Engineering LAGLIDADG homing endonucleases: notes, citations and suggestions

1. Target and enzyme choice.

The single most important step in generating an effective gene targeting nuclease from a LAGLIDADG homing endonuclease (termed an 'LHE', or 'meganuclease') is the initial choice of the target site/endonuclease combination. If you use the search strategies provided by LAHEDES against a genomic target region, you will be provided with a large list of potential starting combinations for engineering. You will probably start your assessment by asking questions about the genomic location of potential targets, such as:

a. Is the target in an exon, intron, or noncoding region? Is it in an early exon where a disruption is likely to lead to a true functional knockout? If gene correction is desired, is the target proximal to the disease-associated mutation(s) to be converted?

b. Is the target strongly conserved, if the LHE is to be used for conversion of a population of related allelic targets?

Then, the most important choice you can make is to identify a combination of target and LHE scaffold that is highly amenable to subsequent engineering for highly efficient and specific recognition and cleavage. The variables of greatest importance are:

a. How many substitutions are found in the desired DNA target site, relative to the sequence of the initial DNA target for the wild-type LHE? Where are those substitutions within the target, and how are they distributed?

b. How well characterized is the LHE that you would need to engineer? Will I be reliant upon a crystal structure or a homology model for engineering choices? Does it have a proven track record of successful engineering (such as I-CreI, I-MsoI, I-OnuI and I-AniI)?

Consider the following three potential combinations of enzymes and /target sites within the human MAO-B gene, each identified as a result of a target site "Identity Search" conducted using all available LHEs in the LAHEDES webserver:

I-CkaMI

Natural Target Score: 0.00311348

Gene	Position	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	Score
>gi|156630984|ref|NM_000898.4|	r 24/2611	A	A	t	c	A	a	C	t	T	T	T	A	t	t	t	t	T	A	A	T	T	A	3.11348e-11

I-OnuI

Natural Target Score: 0.00311348

Gene	Position	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	Score
>gi|156630984|ref|NM_000898.4|	r 2270/2611	g	g	T	C	C	A	C	a	T	A	T	T	t	A	A	C	C	T	T	T	T	g	3.11348e-08

I-PanMI

Natural Target Score: 0.00311348

Gene	Position	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7	8	9	10	11	Score
>gi|156630984|ref|NM_000898.4|	f 1971/2611	t	g	c	C	C	T	C	A	a	A	A	T	C	t	T	T	A	g	a	A	t	a	3.11348e-12

ABOVE: potential target sites identified in the human monoamine oxidase B (MAO-B) gene, when searching all the LHEs in the database using the '*Identity Search*' algorithm. Uppercase bases are those that are identical to the starting wild-type target site, whereas lowercase bases are those that differ. Blue bases are those that are the same at the crucial 'central 4' positions in the target.

Which is the best one to choose? Your considerations would include the following:

a. The best enzyme for engineering: I-OnuI. It has been thoroughly characterized biochemically and structurally (with a DNA-bound structure to about 2.3 Å resolution) and has been successfully used in past work to generate active gene-targeting enzymes (Takeuchi *et al.* (2011) *PNAS USA* **108** (32): 13077 - 82).

b. Fewest mismatches between the 'wild-type' target and the desired target: Engineering I-OnuI to recognize its best target in MAO-B would require alteration of its specificity at a total of 5 basepair positions (-11, -10, -4, +2 and +11). In contrast, I-CkaMI would require changing its specificity at 8 positions, and I-PanMI would require 9.

c. Best match of basepair identity across the 'central 4' positions. Generally, it is preferable that the target be matched across the entire 'central 4' basepairs at the center of the target (positions -2, -1, +1 and +2), because specificity at those positions is primarily influenced by DNA bending and corresponding indirect readout of DNA conformational preferences. We would recommend (i) that no more than one mismatch be considered across those positions, (ii) that the mismatches be limited to transitions (A ⇆ G or T ⇆ C) rather than transversions, and (iii) that the mismatches correspond to known tolerated substitutions by the wild-type enzyme (which can be estimated for a subset of the LHEs in the database by examining specificity profiles (PWMs) in the browser).

In the example above, I-CkaMI is matched at all four (blue) basepairs in the center of the target site, whereas I-OnuI and I-PanMI both must accommodate a +2C to +2T alteration. Fortunately, in both of

those latter cases the necessary substitutions is a transition. As well, examination of the I-OnuI PWM indicates that a 'C' at this position is tolerated as well (if not better) than a T.

d. Distribution of mismatches across the target: As a general rule, multiple mismatches in the desired basepair target at positions +/- 11, 10, and 9 can collectively be accommodated by selection of enzyme variants that are culled from randomized amino acid positions that are distributed across a single contact region spanning a surface loops between β -strands β1 and β2. As well, LHEs generally display elevated promiscuity at positions 11 and 10. Therefore, a combination of the fewest number of DNA basepair mismatches overall, with the majority of those mismatches located at the "outer 3" basepairs in each half-site, are most likely to yield relatively successful redesign efforts.

