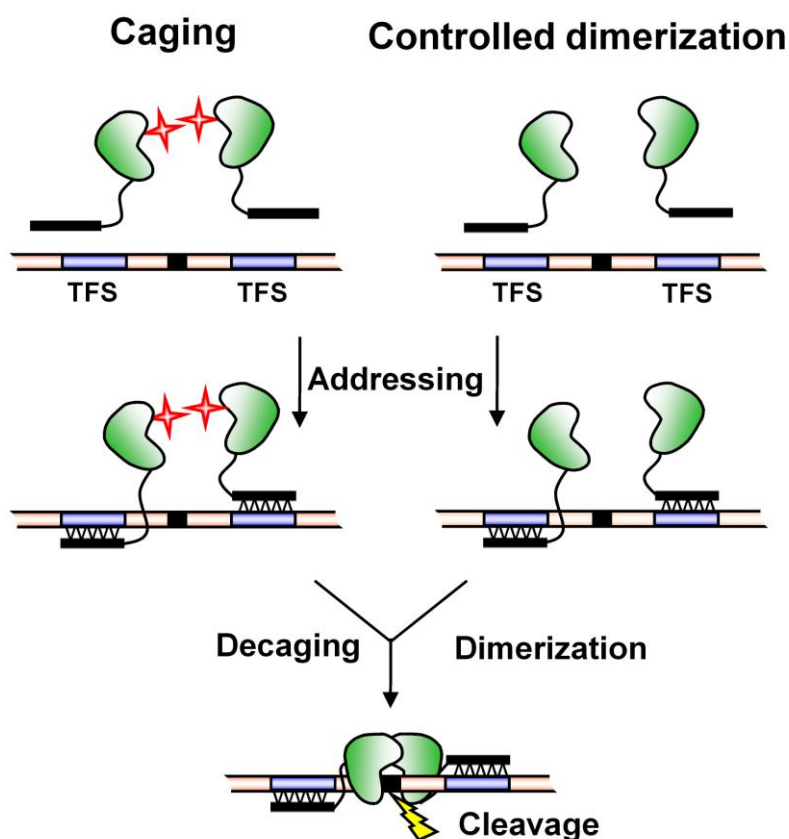
Figure 29. **Experimental strategy for generation of restriction enzyme-TFO conjugates with the adjustable activity.** The active restriction endonuclease homodimer can be inactivated either by caging (left panel, "caging") or by mutations at the dimerization interface (right panel, "controlled dimerization"). After triple helix is formed, the activity of caged restriction endonuclease would be restored upon decaging, while the enzyme with attenuated dimerization interface should form an active dimer due to increased local protein concentration.

In this work we used a triplex forming oligonucleotide (TFO) as a DNA binding mode that attached to the REase performs addressing of an enzyme to the defined locus on DNA thus increasing its specificity. Also we proposed two strategies to control the activity of the REase-TFO conjugates: (i) a caging strategy to introduce inactivating modifications that could be removed by an external signal (preferably by light) after formation of a stable triple helix (Figure 29, "caging"); (ii) a controlled oligomerization strategy to trigger dimerization of the catalytically inactive REase monomers brought to close proximity by triple helix formation at the target site (Figure 29, "controlled dimerization").

93

3.3.1.   Engineering a caged REase-TFO conjugate

To generate an addressable and photoswitchable REase-TFO conjugate, we selected MunI as a model system. MunI is well characterized both in terms of structure (Deibert et al, 1999) and function (Lagunavicius et al, 1997; Sasnauskas et al, 1999). It is a homodimeric protein that recognizes the palindromic DNA sequence 5′-C/AATTG-3′ and uses two symmetrically positioned active sites to cut both DNA strands in a Mg^{2+}-dependent reaction ("/" marks the cleavage position).

A native MunI dimer contains two Cys residues (one per monomer) at position 123. These Cys are buried at the dimerization interface and therefore are not available for TFO coupling. Instead, they are an excellent target for caging with a bulky photoremovable compound (see 3.2.2.). Disruption of MunI dimer in the presence of 1 M GdmCl followed by attachment of the nitrobenzyl (NB) moiety to Cys123 resulted in an inactive MunI monomer that could be reactivated by UV light. Coupling of TFO to MunI by standard thiol chemistry would require an extra solvent accessible Cys residue located distantly from the DNA binding cleft. To this end we engineered a hexahistidine tag followed by four glycines and a cysteine (Eisenschmidt et al, 2005) at the C-terminus of native MunI.

To selectively modify the subunit interface cysteine with the photoremovable compound and the C-terminal cysteine with a triplex forming oligonucleotide, we took advantage of different accessibility of these cysteine residues (Figure 30). First, a sample of His-tagged MunI was incubated with copper phenantroline (see 2.2.12.1.). This resulted in efficient formation (the yield close to 95% according to SDS-PAGE analysis, Figure 31A) of intersubunit disulfide bonds. Very low yield of cross-linked species (<5%, data not shown) in the case of native MunI lacking the C-terminal His_6Gly_4Cys sequence suggests that the disulfide bonds were formed almost exclusively between the C-terminal cysteines. This conclusion is also consistent with the available MunI structural data. Indeed, the length of the highly flexible C-

terminal His_6Gly_4C sequences should be sufficient to allow direct contacts between the terminal cysteines of different MunI subunits. Conversely, the dimerization interface cysteines are completely buried within the hydrophobic core of the protein and therefore are protected from disulfide bond formation (Deibert et al, 1999).
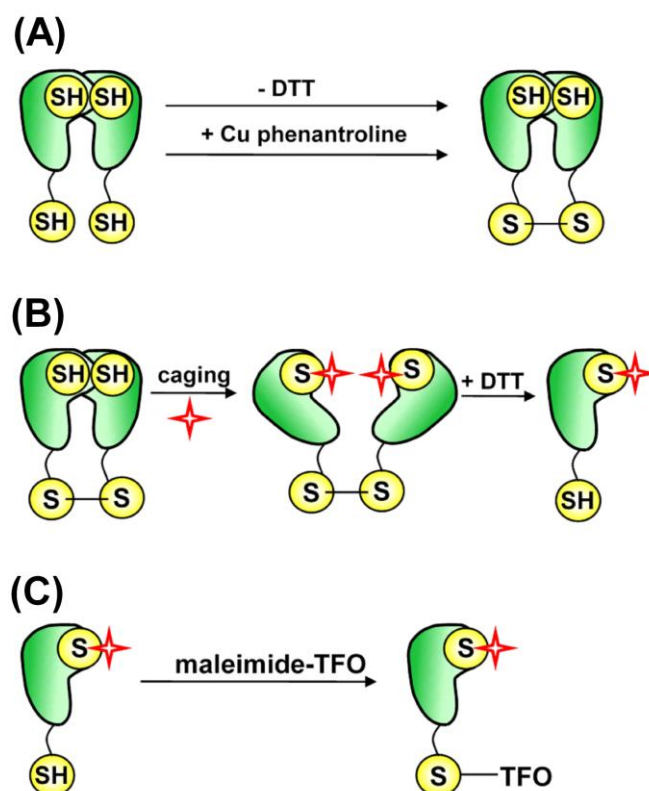

Figure 30. Generation of the caged MunI-TFO conjugate. In the first step, the C-terminal cysteines were protected from thiol-reactive reagents by forming intersubunit disulfide bonds (A). Next, the MunI dimerization interface residues Cys123 where modified by caging with 2-nitrobenzylbromide. After cleavage of intersubunit disulfide bond with DTT, a monomeric caged MunI with a free C-terminal cysteine was obtained (B) that was used for coupling with a triplex forming oligonucleotide *via* a maleimide group (C).

To modify the dimerization interface cysteines with the photoremovable 2-nitrobenzyl (NB) group, we followed the previously described procedure (see 2.2.9.). The Cys123 residues were exposed to the solvent by partial denaturation of MunI in 1 M GdmCl solution and covalently modified with 2-nitrobenzyl bromide (Figure 30B).
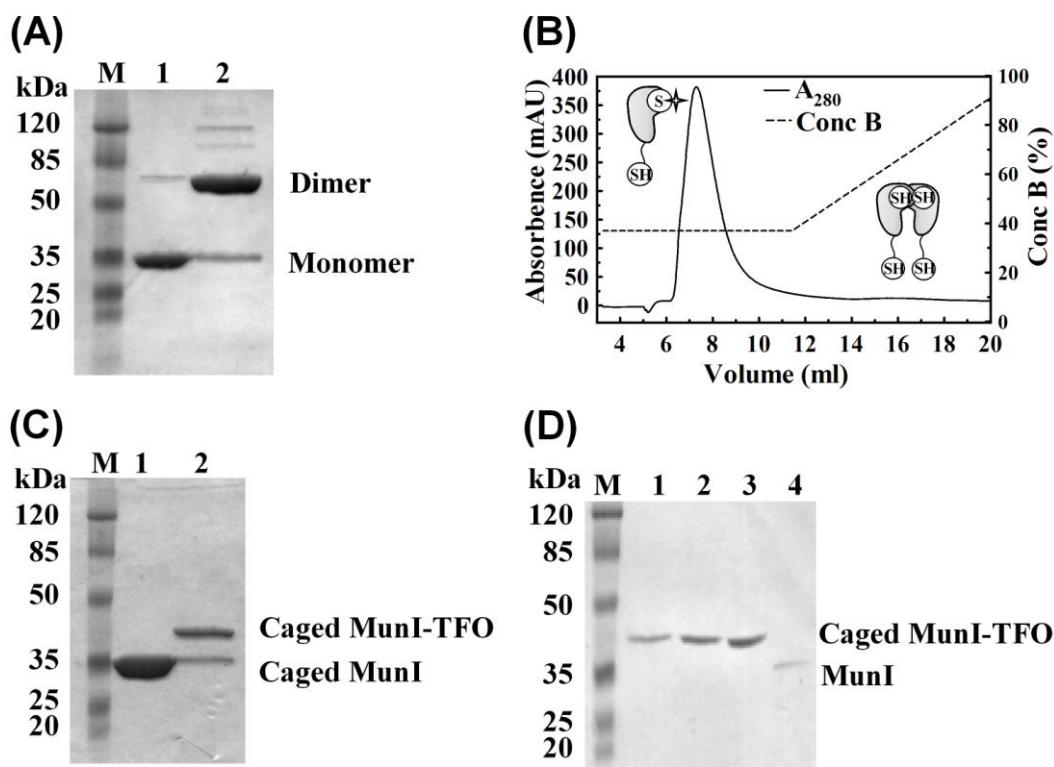
95

Figure 31. Generation of the caged MunI-TFO conjugate. (A) SDS-PAGE analysis of His-tagged MunI before (lane 1) and after (lane 2) intersubunit disulfide bond formation reaction. The yield of the dimer is ~95%. (B) Separation of the monomeric caged MunI from the uncaged homodimer. The caged monomer was eluted from the 1 ml HiTrap column (GE Healthcare, Uppsala, Sweden) using the Equilibration buffer IV containing 0.35 M NaCl and 0.2 mM tris(2-carboxyethyl)phosphine (TCEP). The uncaged homodimer (less than 5% of total protein) was eluted using a linear 0.35-1.0 M NaCl gradient. (C) SDS-PAGE analysis of caged MunI before (lane 1) and after (lane 2) coupling of triplex forming oligonucleotide. (D) Analysis of purified caged MunI-TFO conjugate. Lanes 1, 2 and 3 contain 1, 2 and 3 μL of the purified caged MunI-TFO conjugate solution, respectively; lane 4 contains 50 ng of His-tagged MunI.

During this step the C-terminal cysteines were protected from NBB modification due to the disulfide bond formation. Subsequent dialysis against buffer containing DTT removed the GdmCl with the caging reagent and restored the C-terminal cysteines. Ensuing chromatography on heparin sepharose in the pH 6.0 buffer enabled separation of the caged monomeric MunI bearing a single His tag from the unreacted dimer (see 2.2.12.1.). The yield of the caging reaction was close to 100% as judged from the ratio of monomeric and dimeric forms of MunI eluting from the Heparin sepharose

column (Figure 31B). Finally, the purified caged MunI was coupled to TFO by following the procedure described by Eisenschmidt et al (Eisenschmidt et al, 2005). Briefly, the TFO was first coupled with the bifunctional cross-linker GMBS (N-[γ-maleimidobutyryloxy]succinimide ester) *via* amide bond formation between the 5′-terminal C_{12}-amino group of the TFO and the succinimide group of GMBS, and then the TFO-GMBS adduct was coupled with the C-terminal cysteine of MunI *via* the maleimide group of GMBS to give the desired caged MunI-TFO product (Figure 30C and Figure 31C). MunI-TFO conjugate was separated from unreacted MunI and TFO by ion exchange chromatography (Figure 31D). In agreement with complete inactivation of MunI by caging (Figure 26 and Figure 27B), the caged MunI-TFO conjugate was catalytically inactive (data not shown).

3.3.1.1. DNA cleavage by the MunI-TFO conjugate

The MunI-TFO conjugate differs from the scPvuII-TFO adduct used in the pioneering study of Eisenschmidt and colleagues (Eisenschmidt et al, 2005) in the stoichiometry of the DNA-conjugate interaction. A single scPvuII-TFO molecule, containing two catalytic centers, is able to recognize and cleave the bipartite "addressed" DNA site consisting of the PvuII site and an adjacent triplex forming site. A caged MunI-TFO monomer contains only a single active site, therefore two enzyme-TFO conjugates must be guided to a single site in order to form an active homodimer capable of double-stranded DNA cleavage. This would require a tripartite recognition sequence consisting of the MunI recognition site embedded between two symmetrically positioned triplex forming sequences that match the 2-fold symmetry of the decaged homodimeric MunI-TFO conjugate.

Activity and specificity of the TFO-coupled MunI was tested on the equimolar mixture of 1880 and 3675 bp DNA fragments. The 1880 bp fragment contained the tripartite "addressed" site consisting of two triplex

forming sequences positioned 9 bp downstream and upstream of the MunI recognition site (Figure 32A).
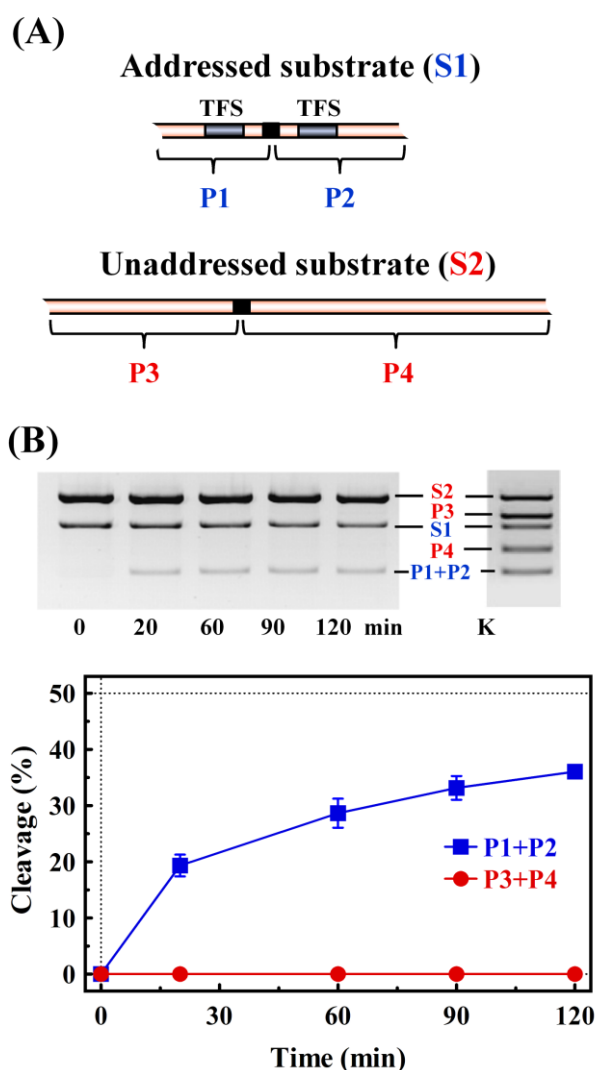
(A)

Addressed substrate (S1)

Unaddressed substrate (S2)

(B)

Figure 32. Addressed DNA cleavage by the caged MunI-TFO conjugate. (A) DNA substrates. The 1880 bp "addressed" substrate containing a MunI recognition site 5′-CAATTG-3′ embedded between two triplex forming sequences (TFSs), and the 3675 bp "unaddressed" substrate containing a MunI site without the adjacent TFS, were used in competition. (B) DNA cleavage. The equimolar mixture of "addressed" and "unaddressed" substrates was preincubated with caged MunI-TFO at 4 °C to allow triplex formation in a Triplex buffer I containing 80 mM Tris-phosphate (pH 6.1 at 25 °C), 1.0 M NaCl and 4 mM Mg^{2+}, 1mM DTT. After irradiation by UV light and dilution the cleavage kinetics was determined. Only the "addressed" substrate is cleaved with appreciable rate (P1+P2 products). The gel lane K contains the mixture of "addressed" and "unaddressed" substrates cleaved with His-tagged MunI lacking a TFO. The data points are presented as mean values from 3 independent experiments ± 1 standard deviation.

The 3675 bp fragment contained a MunI recognition site without the adjacent triplex forming sequences. Control experiments confirmed equal susceptibility

of MunI sites in both fragments to native MunI (Figure 32B, lane "K"). Incubation of the 1880 and 3675 bp fragments with the caged MunI-TFO conjugate in a Triplex buffer I, irradiation of the reaction mixture with UV light and further dilution in the Reaction buffer IV resulted in hydrolysis of the "addressed" MunI site in the 1880 bp fragment; no cleavage of the 3675 bp fragment could be detected even after prolonged incubation (Figure 32B). Assuming that ~30% of the addressed site was cleaved in less than 60 min, and no cleavage of the non-addressed substrate was detected after 3 h, the cleavage rates of the "addressed" and the "non-addressed" substrates must differ at least 100-fold (the estimate is based on the lower detection limit of band intensity analysis of ethidium bromide stained gels, 10 ng or ~1% of total DNA in a gel lane). In conclusion, we have successfully combined MunI caging with TFO coupling to generate a highly specific and photoswitchable DNA cleavage tool.

3.3.2. Engineering an addressable REase with an attenuated dimerization interface

The "addressed" DNA cleavage by a photoswitchable REase still requires a preincubation step to ensure formation of stable DNA triplexes that would guide all caged enzyme to the "addressed" site (Figure 29, "caging"). Otherwise, a free enzyme left in the reaction mixture upon decaging will become engaged in the non-addressed cleavage reactions directed solely by the REase specificity. This limitation can be overcome by the controlled oligomerization approach which makes use of the observation that subunit interactions of some orthodox REases can be attenuated by mutations (Fritsche & Alves, 2004; Zaremba et al, 2006). For example, the R226A mutation at the dimerization/tetramerization interface of restriction endonuclease Bse634I converts the native tetramer into inactive monomers (Zaremba et al, 2006) which are still able to form a functional dimer at increased protein concentrations. Fusion of the two REase monomers to different TFO should bring inactive monomers into close proximity at the Bse634I recognition site and promote the assembly of the catalytically active dimer. In contrast to the

FokI fusions, which employ a nonspecific nuclease domain as a cleavage module, a monomer of orthodox REase, which bears both catalytic and sequence recognition elements, is used for the fusion with TFO. The attenuated dimerization should completely prevent cleavage of non-addressed DNA sites at low enzyme concentrations, but cleavage of "addressed" sites would be turned on when two TFO-coupled monomers are guided to the "addressed" REase site by TFO-TFS interactions (Figure 30, "controlled dimerization").

3.3.2.1. Generation of Bse634I-TFO conjugate

To generate an addressable REase with an attenuated dimerization interface, we have chosen Bse634I REase as a model system. Wt Bse634I is a stable homotetrameric protein specific for the degenerate DNA sequence 5′-R/CCGGY-3′ (where R stands for A/G and Y for T/C) and cuts DNA after the first nucleotide (cleavage position indicated by "/"). Four subunits of the wt enzyme assemble into a dimer of two primary dimers positioned back-to-back to each other with DNA binding clefts facing into opposite directions (Grazulis et al, 2002). Two DNA sites are required for optimal Bse634I activity with only residual activity level observed on a single recognition site (Zaremba et al, 2005). It was demonstrated that alanine replacement of the R226 residue located at the dimer-dimer interface converts Bse634I into a monomer (Zaremba et al, 2006). DNA binding studies revealed that the R226A mutant preserved the ability to dimerize on cognate DNA forming a specific protein-DNA complex, albeit only at high enzyme concentrations (Zaremba et al, 2006). Dimerization of Bse634I R226A on DNA is also supported by the nonlinear dependence of DNA cleavage activity on enzyme concentration: contrary to the wt enzyme that rapidly hydrolyzes DNA even at low concentrations (Figure 33A), the mutant is almost inactive at 200 nM concentration (the DNA nicking rate is ~0.1 h^{-1}, no double-strand cleavage detectable after 8 h incubation), but regains detectable activity at 500 nM, and full catalytic activity at 5000 nM concentration (the rate of ds DNA cleavage

~4 min^{-1}). Moreover, distinctly from wt Bse634I, the R226A mutant is fully active on DNA substrates with a single recognition site (Figure 33B-D).
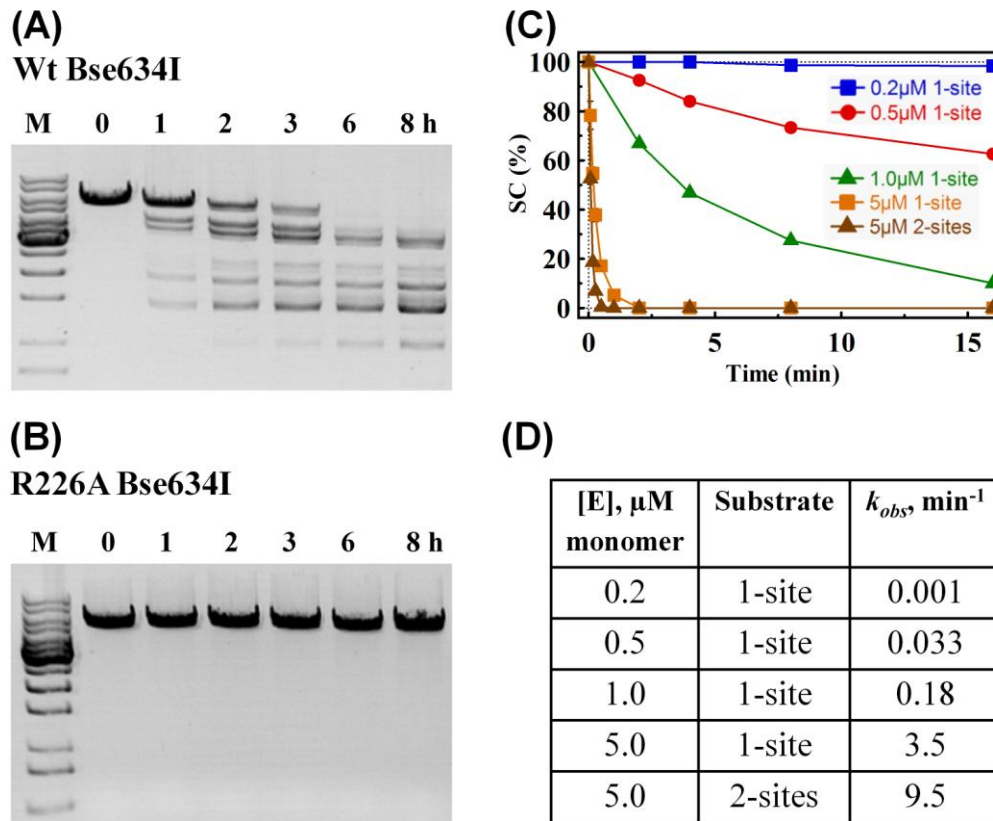

(A) Wt Bse634I

(B) R226A Bse634I

(C)

(D)

[E], μM monomer	Substrate	k_{obs}, min^{-1}
0.2	1-site	0.001
0.5	1-site	0.033
1.0	1-site	0.18
5.0	1-site	3.5
5.0	2-sites	9.5