Again, examining the target sites above, one would predict that the I-OnuI target (with only 2 mismatches found in the 'inner' basepair positions, between -8 and +8) would be the best bet for successful targeting.

Based on the sum total of the criterion listed above, all of the three hits listed would be reasonable and relatively straightforward choices for redesign. In the absence of additional considerations regarding gene structure and sequence constraints, the I-OnuI target would clearly be the top choice.

2. Identification of contact modules for redesign and selection

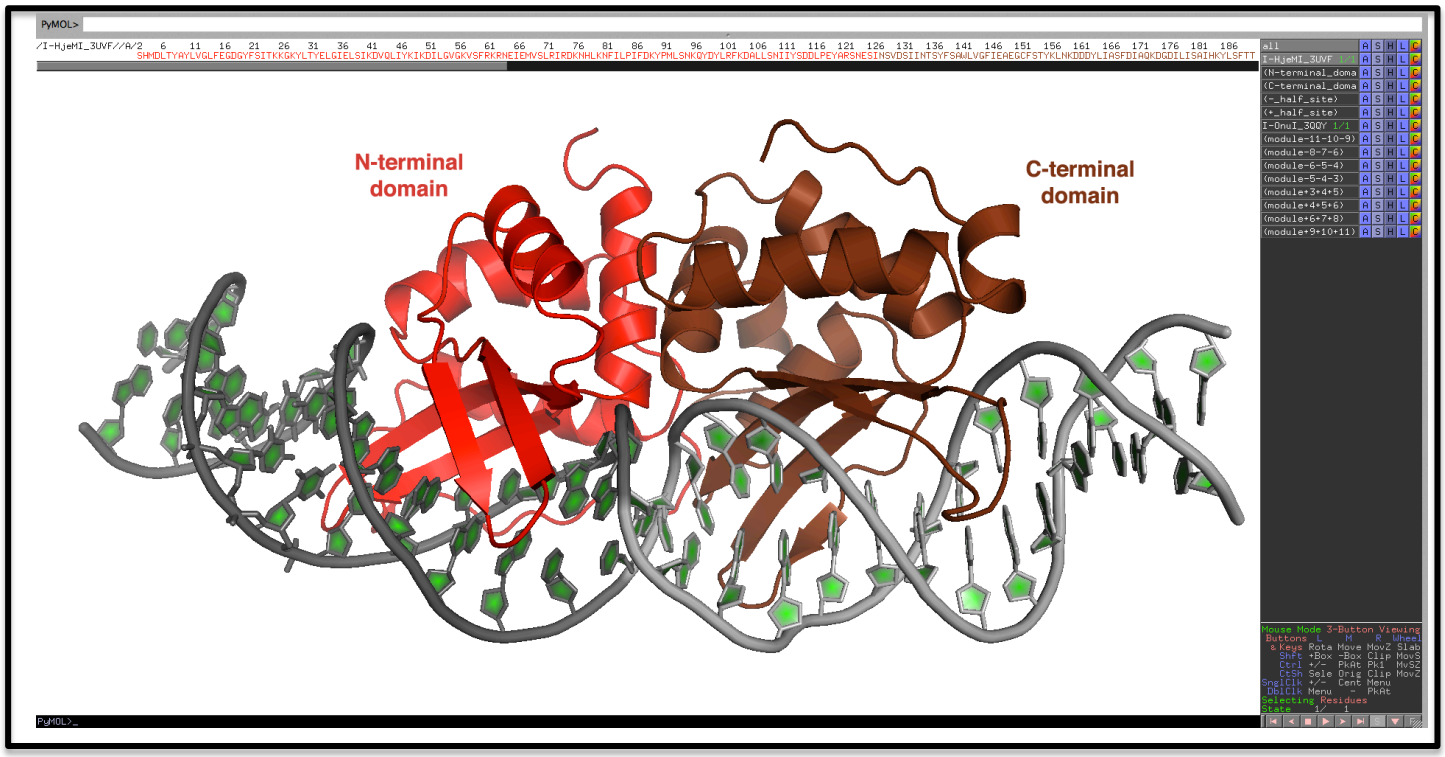
The structures of a large number of LHE-DNA complexes have been determined at high resolution. There have been many corresponding attempts to survey and catalogue the identity of contacts made in those complexes. These studies have repeatedly indicated that while certain DNA nucleotides display preferences for specific residue contact patterns (such as the use of arginine side chains to make direct contacts to guanine bases, and the use of glutamine or asparagine to make direct contacts to adenine bases), there is no simple, reliable "code" that can definitively predict individual protein-DNA interactions or the effect of single amino acid substitutions on activity, affinity and specificity. This is because the contributions of individual side-chain contacts to the DNA on either binding affinity or cleavage activity are highly dependent upon the surrounding structural and chemical context of the protein-DNA interface.