Figure 33. Catalytic activity of wt Bse634I and the monomeric mutant R226A. (A-B) Cleavage of a linear 7-site 5.5 kbp DNA substrate by wt Bse634I and the R226A mutant. 50 nM of DNA were incubated with 200 nM of the tetrameric wt Bse634I (A) or 200 nM of the monomeric R226A mutant (B) for 0-8 hours at 37 °C in 80 mM Tris-phosphate (pH 6.1 at 37 °C), 2 mM spermine, 2 mM MgCl$_2$ and 10 mM DTT. Reactions were quenched with loading dye solution and analyzed by agarose electrophoresis. (C-D) DNA nicking by the monomeric Bse634I R226A mutant at various enzyme concentrations. The reactions were performed at 25 °C on the supercoiled plasmid substrates pUCAT1 and pUCAT2 bearing a single or two 5′-ACCGGT-3′ recognition sites in the optimal Bse634I Reaction buffer II (10 mM Tris-HCl (pH 8.5), 100 mM KCl, 10 mM MgCl$_2$, and 0.1 mg/ml BSA). Initiation of the reactions, sample collection and data analysis was performed as described by Zaremba et al. (Zaremba et al, 2005). The cleavage profiles of the supercoiled substrate at various enzyme concentrations (in terms of monomer) are shown in (C); the determined nicking rate constants are summarized in (D). The higher rate of the 2-site plasmid cleavage results from simultaneous action of two Bse634I molecules bound to the two sites.
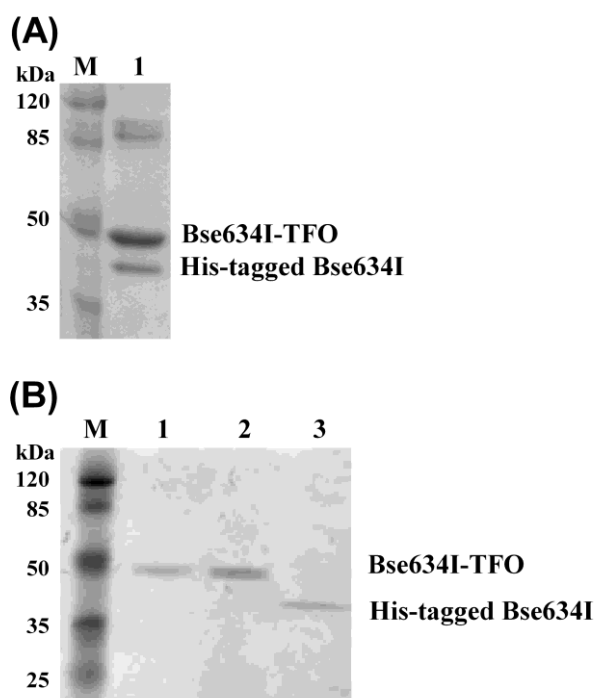
Figure 34. Generation of the Bse634I-TFO conjugate. (A) SDS-PAGE analysis of His-tagged Bse634I sample after the TFO coupling reaction. (B) SDS-PAGE analysis of purified Bse634I-TFO conjugate. Lanes 1 and 2 contain 1 and 2 µl of the purified Bse634I-TFO conjugate solution, respectively; lane 3 contains 50 ng of His-tagged Bse634I.

Wt Bse634I monomer contains two Cys residues at positions 10 and 186. In order to limit TFO coupling to the Bse634I R226A mutant at a single position, we replaced the native Bse634I cysteines by Ser (mutations C10S and C186S) and engineered the His_6Gly_4Cys sequence containing a unique cysteine at the C-terminus of the protein. The resulting His-tagged variant of Bse634I R226A+C10S+C186S mutant (His-tagged Bse634I) displayed the same nonlinear dependence of enzymatic activity on the protein concentration as the untagged Bse634I R226A. The Bse634I-TFO conjugates were prepared using two 15 nt triplex forming oligonucleotides (TFO1 and TFO2) containing 5-methyl-2′-deoxycytidine (for deoxycytidine) and 5-[1-propynyl]-2-deoxyuridine (for thymidine) modifications that increase DNA triplex stability at near-neutral pH. Similarly to MunI, the TFOs were coupled to the C-terminal cysteine of His-tagged Bse634I *via* the bifunctional cross-linker GMBS. Ion exchange chromatography enabled separation of the Bse634I-TFO

conjugates from the unreacted protein and triplex forming oligonucleotides (see Figure 34 and "methods" section 2.2.12.1. for details).

3.3.2.2. Activity and specificity of the Bse634I-TFO conjugates

First, was compared the activity of the His-tagged Bse634I and the Bse634I-TFO conjugates on a supercoiled plasmid substrate pUCAT1 (2005 Zaremba) bearing a single Bse634I recognition site with no adjacent triplex forming sequences.
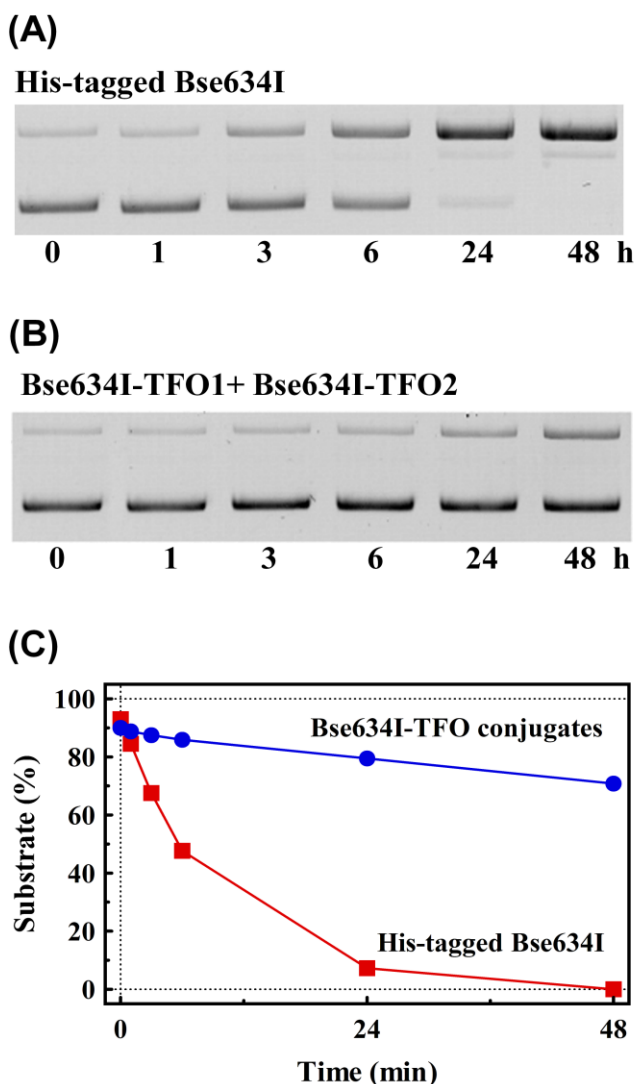
(A)

His-tagged Bse634I

0   1   3   6   24   48 h

(B)

Bse634I-TFO1+ Bse634I-TFO2

0   1   3   6   24   48 h

(C)

Figure 35. Activity of His-tagged Bse634I and Bse634I-TFO conjugates on the one-site plasmid pUCAT1. (A) Cleavage of supercoiled pUCAT1 DNA by His-tagged Bse634I. The reaction contained 50 nM plasmid DNA and 200 nM His-tagged Bse634I monomer in the Reaction buffer III (80 mM Tris-phosphate (pH 7.1 at 25 °C), 2 mM spermine, 2 mM $MgCl_2$ and 10 mM DTT). Reactions were stopped with loading dye solution [50 mM EDTA (pH 8.0 at 25 °C), 0.1% SDS, 50% (v/v) glycerol, 0.01% bromphenol blue] and analyzed by agarose electrophoresis. (B) Cleavage of supercoiled pUCAT1 DNA by the equimolar mixture of Bse634I-TFO1 and Bse634I-TFO2 conjugates. The reaction contained 50 nM plasmid DNA and 200 nM enzyme-TFO conjugates in the Reaction buffer III mentioned above. The samples were analyzed as in panel (A). (C) Cleavage of supercoiled DNA by the His-tagged Bse634I (red filled squares) and Bse634I-TFO conjugates (blue filled circles). The solid lines are single-exponential fits to the experimental data derived from panels (A) and (B). The determined rate constant is 0.11 h-1 for the His-tagged Bse634I and 0.005 h-1 for the Bse634I-TFO conjugates.
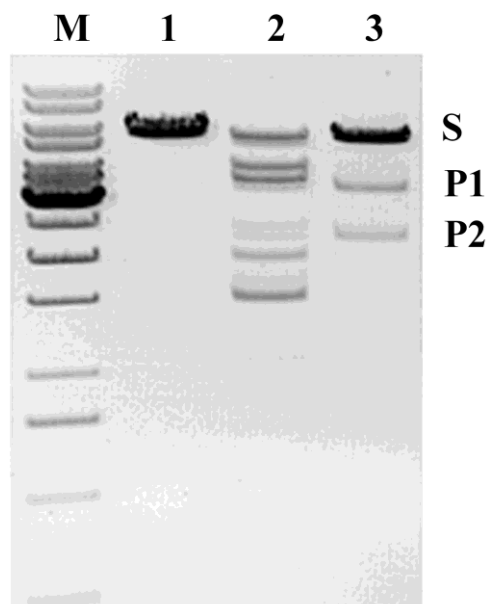
103

Figure 36. DNA cleavage by the His-tagged Bse634I and the Bse634I-TFO conjugates. 50 nM of the 5.5 kbp linear DNA fragment containing a single addressed Bse634I site and 6 non-addressed sites was incubated with 2000 nM of His-tagged monomeric Bse634I (lane "2") or 200 nM of Bse634I-TFO1 and Bse634I-TFO2 conjugates (lane "3") at 37 °C in 80 mM Tris-phosphate (pH 6.1 at 37 °C), 2 mM spermine, 2 mM MgCl$_2$ and 10 mM DTT for 2 hours. Lane "1" contains uncleaved substrate, lane "M" contains "GeneRuler 1 kb DNA Ladder" ("Thermo Fisher Scientific").

TFO conjugation to the C-terminus of Bse634I reduced the DNA nicking rate approximately 20-fold (to ~0.005 h^{-1} with 200 nM of Bse634I-TFO conjugate, Figure 35). Presumably, the repulsion between the negatively charged TFO and the DNA duplex or between two TFOs impedes association of Bse634I monomers with the recognition site.

To assess the ability of Bse634I-TFO conjugates to specifically recognize and hydrolyze the addressed DNA sites, we employed a 5.5 kbp linear DNA fragment containing 7 Bse634I recognition sites. One of these sites was flanked by two different 15 bp pyrimidine tracts (Figure 37A) that form triple helixes with TFO1 and TFO2. A control experiment with 2 μM of His-tagged Bse634I confirmed that the enzyme without the attached TFO has no preference for the "addressed" site (Figure 36).
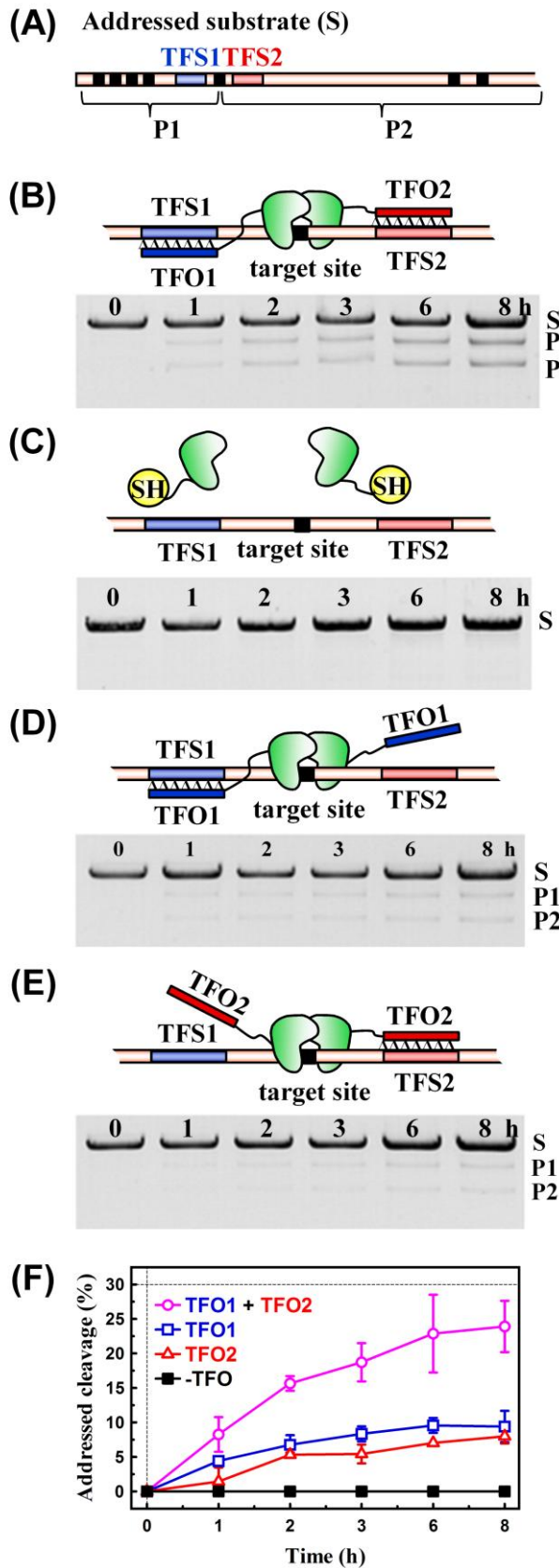
Figure 37. Addressed DNA cleavage by the Bse634I-TFO conjugates. The 5.5 kbp linear DNA substrate, containing one "addressed" Bse634I site [having triplex forming sequences TFS1 (blue box) and TFS2 (red box) 9 bp away from the Bse634I site 5′-ACCGGT-3′] and six "unaddressed" sites (A) was incubated with equimolar mixture of Bse634I-TFO1 and Bse634I-TFO2 conjugates at 37 °C (B). Only the "addressed" cleavage products (the 2298 bp and 3257 bp fragments P1 and P2) are formed during 1-8 h incubation. Control experiments were performed with His-tagged Bse634I lacking a TFO (C) and individual Bse634I-TFO1 or Bse634I-TFO2 conjugates (D-E). The extent of DNA cleavage in different experiments is summarized in panel (F). The data points are presented as mean values from 3 (TFO1+TFO2) or 2 (TFO1, TFO2) independent experiments ± 1 standard deviation.

105

Incubation of the substrate with equimolar mixture of Bse634I-TFO1 and Bse634I-TFO2 conjugates (total conjugate concentration 200 nM) in a Triplex buffer II (80 mM Tris-phosphate (pH 6.1 at 37 °C), 2 mM spermine, 2 mM MgCl$_2$ and 10 mM DTT) that is compatible with both DNA triplex formation and DNA hydrolysis resulted in relatively slow appearance of the 3257 and 2298 bp DNA fragments that correspond to cleavage of the addressed Bse634I site (Figure 37B). Approximately 25% of the substrate was cleaved during 8 h incubation time without any detectable non-addressed cleavage products (Figure 37F). Assuming that the 5.5 kbp substrate contains a 6-fold excess of non-addressed Bse634I sites over the addressed site, and taking into account the detection limit of the ethidium bromide stained DNA bands (~10 ng or ~1% of total DNA in an agarose gel lane), the cleavage rates of the "addressed" and the "non-addressed" Bse634I sites must differ at least 100-fold.

To verify that both TFO1-TFS1 and TFO2-TFS2 interactions are required for the "addressed" cleavage reaction, we performed control experiments using (i) the His-tagged Bse634I lacking a conjugated TFO; (ii) only the Bse634I-TFO1 conjugate, and (iii) only the Bse634I-TFO2 conjugate. Experiments with His-tagged Bse634I lacking a conjugated TFO yielded no double-strand breaks during 8 h incubation (Figure 37C), confirming that TFO-TFS interactions are obligatory for the assembly of Bse634I monomers on DNA. Individual Bse634I-TFO cross-links generated detectable amounts of the addressed cleavage products (Figure 37D, E), but cleavage extent was 2.5-fold (Bse634I-TFO1) and 3-fold (Bse634I-TFO2) lower than in the experiment with the equimolar mixture of both Bse634I-TFO1 and Bse634I-TFO2 conjugates. This suggests that under our reaction conditions (200 nM Bse634I-TFO conjugate) the catalytically competent Bse634I dimers assemble on the "addressed" site by two pathways. In the optimal scenario, both Bse634I monomers are guided to the "addressed" site by the TFO-TFS interactions, resulting in the highest extent of DNA cleavage. In a less favorable scenario, a single Bse634I

monomer guided to the addressed site by triplex formation captures the second Bse634I monomer from free solution.

CONCLUSIONS

1. Subunit exchange rate of the MunI restriction endonuclease dimer has been evaluated and experimental conditions that stabilize the MunI dimer and prevent subunit exchange have been established.

2. GdmCl-induced denaturation pathway of the MunI restriction endonuclease has been established. MunI denaturation induced by GdmCl occurs in two separate stages: first MunI dimer dissociates into the monomers, followed by the monomer unfolding.

3. Disulfide bond introduced at the dimer interface of MunI acts as a red-ox switch which controls cleavage activity. MunI shows no activity in the cross-linked state but addition of the reducing agent fully restores the catalytic activity.

4. Introducing of the photoremovable cage compound at the MunI interface disrupts the MunI dimer into an inactive monomer. Irradiation with UV light removes cage compound and restores the catalytic activity of MunI.

5. Restriction endonuclease Bse634I-TFO conjugates with controllable catalytic activity were engineered for addressed DNA cleavage.

6. Restriction enzyme-TFO conjugates with spatial and temporal activity control provide an experimental platform for engineering novel tools for gene targeting.

LIST OF PUBLICATIONS

The thesis is based on the following original publications:

1. **Silanskas, A.**, Zaremba, M., Sasnauskas, G., Siksnys, V. (2012) Catalytic activity control of restriction endonuclease--triplex forming oligonucleotide conjugates. *Bioconjug Chem* **23**(2)**:** 203-211.
2. **Silanskas, A.**, Foss, M., Wende, W., Urbanke, C., Lagunavicius, A., Pingoud, A., Siksnys, V. (2011) Photocaged variants of the MunI and PvuII restriction enzymes. *Biochemistry* **50**(14)**:** 2800-2807.

Other publication:

3. Geel, T.M., Meiss, G., van der Gun, B.T., Kroesen, B.J., de Leij, L.F., Zaremba, M., **Silanskas, A.**, Kokkinidis, M., Pingoud, A., Ruiters, M.H., McLaughlin, P.M., Rots, M.G. (2009) Endonucleases induced TRAIL-insensitive apoptosis in ovarian carcinoma cells. *Exp Cell Res* **315:** 2487-2495.

Patent application:

4. Siksnys, V., Zaremba, M., Sasnauskas, G., **Silanskas, A.** (2011) Generation of restriction endonuclease triplex forming oligonucleotide conjugates with switchable activity. USA provisional patent application Nr.: 61/522,310.

CONFERENCE PRESENTATIONS

1. **Silanskas, A.**, Urbanke, C., Siksnys, V. Type II restriction endonuclease under light control. 6th New England Biolabs Meeting on Restriction/Modification, Jacobs University, Bremen, Germany; 2010.08.01-06.

2. **Silanskas, A.**, Lagunavicius, A., Urbanke, C., Siksnys, V. Reversible redox switch at the dimer interface of MunI restriction enzyme. 5th New England Biolabs Meeting on Restriction/Modification; University of Bristol, Bristol, UK, 2004.09.04-08.

FINANCIAL SUPPORT

ACKNOWLEDGEMENT

Mališauskas for multiple encouragements to strive for faster doctoral dissertation defence.

My greatest thanks to my parents, my brother and parents-in-law for their concern and understanding.

My deepest and warmest thanks go to my wife Marija for her attention and support sharing those happy and sad moments over these years, and to my children Ugnė and Ignas whose smile and laugh motivate me to reach the aims.

REFERENCES

1. Accurso FJ (2006) Update in cystic fibrosis 2005. *Am J Respir Crit Care Med* **173:** 944-947

2. Aihara H, Miyazaki J (1998) Gene transfer into muscle by electroporation in vivo. *Nat Biotechnol* **16:** 867-870

3. Alwin S, Gere MB, Guhl E, Effertz K, Barbas CF, 3rd, Segal DJ, Weitzman MD, Cathomen T (2005) Custom zinc-finger nucleases for use in human cells. *Mol Ther* **12:** 610-617

4. Ardehali A, Fyfe A, Laks H, Drinkwater DC, Jr., Qiao JH, Lusis AJ (1995) Direct gene transfer into donor hearts at the time of harvest. *J Thorac Cardiovasc Surg* **109:** 716-719; discussion 719-720

5. Arimondo PB, Thomas CJ, Oussedik K, Baldeyrou B, Mahieu C, Halby L, Guianvarc'h D, Lansiaux A, Hecht SM, Bailly C, Giovannangeli C (2006) Exploring the cellular activity of camptothecin-triple-helix-forming oligonucleotide conjugates. *Mol Cell Biol* **26:** 324-333

6. Arnould S, Chames P, Perez C, Lacroix E, Duclert A, Epinat JC, Stricher F, Petit AS, Patin A, Guillier S, Rolland S, Prieto J, Blanco FJ, Bravo J, Montoya G, Serrano L, Duchateau P, Paques F (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J Mol Biol* **355:** 443-458

7. Arnould S, Perez C, Cabaniols JP, Smith J, Gouble A, Grizot S, Epinat JC, Duclert A, Duchateau P, Paques F (2007) Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J Mol Biol* **371:** 49-65

8. Asensio JL, Brown T, Lane AN (1999) Solution conformation of a parallel DNA triple helix with 5' and 3' triplex-duplex junctions. *Structure* **7:** 1-11

9. Barry JK, Matthews KS (1999) Thermodynamic analysis of unfolding and dissociation in lactose repressor protein. *Biochemistry* **38:** 6520-6528

10. Beerli RR, Schopfer U, Dreier B, Barbas CF, 3rd (2000) Chemically regulated zinc finger transcription factors. *J Biol Chem* **275:** 32617-32627

11. Belfort M, Perlman PS (1995) Mechanisms of intron mobility. *J Biol Chem* **270:** 30237-30240

12. Belfort M, Roberts RJ (1997) Homing endonucleases: keeping the house in order. *Nucleic Acids Res* **25:** 3379-3388

13. Berg JM (1988) Proposed structure for the zinc-binding domains from transcription factor IIIA and related proteins. *Proc Natl Acad Sci U S A* **85:** 99-102

14. Beumer K, Bhattacharyya G, Bibikova M, Trautman JK, Carroll D (2006) Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics* **172:** 2391-2403

15. Bibikova M, Beumer K, Trautman JK, Carroll D (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science* **300:** 764

16. 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

17. Bibikova M, Golic M, Golic KG, Carroll D (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics* **161:** 1169-1175

18. Bigey P, Pratviel G, Meunier B (1995) Cleavage of double-stranded DNA by 'metalloporphyrin-linker-oligonucleotide' molecules: influence of the linker. *Nucleic Acids Res* **23:** 3894-3900

19. Birg F, Praseuth D, Zerial A, Thuong NT, Asseline U, Le Doan T, Helene C (1990) Inhibition of simian virus 40 DNA replication in CV-1 cells by an oligodeoxynucleotide covalently linked to an intercalating agent. *Nucleic Acids Res* **18:** 2901-2908

20. Bitinaite J, Wah DA, Aggarwal AK, Schildkraut I (1998) FokI dimerization is required for DNA cleavage. *Proc Natl Acad Sci U S A* **95:** 10570-10575

21. Blume SW, Lebowitz J, Zacharias W, Guarcello V, Mayfield CA, Ebbinghaus SW, Bates P, Jones DE, Jr., Trent J, Vigneswaran N, Miller DM (1999) The integral divalent cation within the intermolecular

purine*purine. pyrimidine structure: a variable determinant of the potential for and characteristics of the triple helical association. *Nucleic Acids Res* **27:** 695-702

22. Boch J, Bonas U (2010) Xanthomonas AvrBs3 family-type III effectors: discovery and function. *Annu Rev Phytopathol* **48:** 419-436

23. 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

24. Bolduc JM, Spiegel PC, Chatterjee P, Brady KL, Downing ME, Caprara MG, Waring RB, Stoddard BL (2003) Structural and biochemical analyses of DNA and RNA binding by a bifunctional homing endonuclease and group I intron splicing factor. *Genes Dev* **17:** 2875-2888

25. Bolon DN, Grant RA, Baker TA, Sauer RT (2005) Specificity versus stability in computational protein design. *Proc Natl Acad Sci U S A* **102:** 12724-12729

26. Bolotin M, Coen D, Deutsch J, Dujon B, Netter P, Petrochilo E, Slonimski PP (1971) La recombinaison des mitochondries chez Saccharomyces
27. cerevisiae. *Bull Inst Pasteur* **69:** 215–239

28. Bos JL, Heyting C, Borst P, Arnberg AC, Van Bruggen EF (1978) An insert in the single gene for the large ribosomal RNA in yeast mitochondrial DNA. *Nature* **275:** 336-338

29. Brunet E, Alberti P, Perrouault L, Babu R, Wengel J, Giovannangeli C (2005) Exploring cellular activity of locked nucleic acid-modified triplex-forming oligonucleotides and defining its molecular basis. *J Biol Chem* **280:** 20076-20085

30. Carr MJ, Kajon AE, Lu X, Dunford L, O'Reilly P, Holder P, De Gascun CF, Coughlan S, Connell J, Erdman DD, Hall WW (2011) Deaths associated with human adenovirus-14p1 infections, Europe, 2009-2010. *Emerg Infect Dis* **17:** 1402-1408

31. Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ, Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39:** e82

32. Chames P, Epinat JC, Guillier S, Patin A, Lacroix E, Paques F (2005) In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination. *Nucleic Acids Res* **33:** e178

33. Chandrasegaran S, Smith J (1999) Chimeric restriction enzymes: what is next? *Biol Chem* **380:** 841-848

34. Chevalier B, Sussman D, Otis C, Noel AJ, Turmel M, Lemieux C, Stephens K, Monnat RJ, Jr., Stoddard BL (2004) Metal-dependent DNA cleavage mechanism of the I-CreI LAGLIDADG homing endonuclease. *Biochemistry* **43:** 14015-14026

35. Chevalier B, Turmel M, Lemieux C, Monnat RJ, Jr., Stoddard BL (2003) Flexible DNA target site recognition by divergent homing endonuclease isoschizomers I-CreI and I-MsoI. *J Mol Biol* **329:** 253-269

36. Chevalier BS, Kortemme T, Chadsey MS, Baker D, Monnat RJ, Stoddard BL (2002) Design, activity, and structure of a highly specific artificial endonuclease. *Mol Cell* **10:** 895-905

37. Chevalier BS, Stoddard BL (2001) Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility. *Nucleic Acids Res* **29:** 3757-3774

38. Choo Y, Klug A (1994a) Selection of DNA binding sites for zinc fingers using rationally randomized DNA reveals coded interactions. *Proc Natl Acad Sci U S A* **91:** 11168-11172

39. Choo Y, Klug A (1994b) Toward a code for the interactions of zinc fingers with DNA: selection of randomized fingers displayed on phage. *Proc Natl Acad Sci U S A* **91:** 11163-11167

40. Choo Y, Sanchez-Garcia I, Klug A (1994) In vivo repression by a site-specific DNA-binding protein designed against an oncogenic sequence. *Nature* **372:** 642-645

41. Chou C, Deiters A (2011) Light-activated gene editing with a photocaged zinc-finger nuclease. *Angew Chem Int Ed Engl* **50:** 6839-6842

42. Chou C, Young DD, Deiters A (2009) A light-activated DNA polymerase. *Angew Chem Int Ed Engl* **48:** 5950-5953

43. Christian M, Cermak T, Doyle EL, Schmidt C, Zhang F, Hummel A, Bogdanove AJ, Voytas DF (2010) Targeting DNA double-strand breaks with TAL effector nucleases. *Genetics* **186:** 757-761

44. Christy BA, Lau LF, Nathans D (1988) A gene activated in mouse 3T3 cells by serum growth factors encodes a protein with "zinc finger" sequences. *Proc Natl Acad Sci U S A* **85:** 7857-7861

45. Collins FS, Lander ES, Rogers J, Waterson RH (2004) Finishing the euchromatic sequence of the human genome. *Nature* **431:** 931-945

46. Cornu TI, Thibodeau-Beganny S, Guhl E, Alwin S, Eichtinger M, Joung JK, Cathomen T (2008) DNA-binding specificity is a major determinant of the activity and toxicity of zinc-finger nucleases. *Mol Ther* **16:** 352-358

47. Dalgaard JZ, Klar AJ, Moser MJ, Holley WR, Chatterjee A, Mian IS (1997) Statistical modeling and analysis of the LAGLIDADG family of site-specific endonucleases and identification of an intein that encodes a site-specific endonuclease of the HNH family. *Nucleic Acids Res* **25:** 4626-4638

48. Deibert M, Grazulis S, Janulaitis A, Siksnys V, Huber R (1999) Crystal structure of MunI restriction endonuclease in complex with cognate DNA at 1.7 A resolution. *Embo J* **18:** 5805-5816

49. Deng D, Yan C, Pan X, Mahfouz M, Wang J, Zhu JK, Shi Y, Yan N (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science* **335:** 720-723

50. Dervan PB (1986) Design of sequence-specific DNA-binding molecules. *Science* **232:** 464-471

51. Doyon JB, Pattanayak V, Meyer CB, Liu DR (2006) Directed evolution and substrate specificity profile of homing endonuclease I-SceI. *J Am Chem Soc* **128:** 2477-2484

52. Duca M, Vekhoff P, Oussedik K, Halby L, Arimondo PB (2008) The triple helix: 50 years later, the outcome. *Nucleic Acids Res* **36:** 5123-5138

53. Dujon B (1989) Group I introns as mobile genetic elements: facts and mechanistic speculations--a review. *Gene* **82:** 91-114

54. Dupureur CM, Dominguez MA, Jr. (2001) The PD...(D/E)XK motif in restriction enzymes: a link between function and conformation. *Biochemistry* **40:** 387-394

55. Eisenschmidt K, Lanio T, Simoncsits A, Jeltsch A, Pingoud V, Wende W, Pingoud A (2005) Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Res* **33:** 7039-7047

56. Ellman GL (1959) Tissue sulfhydryl groups. *Arch Biochem Biophys* **82:** 70-77

57. Elrod-Erickson M, Pabo CO (1999) Binding studies with mutants of Zif268. Contribution of individual side chains to binding affinity and specificity in the Zif268 zinc finger-DNA complex. *J Biol Chem* **274:** 19281-19285

58. Elrod-Erickson M, Rould MA, Nekludova L, Pabo CO (1996) Zif268 protein-DNA complex refined at 1.6 A: a model system for understanding zinc finger-DNA interactions. *Structure* **4:** 1171-1180

59. Endo M, Nakayama K, Majima T (2004) Design and synthesis of photochemically controllable restriction endonuclease BamHI by manipulating the salt-bridge network in the dimer interface. *J Org Chem* **69:** 4292-4298

60. Engler C, Gruetzner R, Kandzia R, Marillonnet S (2009) Golden gate shuffling: a one-pot DNA shuffling method based on type IIs restriction enzymes. *PLoS One* **4:** e5553

61. Engler C, Kandzia R, Marillonnet S (2008) A one pot, one step, precision cloning method with high throughput capability. *PLoS One* **3:** e3647

62. Epinat JC, Arnould S, Chames P, Rochaix P, Desfontaines D, Puzin C, Patin A, Zanghellini A, Paques F, Lacroix E (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res* **31:** 2952-2962

63. Escude C, Nguyen CH, Kukreti S, Janin Y, Sun JS, Bisagni E, Garestier T, Helene C (1998) Rational design of a triple helix-specific intercalating ligand. *Proc Natl Acad Sci U S A* **95:** 3591-3596

64. Escude C, Sun JS, Rougee M, Garestier T, Helene C (1992) Stable triple helices are formed upon binding of RNA oligonucleotides and their 2'-

O-methyl derivatives to double-helical DNA. *C R Acad Sci III* **315:** 521-525

65. Faye G, Dennebouy N, Kujawa C, Jacq C (1979) Inserted sequence in the mitochondrial 23S ribosomal RNA gene of the yeast Saccharomyces cerevisiae. *Mol Gen Genet* **168:** 101-109

66. Farhood H, Serbina N, Huang L (1995) The role of dioleoyl phosphatidylethanolamine in cationic liposome mediated gene transfer. *Biochim Biophys Acta* **1235:** 289-295

67. Faruqi AF, Egholm M, Glazer PM (1998) Peptide nucleic acid-targeted mutagenesis of a chromosomal gene in mouse cells. *Proc Natl Acad Sci U S A* **95:** 1398-1403

68. Felsenfeld G, Rich A (1957) Studies on the formation of two- and three-stranded polyribonucleotides. *Biochim Biophys Acta* **26:** 457-468

69. Flick KE, Jurica MS, Monnat RJ, Jr., Stoddard BL (1998) DNA binding and cleavage by the nuclear intron-encoded homing endonuclease I-PpoI. *Nature* **394:** 96-101

70. Floris R, Scaggiante B, Manzini G, Quadrifoglio F, Xodo LE (1999) Effect of cations on purine.purine.pyrimidine triple helix formation in mixed-valence salt solutions. *Eur J Biochem* **260:** 801-809

71. Fox KR (2000) Targeting DNA with triplexes. *Curr Med Chem* **7:** 17-37

72. Francois JC, Saison-Behmoaras T, Barbier C, Chassignol M, Thuong NT, Helene C (1989a) Sequence-specific recognition and cleavage of duplex DNA *via* triple-helix formation by oligonucleotides covalently linked to a phenanthroline-copper chelate. *Proc Natl Acad Sci U S A* **86:** 9702-9706

73. Francois JC, Saison-Behmoaras T, Chassignol M, Thuong NT, Helene C (1989b) Sequence-targeted cleavage of single- and double-stranded DNA by oligothymidylates covalently linked to 1,10-phenanthroline. *J Biol Chem* **264:** 5891-5898

74. Fritsche P, Alves J (2004) A monomeric mutant of restriction endonuclease EcoRI nicks DNA without sequence specificity. *Biol Chem* **385:** 975-985

75. Geel TM, Meiss G, van der Gun BT, Kroesen BJ, de Leij LF, Zaremba M, Silanskas A, Kokkinidis M, Pingoud A, Ruiters MH, McLaughlin

PM, Rots MG (2009) Endonucleases induced TRAIL-insensitive apoptosis in ovarian carcinoma cells. *Exp Cell Res* **315:** 2487-2495

76. Geese WJ, Kwon YK, Wen X, Waring RB (2003) In vitro analysis of the relationship between endonuclease and maturase activities in the bifunctional group I intron-encoded protein, I-AniI. *Eur J Biochem* **270:** 1543-1554

77. Geissler R, Scholze H, Hahn S, Streubel J, Bonas U, Behrens SE, Boch J (2011) Transcriptional activators of human genes with programmable DNA-specificity. *PLoS One* **6:** e19509

78. Gimble FS, Moure CM, Posey KL (2003) Assessing the plasticity of DNA target site recognition of the PI-SceI homing endonuclease using a bacterial two-hybrid selection system. *J Mol Biol* **334:** 993-1008

79. Giovannangeli C, Thuong NT, Helene C (1992) Oligodeoxynucleotide-directed photo-induced cross-linking of HIV proviral DNA *via* triple-helix formation. *Nucleic Acids Res* **20:** 4275-4281

80. Goeldner M, Givens R (eds) (2005) *Dynamic studies in biology: phototriggers, photoswitches and caged biomolecules*. Weinheim: Wiley-VCH

81. Gowers DM, Bijapur J, Brown T, Fox KR (1999) DNA triple helix formation at target sites containing several pyrimidine interruptions: stabilization by protonated cytosine or 5-(1-propargylamino)dU. *Biochemistry* **38:** 13747-13758

82. Grazulis S, Deibert M, Rimseliene R, Skirgaila R, Sasnauskas G, Lagunavicius A, Repin V, Urbanke C, Huber R, Siksnys V (2002) Crystal structure of the Bse634I restriction endonuclease: comparison of two enzymes recognizing the same DNA sequence. *Nucleic Acids Res* **30:** 876-885

83. Grieger JC, Samulski RJ (2005) Adeno-associated virus as a gene therapy vector: vector development, production and clinical applications. *Adv Biochem Eng Biotechnol* **99:** 119-145

84. Grigoriev M, Praseuth D, Robin P, Hemar A, Saison-Behmoaras T, Dautry-Varsat A, Thuong NT, Helene C, Harel-Bellan A (1992) A triple helix-forming oligonucleotide-intercalator conjugate acts as a transcriptional repressor *via* inhibition of NF kappa B binding to interleukin-2 receptor alpha-regulatory sequence. *J Biol Chem* **267:** 3389-3395

85. Grizot S, Epinat JC, Thomas S, Duclert A, Rolland S, Paques F, Duchateau P (2010) Generation of redesigned homing endonucleases comprising DNA-binding domains derived from two different scaffolds. *Nucleic Acids Res* **38:** 2006-2018

86. Grizot S, Smith J, Daboussi F, Prieto J, Redondo P, Merino N, Villate M, Thomas S, Lemaire L, Montoya G, Blanco FJ, Paques F, Duchateau P (2009) Efficient targeting of a SCID gene by an engineered single-chain homing endonuclease. *Nucleic Acids Res* **37:** 5405-5419

87. Gruen M, Chang K, Serbanescu I, Liu DR (2002) An in vivo selection system for homing endonuclease activity. *Nucleic Acids Res* **30:** e29

88. Gu K, Yang B, Tian D, Wu L, Wang D, Sreekala C, Yang F, Chu Z, Wang GL, White FF, Yin Z (2005) R gene expression induced by a type-III effector triggers disease resistance in rice. *Nature* **435:** 1122-1125

89. Haq I, O'Brien R, Lagunavicius A, Siksnys V, Ladbury JE (2001) Specific DNA recognition by the type II restriction endonuclease MunI: the effect of pH. *Biochemistry* **40:** 14960-14967

90. Heath PJ, Stephens KM, Monnat RJ, Jr., Stoddard BL (1997) The structure of I-Crel, a group I intron-encoded homing endonuclease. *Nat Struct Biol* **4:** 468-476

91. Heller R, Jaroszeski M, Atkin A, Moradpour D, Gilbert R, Wands J, Nicolau C (1996) In vivo gene electroinjection and expression in rat liver. *FEBS Lett* **389:** 225-228

92. Heuer H, Yin YN, Xue QY, Smalla K, Guo JH (2007) Repeat domain diversity of avrBs3-like genes in Ralstonia solanacearum strains and association with host preferences in the field. *Appl Environ Microbiol* **73:** 4379-4384

93. Hickman MA, Malone RW, Lehmann-Bruinsma K, Sih TR, Knoell D, Szoka FC, Walzem R, Carlson DM, Powell JS (1994) Gene expression following direct injection of DNA into liver. *Hum Gene Ther* **5:** 1477-1483

94. Horner SM, DiMaio D (2007) The DNA binding domain of a papillomavirus E2 protein programs a chimeric nuclease to cleave integrated human papillomavirus DNA in HeLa cervical carcinoma cells. *J Virol* **81:** 6254-6264

95. Yagi H, Ogura T, Mizukami H, Urabe M, Hamada H, Yoshikawa H, Ozawa K, Kume A (2011) Complete restoration of phenylalanine oxidation in phenylketonuria mouse by a self-complementary adeno-associated virus vector. *J Gene Med* **13:** 114-122

96. Yoo JJ, Soker S, Lin LF, Mehegan K, Guthrie PD, Atala A (1999) Direct in vivo gene transfer to urological organs. *J Urol* **162:** 1115-1118

97. Jacquier A, Dujon B (1985) An intron-encoded protein is active in a gene conversion process that spreads an intron into a mitochondrial gene. *Cell* **41:** 383-394

98. Jain A, Wang G, Vasquez KM (2008) DNA triple helices: biological consequences and therapeutic potential. *Biochimie* **90:** 1117-1130

99. Jenkins FJ, Turner SL (1996) Herpes simplex virus: a tool for neuroscientists. *Front Biosci* **1:** d241-247

100.   Jurica MS, Monnat RJ, Jr., Stoddard BL (1998) DNA recognition and cleavage by the LAGLIDADG homing endonuclease I-CreI. *Mol Cell* **2:** 469-476

101.   Kelly EB (ed) (2007) *Gene Therapy*: Greenwood Publishing Group

102.   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

103.   Kim YG, Chandrasegaran S (1994) Chimeric restriction endonuclease. *Proc Natl Acad Sci U S A* **91:** 883-887

104.   Kim YG, Kim PS, Herbert A, Rich A (1997) Construction of a Z-DNA-specific restriction endonuclease. *Proc Natl Acad Sci U S A* **94:** 12875-12879

105.   Kim YG, Smith J, Durgesha M, Chandrasegaran S (1998) Chimeric restriction enzyme: Gal4 fusion to FokI cleavage domain. *Biol Chem* **379:** 489-495

106.   Krawczyk SH, Milligan JF, Wadwani S, Moulds C, Froehler BC, Matteucci MD (1992) Oligonucleotide-mediated triple helix formation using an N3-protonated deoxycytidine analog exhibiting pH-independent binding within the physiological range. *Proc Natl Acad Sci U S A* **89:** 3761-3764

107. Kresina TF (ed) (2001) *AN INTRODUCTION TO MOLECULAR MEDICINE AND GENE THERAPY*: A JOHN WILEY & SONS, INC.

108. Krishna SS, Majumdar I, Grishin NV (2003) Structural classification of zinc fingers: survey and summary. *Nucleic Acids Res* **31:** 532-550

109. Lagunavicius A, Grazulis S, Balciunaite E, Vainius D, Siksnys V (1997) DNA binding specificity of MunI restriction endonuclease is controlled by pH and calcium ions: involvement of active site carboxylate residues. *Biochemistry* **36:** 11093-11099

110. Lambowitz AM, Belfort M (1993) Introns as mobile genetic elements. *Annu Rev Biochem* **62:** 587-622

111. Landgraf R, Chen CH, Sigman DS (1994) Oligonucleotide-directed nucleic acid scission by micrococcal nuclease. *Biochemistry* **33:** 10607-10615

112. Lanzov VA (1999) Gene targeting for gene therapy: prospects. *Mol Genet Metab* **68:** 276-282

113. Lariguet P, Dunand C, Herzog M, Vachon G (1999) APETALA3-nuclease hybrid protein: a potential tool for APETALA3 target gene mutagenesis. *Plant Sci* **148:** 19-30

114. Le Doan T, Perrouault L, Praseuth D, Habhoub N, Decout JL, Thuong NT, Lhomme J, Helene C (1987) Sequence-specific recognition, photocrosslinking and cleavage of the DNA double helix by an oligo-[alpha]-thymidylate covalently linked to an azidoproflavine derivative. *Nucleic Acids Res* **15:** 7749-7760

115. Lee JS, Woodsworth ML, Latimer LJ, Morgan AR (1984) Poly(pyrimidine) . poly(purine) synthetic DNAs containing 5-methylcytosine form stable triplexes at neutral pH. *Nucleic Acids Res* **12:** 6603-6614

116. Lee MS, Gippert GP, Soman KV, Case DA, Wright PE (1989) Three-dimensional solution structure of a single zinc finger DNA-binding domain. *Science* **245:** 635-637

117. Li L, Wu LP, Chandrasegaran S (1992) Functional domains in Fok I restriction endonuclease. *Proc Natl Acad Sci U S A* **89:** 4275-4279

118. Li T, Huang S, Zhao X, Wright DA, Carpenter S, Spalding MH, Weeks DP, Yang B (2011) Modularly assembled designer TAL effector

nucleases for targeted gene knockout and gene replacement in eukaryotes. *Nucleic Acids Res* **39:** 6315-6325

119.   Lippow SM, Aha PM, Parker MH, Blake WJ, Baynes BM, Lipovsek D (2009) Creation of a type IIS restriction endonuclease with a long recognition sequence. *Nucleic Acids Res* **37:** 3061-3073

120.   Liu PQ, Rebar EJ, Zhang L, Liu Q, Jamieson AC, Liang Y, Qi H, Li PX, Chen B, Mendel MC, Zhong X, Lee YL, Eisenberg SP, Spratt SK, Case CC, Wolffe AP (2001) Regulation of an endogenous locus using a panel of designed zinc finger proteins targeted to accessible chromatin regions. Activation of vascular endothelial growth factor A. *J Biol Chem* **276:** 11323-11334

121.   Lloyd A, Plaisier CL, Carroll D, Drews GN (2005) Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proc Natl Acad Sci U S A* **102:** 2232-2237

122.   Lucas P, Otis C, Mercier JP, Turmel M, Lemieux C (2001) Rapid evolution of the DNA-binding site in LAGLIDADG homing endonucleases. *Nucleic Acids Res* **29:** 960-969

123.   Mahfouz MM, Li L, Shamimuzzaman M, Wibowo A, Fang X, Zhu JK (2011) De novo-engineered transcription activator-like effector (TALE) hybrid nuclease with novel DNA binding specificity creates double-strand breaks. *Proc Natl Acad Sci U S A* **108:** 2623-2628

124.   Majumdar A, Muniandy PA, Liu J, Liu JL, Liu ST, Cuenoud B, Seidman MM (2008) Targeted gene knock in and sequence modulation mediated by a psoralen-linked triplex-forming oligonucleotide. *J Biol Chem* **283:** 11244-11252

125.   Mak AN, Bradley P, Cernadas RA, Bogdanove AJ, Stoddard BL (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science* **335:** 716-719

126.   Marchand C, Bailly C, Nguyen CH, Bisagni E, Garestier T, Helene C, Waring MJ (1996) Stabilization of triple helical DNA by a benzopyridoquinoxaline intercalator. *Biochemistry* **35:** 5022-5032

127.   Mastrobattista E, Hennink WE (2012) Polymers for gene delivery: Charged for success. *Nat Mater* **11:** 10-12

128.   Maurisse R, Feugeas JP, Biet E, Kuzniak I, Leboulch P, Dutreix M, Sun JS (2002) A new method (GOREC) for directed mutagenesis and gene repair by homologous recombination. *Gene Ther* **9:** 703-707

129.   Mergny JL, Duval-Valentin G, Nguyen CH, Perrouault L, Faucon B, Rougee M, Montenay-Garestier T, Bisagni E, Helene C (1992) Triple helix-specific ligands. *Science* **256:** 1681-1684

130.   Merickel SK, Sanders ER, Vazquez-Ibar JL, Johnson RC (2002) Subunit exchange and the role of dimer flexibility in DNA binding by the Fis protein. *Biochemistry* **41:** 5788-5798

131.   Michel T, Debart F, Heitz F, Vasseur JJ (2005) Highly stable DNA triplexes formed with cationic phosphoramidate pyrimidine alpha-oligonucleotides. *Chembiochem* **6:** 1254-1262

132.   Michel T, Martinand-Mari C, Debart F, Lebleu B, Robbins I, Vasseur JJ (2003) Cationic phosphoramidate alpha-oligonucleotides efficiently target single-stranded DNA and RNA and inhibit hepatitis C virus IRES-mediated translation. *Nucleic Acids Res* **31:** 5282-5290

133.   Miller DL, Schildbach JF (2003) Evidence for a monomeric intermediate in the reversible unfolding of F factor TraM. *J Biol Chem* **278:** 10400-10407

134.   Miller J, McLachlan AD, Klug A (1985) Repetitive zinc-binding domains in the protein transcription factor IIIA from Xenopus oocytes. *Embo J* **4:** 1609-1614

135.   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

136.   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

137.   Milligan JF, Krawczyk SH, Wadwani S, Matteucci MD (1993) An anti-parallel triple helix motif with oligodeoxynucleotides containing 2'-deoxyguanosine and 7-deaza-2'-deoxyxanthosine. *Nucleic Acids Res* **21:** 327-333

138.   Mills M, Arimondo PB, Lacroix L, Garestier T, Klump H, Mergny JL (2002) Chemical modification of the third strand: differential effects

on purine and pyrimidine triple helix formation. *Biochemistry* **41:** 357-366

139.   Mir LM, Banoun H, Paoletti C (1988) Introduction of definite amounts of nonpermeant molecules into living cells after electropermeabilization: direct access to the cytosol. *Exp Cell Res* **175:** 15-25

140.   Mombaerts P, Clarke AR, Hooper ML, Tonegawa S (1991) Creation of a large genomic deletion at the T-cell antigen receptor beta-subunit locus in mouse embryonic stem cells by gene targeting. *Proc Natl Acad Sci U S A* **88:** 3084-3087

141.   Moolten FL (1994) Drug sensitivity ("suicide") genes for selective cancer chemotherapy. *Cancer Gene Ther* **1:** 279-287

142.   Morbitzer R, Romer P, Boch J, Lahaye T (2010) Regulation of selected genome loci using de novo-engineered transcription activator-like effector (TALE)-type transcription factors. *Proc Natl Acad Sci U S A* **107:** 21617-21622

143.   Moscou MJ, Bogdanove AJ (2009) A simple cipher governs DNA recognition by TAL effectors. *Science* **326:** 1501

144.   Moure CM, Gimble FS, Quiocho FA (2008) Crystal structures of I-SceI complexed to nicked DNA substrates: snapshots of intermediates along the DNA cleavage reaction pathway. *Nucleic Acids Res* **36:** 3287-3296

145.   Mumper RJ, Duguid JG, Anwer K, Barron MK, Nitta H, Rolland AP (1996) Polyvinyl derivatives as novel interactive polymers for controlled gene delivery to muscle. *Pharm Res* **13:** 701-709

146.   Nakayama K, Endo M, Majima T (2005) A hydrophilic azobenzene-bearing amino acid for photochemical control of a restriction enzyme BamHI. *Bioconjug Chem* **16:** 1360-1366

147.   Nishikawa M, Huang L (2001) Nonviral vectors in the new millennium: delivery barriers in gene transfer. *Hum Gene Ther* **12:** 861-870

148.   Olivas WM, Maher LJ, 3rd (1995) Overcoming potassium-mediated triplex inhibition. *Nucleic Acids Res* **23:** 1936-1941

149.   Pabo CO, Peisach E, Grant RA (2001) Design and selection of novel Cys2His2 zinc finger proteins. *Annu Rev Biochem* **70:** 313-340

150.   Papworth M, Moore M, Isalan M, Minczuk M, Choo Y, Klug A (2003) Inhibition of herpes simplex virus 1 gene expression by designer zinc-finger transcription factors. *Proc Natl Acad Sci U S A* **100:** 1621-1626

151.   Paques F, Duchateau P (2007) Meganucleases and DNA double-strand break-induced recombination: perspectives for gene therapy. *Curr Gene Ther* **7:** 49-66

152.   Park SY, Quezada CM, Bilwes AM, Crane BR (2004) Subunit exchange by CheA histidine kinases from the mesophile Escherichia coli and the thermophile Thermotoga maritima. *Biochemistry* **43:** 2228-2240

153.   Pavletich NP, Pabo CO (1991) Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 A. *Science* **252:** 809-817

154.   Pei D, Corey DR, Schultz PG (1990) Site-specific cleavage of duplex DNA by a semisynthetic nuclease *via* triple-helix formation. *Proc Natl Acad Sci U S A* **87:** 9858-9862

155.   Perrouault L, Asseline U, Rivalle C, Thuong NT, Bisagni E, Giovannangeli C, Le Doan T, Helene C (1990) Sequence-specific artificial photo-induced endonucleases based on triple helix-forming oligonucleotides. *Nature* **344:** 358-360

156.   Petersen M, Wengel J (2003) LNA: a versatile tool for therapeutics and genomics. *Trends Biotechnol* **21:** 74-81

157.   Phipps AK, Tarkoy M, Schultze P, Feigon J (1998) Solution structure of an intramolecular DNA triplex containing 5-(1-propynyl)-2'-deoxyuridine residues in the third strand. *Biochemistry* **37:** 5820-5830

158.   Pilch DS, Waring MJ, Sun JS, Rougee M, Nguyen CH, Bisagni E, Garestier T, Helene C (1993) Characterization of a triple helix-specific ligand. BePI (3-methoxy-7H-8-methyl-11- [(3'-amino)propylamino]-benzo[e]pyrido[4,3-b]indole) intercalates into both double-helical and triple-helical DNA. *J Mol Biol* **232:** 926-946

159.   Pingoud A, Fuxreiter M, Pingoud V, Wende W (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol Life Sci* **62:** 685-707

160. Pingoud A, Silva G, Wende W (2011) Meganucleases: promising tools for gene therapy. In *Modern Biopharmaceuticals. Recent Success Stories*, Knäblein J (ed), Vol. (in press). Wiley-VCH

161. Porteus MH (2006) Mammalian gene targeting with designed zinc finger nucleases. *Mol Ther* **13:** 438-446

162. Porteus MH, Baltimore D (2003) Chimeric nucleases stimulate gene targeting in human cells. *Science* **300:** 763

163. Porteus MH, Carroll D (2005) Gene targeting using zinc finger nucleases. *Nat Biotechnol* **23:** 967-973

164. Praseuth D, Perrouault L, Le Doan T, Chassignol M, Thuong N, Helene C (1988) Sequence-specific binding and photocrosslinking of alpha and beta oligodeoxynucleotides to the major groove of DNA *via* triple-helix formation. *Proc Natl Acad Sci U S A* **85:** 1349-1353

165. 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

166. 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

167. Radhakrishnan I, Patel DJ (1994) Solution structure of a pyrimidine.purine.pyrimidine DNA triplex containing T.AT, C+.GC and G.TA triples. *Structure* **2:** 17-32

168. Rahman SM, Seki S, Obika S, Yoshikawa H, Miyashita K, Imanishi T (2008) Design, synthesis, and properties of 2',4'-BNA(NC): a bridged nucleic acid analogue. *J Am Chem Soc* **130:** 4886-4896

169. Ramirez CL, Foley JE, Wright DA, Muller-Lerch F, Rahman SH, Cornu TI, Winfrey RJ, Sander JD, Fu F, Townsend JA, Cathomen T, Voytas DF, Joung JK (2008) Unexpected failure rates for modular assembly of engineered zinc fingers. *Nat Methods* **5:** 374-375

170. Rao TS, Durland RH, Seth DM, Myrick MA, Bodepudi V, Revankar GR (1995) Incorporation of 2'-deoxy-6-thioguanosine into G-rich oligodeoxyribonucleotides inhibits G-tetrad formation and facilitates triplex formation. *Biochemistry* **34:** 765-772

171. Rathert P, Rasko T, Roth M, Slaska-Kiss K, Pingoud A, Kiss A, Jeltsch A (2007) Reversible inactivation of the CG specific SssI DNA

(cytosine-C5)-methyltransferase with a photocleavable protecting group. *Chembiochem* **8:** 202-207

172.   Rebar EJ, Huang Y, Hickey R, Nath AK, Meoli D, Nath S, Chen B, Xu L, Liang Y, Jamieson AC, Zhang L, Spratt SK, Case CC, Wolffe A, Giordano FJ (2002) Induction of angiogenesis in a mouse model using engineered transcription factors. *Nat Med* **8:** 1427-1432

173.   Redondo P, Prieto J, Munoz IG, Alibes A, Stricher F, Serrano L, Cabaniols JP, Daboussi F, Arnould S, Perez C, Duchateau P, Paques F, Blanco FJ, Montoya G (2008) Molecular basis of xeroderma pigmentosum group C DNA recognition by engineered meganucleases. *Nature* **456:** 107-111

174.   Reynolds L, Ullman C, Moore M, Isalan M, West MJ, Clapham P, Klug A, Choo Y (2003) Repression of the HIV-1 5' LTR promoter and inhibition of HIV-1 replication by using engineered zinc-finger transcription factors. *Proc Natl Acad Sci U S A* **100:** 1615-1620

175.   Renner C, Moroder L (2006) Azobenzene as conformational switch in model peptides. *Chembiochem* **7:** 868-878

176.   Repin VE, Lebedev LR, Puchkova L, Serov GD, Tereschenko T, Chizikov VE, Andreeva I (1995) New restriction endonucleases from thermophilic soil bacteria. *Gene* **157:** 321-322

177.   Rogers FA, Manoharan M, Rabinovitch P, Ward DC, Glazer PM (2004) Peptide conjugates for chromosomal gene targeting by triplex-forming oligonucleotides. *Nucleic Acids Res* **32:** 6595-6604

178.   Rogers FA, Vasquez KM, Egholm M, Glazer PM (2002) Site-directed recombination *via* bifunctional PNA-DNA conjugates. *Proc Natl Acad Sci U S A* **99:** 16695-16700

179.   Rols MP, Delteil C, Golzio M, Dumond P, Cros S, Teissie J (1998) In vivo electrically mediated protein and gene transfer in murine melanoma. *Nat Biotechnol* **16:** 168-171

180.   Rosen LE, Morrison HA, Masri S, Brown MJ, Springstubb B, Sussman D, Stoddard BL, Seligman LM (2006) Homing endonuclease I-CreI derivatives with novel DNA target specificities. *Nucleic Acids Res* **34:** 4791-4800

181.   Rubin GM, Yandell MD, Wortman JR, Gabor Miklos GL, Nelson CR, Hariharan IK, Fortini ME, Li PW, Apweiler R, Fleischmann W, Cherry JM, Henikoff S, Skupski MP, Misra S, Ashburner M, Birney E,

Boguski MS, Brody T, Brokstein P, Celniker SE, Chervitz SA, Coates D, Cravchik A, Gabrielian A, Galle RF, Gelbart WM, George RA, Goldstein LS, Gong F, Guan P, Harris NL, Hay BA, Hoskins RA, Li J, Li Z, Hynes RO, Jones SJ, Kuehl PM, Lemaitre B, Littleton JT, Morrison DK, Mungall C, O'Farrell PH, Pickeral OK, Shue C, Vosshall LB, Zhang J, Zhao Q, Zheng XH, Lewis S (2000) Comparative genomics of the eukaryotes. *Science* **287:** 2204-2215

182.   Ruminy P, Derambure C, Chandrasegaran S, Salier JP (2001) Long-range identification of hepatocyte nuclear factor-3 (FoxA) high and low-affinity binding sites with a chimeric nuclease. *J Mol Biol* **310:** 523-535

183.   Rusling DA, Powers VE, Ranasinghe RT, Wang Y, Osborne SD, Brown T, Fox KR (2005) Four base recognition by triplex-forming oligonucleotides at physiological pH. *Nucleic Acids Res* **33:** 3025-3032

184.   Salanoubat M, Genin S, Artiguenave F, Gouzy J, Mangenot S, Arlat M, Billault A, Brottier P, Camus JC, Cattolico L, Chandler M, Choisne N, Claudel-Renard C, Cunnac S, Demange N, Gaspin C, Lavie M, Moisan A, Robert C, Saurin W, Schiex T, Siguier P, Thebault P, Whalen M, Wincker P, Levy M, Weissenbach J, Boucher CA (2002) Genome sequence of the plant pathogen Ralstonia solanacearum. *Nature* **415:** 497-502

185.   Sambrook J, Fritsch, E.F., & Maniatis, T. (ed) (1989) *Molecular Cloning, A Laboratory manual.* Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press

186.   Sasnauskas G, Jeltsch A, Pingoud A, Siksnys V (1999) Plasmid DNA cleavage by MunI restriction enzyme: single-turnover and steady-state kinetic analysis. *Biochemistry* **38:** 4028-4036

187.   Scalley-Kim M, McConnell-Smith A, Stoddard BL (2007) Coevolution of a homing endonuclease and its host target sequence. *J Mol Biol* **372:** 1305-1319

188.   Scaria PV, Shafer RH (1991) Binding of ethidium bromide to a DNA triple helix. Evidence for intercalation. *J Biol Chem* **266:** 5417-5423

189.   Schierling B, Noel AJ, Wende W, Hien le T, Volkov E, Kubareva E, Oretskaya T, Kokkinidis M, Rompp A, Spengler B, Pingoud A (2010) Controlling the enzymatic activity of a restriction enzyme by light. *Proc Natl Acad Sci U S A* **107:** 1361-1366

190. Schneider F, Hammarstrom P, Kelly JW (2001) Transthyretin slowly exchanges subunits under physiological conditions: A convenient chromatographic method to study subunit exchange in oligomeric proteins. *Protein Sci* **10:** 1606-1613

191. Scholze H, Boch J (2011) TAL effectors are remote controls for gene activation. *Curr Opin Microbiol* **14:** 47-53

192. Schwartz B, Benoist C, Abdallah B, Rangara R, Hassan A, Scherman D, Demeneix BA (1996) Gene transfer by naked DNA into adult mouse brain. *Gene Ther* **3:** 405-411

193. Searles MA, Lu D, Klug A (2000) The role of the central zinc fingers of transcription factor IIIA in binding to 5 S RNA. *J Mol Biol* **301:** 47-60

194. Seidman MM, Puri N, Majumdar A, Cuenoud B, Miller PS, Alam R (2005) The development of bioactive triple helix-forming oligonucleotides. *Ann N Y Acad Sci* **1058:** 119-127

195. Seligman LM, Chisholm KM, Chevalier BS, Chadsey MS, Edwards ST, Savage JH, Veillet AL (2002) Mutations altering the cleavage specificity of a homing endonuclease. *Nucleic Acids Res* **30:** 3870-3879

196. Shafer RH (1998) Stability and structure of model DNA triplexes and quadruplexes and their interactions with small ligands. *Prog Nucleic Acid Res Mol Biol* **59:** 55-94

197. Shen BW, Landthaler M, Shub DA, Stoddard BL (2004) DNA binding and cleavage by the HNH homing endonuclease I-HmuI. *J Mol Biol* **342:** 43-56

198. Shimizu M, Konishi A, Shimada Y, Inoue H, Ohtsuka E (1992) Oligo(2'-O-methyl)ribonucleotides. Effective probes for duplex DNA. *FEBS Lett* **302:** 155-158

199. Sikes ML, O'Malley BW, Jr., Finegold MJ, Ledley FD (1994) In vivo gene transfer into rabbit thyroid follicular cells by direct DNA injection. *Hum Gene Ther* **5:** 837-844

200. Siksnys V, Pleckaityte M (1992) Role of the reactive cysteine residue in restriction endonuclease Cfr9I. *Biochim Biophys Acta* **1160:** 199-205

201.   Siksnys V, Skirgaila R, Sasnauskas G, Urbanke C, Cherny D, Grazulis S, Huber R (1999) The Cfr10I restriction enzyme is functional as a tetramer. *J Mol Biol* **291:** 1105-1118

202.   Siksnys V, Zareckaja N, Vaisvila R, Timinskas A, Stakenas P, Butkus V, Janulaitis A (1994) CAATTG-specific restriction-modification munI genes from Mycoplasma: sequence similarities between R.MunI and R.EcoRI. *Gene* **142:** 1-8

203.   Silanskas A, Foss M, Wende W, Urbanke C, Lagunavicius A, Pingoud A, Siksnys V (2011) Photocaged variants of the MunI and PvuII restriction enzymes. *Biochemistry* **50:** 2800-2807

204.   Silva G, Poirot L, Galetto R, Smith J, Montoya G, Duchateau P, Paques F (2011) Meganucleases and other tools for targeted genome engineering: perspectives and challenges for gene therapy. *Curr Gene Ther* **11:** 11-27

205.   Silva GH, Belfort M (2004) Analysis of the LAGLIDADG interface of the monomeric homing endonuclease I-DmoI. *Nucleic Acids Res* **32:** 3156-3168

206.   Silva GH, Belfort M, Wende W, Pingoud A (2006) From monomeric to homodimeric endonucleases and back: engineering novel specificity of LAGLIDADG enzymes. *J Mol Biol* **361:** 744-754

207.   Simon P, Cannata F, Perrouault L, Halby L, Concordet JP, Boutorine A, Ryabinin V, Sinyakov A, Giovannangeli C (2008) Sequence-specific DNA cleavage mediated by bipyridine polyamide conjugates. *Nucleic Acids Res* **36:** 3531-3538

208.   Simoncsits A, Tjornhammar ML, Rasko T, Kiss A, Pongor S (2001) Covalent joining of the subunits of a homodimeric type II restriction endonuclease: single-chain PvuII endonuclease. *J Mol Biol* **309:** 89-97

209.   Sims PA, Menefee AL, Larsen TM, Mansoorabadi SO, Reed GH (2006) Structure and catalytic properties of an engineered heterodimer of enolase composed of one active and one inactive subunit. *J Mol Biol* **355:** 422-431

210.   Singleton SF, Dervan PB (1993) Equilibrium association constants for oligonucleotide-directed triple helix formation at single DNA sites: linkage to cation valence and concentration. *Biochemistry* **32:** 13171-13179

211.   Smith J, Grizot S, Arnould S, Duclert A, Epinat JC, Chames P, Prieto J, Redondo P, Blanco FJ, Bravo J, Montoya G, Paques F, Duchateau P (2006) A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res* **34:** e149

212.   Smithies O, Gregg RG, Boggs SS, Koralewski MA, Kucherlapati RS (1985) Insertion of DNA sequences into the human chromosomal beta-globin locus by homologous recombination. *Nature* **317:** 230-234

213.   Snyder RO, Miao CH, Patijn GA, Spratt SK, Danos O, Nagy D, Gown AM, Winther B, Meuse L, Cohen LK, Thompson AR, Kay MA (1997) Persistent and therapeutic concentrations of human factor IX in mice after hepatic gene transfer of recombinant AAV vectors. *Nat Genet* **16:** 270-276

214.   Sobott F, Benesch JL, Vierling E, Robinson CV (2002) Subunit exchange of multimeric protein complexes. Real-time monitoring of subunit exchange between small heat shock proteins by using electrospray mass spectrometry. *J Biol Chem* **277:** 38921-38929

215.   Southwell AL, Patterson PH (2011) Gene therapy in mouse models of huntington disease. *Neuroscientist* **17:** 153-162

216.   Stahl F, Wende W, Jeltsch A, Pingoud A (1996) Introduction of asymmetry in the naturally symmetric restriction endonuclease EcoRV to investigate intersubunit communication in the homodimeric protein. *Proc Natl Acad Sci U S A* **93:** 6175-6180

217.   Stakenas PS, Zaretskaia NM, Manelene ZP, Mauritsas MM, Butkus VV, Ianulaitis AA (1992) [Mycoplasma restriction-modification system MunI and its possible role in pathogenesis processes]. *Mol Biol (Mosk)* **26:** 546-557

218.   Steuer S, Pingoud V, Pingoud A, Wende W (2004) Chimeras of the homing endonuclease PI-SceI and the homologous Candida tropicalis intein: a study to explore the possibility of exchanging DNA-binding modules to obtain highly specific endonucleases with altered specificity. *Chembiochem* **5:** 206-213

219.   Stoddard BL (2005) Homing endonuclease structure and function. *Q Rev Biophys* **38:** 49-95

220.   Stoddard BL (2011) Homing endonucleases: from microbial genetic invaders to reagents for targeted DNA modification. *Structure* **19:** 7-15

221.   Stonehouse TJ, Fox KR (1994) DNase I footprinting of triple helix formation at polypurine tracts by acridine-linked oligopyrimidines: stringency, structural changes and interaction with minor groove binding ligands. *Biochim Biophys Acta* **1218:** 322-330

222.   Strobel SA, Dervan PB (1990) Site-specific cleavage of a yeast chromosome by oligonucleotide-directed triple-helix formation. *Science* **249:** 73-75

223.   Sun BW, Babu BR, Sorensen MD, Zakrzewska K, Wengel J, Sun JS (2004) Sequence and pH effects of LNA-containing triple helix-forming oligonucleotides: physical chemistry, biochemistry, and modeling studies. *Biochemistry* **43:** 4160-4169

224.   Sun JS, Francois JC, Montenay-Garestier T, Saison-Behmoaras T, Roig V, Thuong NT, Helene C (1989) Sequence-specific intercalating agents: intercalation at specific sequences on duplex DNA *via* major groove recognition by oligonucleotide-intercalator conjugates. *Proc Natl Acad Sci U S A* **86:** 9198-9202

225.   Sun JS, Giovannangeli C, Francois JC, Kurfurst R, Montenay-Garestier T, Asseline U, Saison-Behmoaras T, Thuong NT, Helene C (1991) Triple-helix formation by alpha oligodeoxynucleotides and alpha oligodeoxynucleotide-intercalator conjugates. *Proc Natl Acad Sci U S A* **88:** 6023-6027

226.   Sun N, Liang J, Abil Z, Zhao H (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol Biosyst*

227.   Sun X, Zhang N (2010) Cationic polymer optimization for efficient gene delivery. *Mini Rev Med Chem* **10:** 108-125

228.   Sussman D, Chadsey M, Fauce S, Engel A, Bruett A, Monnat R, Jr., Stoddard BL, Seligman LM (2004) Isolation and characterization of new homing endonuclease specificities at individual target site positions. *J Mol Biol* **342:** 31-41

229.   Szczepek M, Brondani V, Buchel 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

230.   Takasugi M, Guendouz A, Chassignol M, Decout JL, Lhomme J, Thuong NT, Helene C (1991) Sequence-specific photo-induced cross-

linking of the two strands of double-helical DNA by a psoralen covalently linked to a triple helix-forming oligonucleotide. *Proc Natl Acad Sci U S A* **88:** 5602-5606

231.   Thomas T, Thomas TJ (1993) Selectivity of polyamines in triplex DNA stabilization. *Biochemistry* **32:** 14068-14074

232.   Thomas TJ, Kulkarni GD, Greenfield NJ, Shirahata A, Thomas T (1996) Structural specificity effects of trivalent polyamine analogues on the stabilization and conformational plasticity of triplex DNA. *Biochem J* **319 (Pt 2):** 591-599

233.   Thompson AJ, Yuan X, Kudlicki W, Herrin DL (1992) Cleavage and recognition pattern of a double-strand-specific endonuclease (I-creI) encoded by the chloroplast 23S rRNA intron of Chlamydomonas reinhardtii. *Gene* **119:** 247-251

234.   Titomirov AV, Sukharev S, Kistanova E (1991) In vivo electroporation and stable transformation of skin cells of newborn mice by plasmid DNA. *Biochim Biophys Acta* **1088:** 131-134

235.   Torigoe H, Hari Y, Sekiguchi M, Obika S, Imanishi T (2001) 2'-O,4'-C-methylene bridged nucleic acid modification promotes pyrimidine motif triplex DNA formation at physiological pH: thermodynamic and kinetic studies. *J Biol Chem* **276:** 2354-2360

236.   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

237.   Urnov FD, Rebar EJ, Holmes MC, Zhang HS, Gregory PD (2010) Genome editing with engineered zinc finger nucleases. *Nat Rev Genet* **11:** 636-646

238.   Van Roey P, Meehan L, Kowalski JC, Belfort M, Derbyshire V (2002) Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. *Nat Struct Biol* **9:** 806-811

239.   Van Roey P, Waddling CA, Fox KM, Belfort M, Derbyshire V (2001) Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. *Embo J* **20:** 3631-3637

240.   Vasquez KM, Wensel TG, Hogan ME, Wilson JH (1996) High-efficiency triple-helix-mediated photo-cross-linking at a targeted site within a selectable mammalian gene. *Biochemistry* **35:** 10712-10719

241.   Wah DA, Bitinaite J, Schildkraut I, Aggarwal AK (1998) Structure of FokI has implications for DNA cleavage. *Proc Natl Acad Sci U S A* **95:** 10564-10569

242.   Wah DA, Hirsch JA, Dorner LF, Schildkraut I, Aggarwal AK (1997) Structure of the multimodular endonuclease FokI bound to DNA. *Nature* **388:** 97-100

243.   Walker FO (2007) Huntington's disease. *Lancet* **369:** 218-228

244.   Wang J, Kim HH, Yuan X, Herrin DL (1997) Purification, biochemical characterization and protein-DNA interactions of the I-CreI endonuclease produced in Escherichia coli. *Nucleic Acids Res* **25:** 3767-3776

245.   Weber E, Gruetzner R, Werner S, Engler C, Marillonnet S (2011) Assembly of designer TAL effectors by Golden Gate cloning. *PLoS One* **6:** e19722

246.   Wolfe SA, Nekludova L, Pabo CO (2000) DNA recognition by Cys2His2 zinc finger proteins. *Annu Rev Biophys Biomol Struct* **29:** 183-212

247.   Wolff JA, Malone RW, Williams P, Chong W, Acsadi G, Jani A, Felgner PL (1990) Direct gene transfer into mouse muscle in vivo. *Science* **247:** 1465-1468

248.   Wright DA, Townsend JA, Winfrey RJ, Jr., Irwin PA, Rajagopal J, Lonosky PM, Hall BD, Jondle MD, Voytas DF (2005) High-frequency homologous recombination in plants mediated by zinc-finger nucleases. *Plant J* **44:** 693-705

249.   Xodo LE, Manzini G, Quadrifoglio F, van der Marel GA, van Boom JH (1991) Effect of 5-methylcytosine on the stability of triple-stranded DNA--a thermodynamic study. *Nucleic Acids Res* **19:** 5625-5631

250.   Zaremba M, Sasnauskas G, Urbanke C, Siksnys V (2005) Conversion of the tetrameric restriction endonuclease Bse634I into a dimer: oligomeric structure-stability-function correlations. *J Mol Biol* **348:** 459-478

251.   Zaremba M, Sasnauskas G, Urbanke C, Siksnys V (2006) Allosteric communication network in the tetrameric restriction endonuclease Bse634I. *J Mol Biol* **363:** 800-12

252.   Zaremba M, Urbanke C, Halford SE, Siksnys V (2004) Generation of the BfiI restriction endonuclease from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. *J Mol Biol* **336:** 81-92

253.   Zhang F, Cong L, Lodato S, Kosuri S, Church GM, Arlotta P (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol* **29:** 149-153

254.   Zhao L, Bonocora RP, Shub DA, Stoddard BL (2007) The restriction fold turns to the dark side: a bacterial homing endonuclease with a PD-(D/E)-XK motif. *Embo J* **26:** 2432-2442