The significance of this point on LHE redesign is clear: structure-based randomization of 'modules' of amino acids involved in DNA recognition, combined with and the incorporation of individual critical mutations within the protein scaffold, is usually necessary to produce variants of wild-type LHEs that display the requisite target site specificity and activity for genome engineering applications.

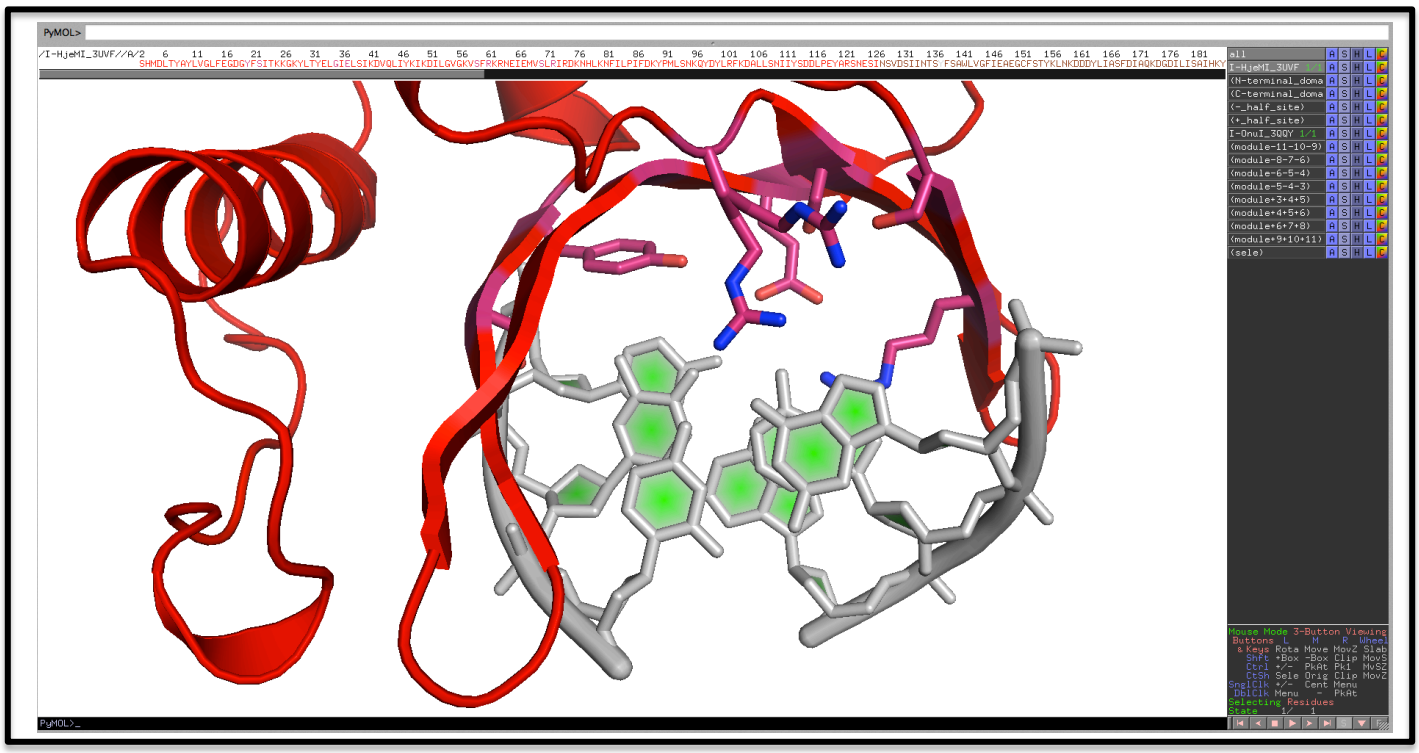
Once you have settled on a combination of an enzyme and a target site for redesign, it is time to settle upon an experimental strategy for LHE redesign and selection. The first step is to identify the amino acids that are most responsible for sequence specificity at those DNA target positions and must therefore be randomized for selection experiments.

LAHEDES provides visual assistance for this step. In the browser page for each enzyme, you will find the FASTA sequence for the protein (to aid in design of coding sequences) as well as links to either a crystal structure or a homology model for the enzyme. A 'Pymol' session file is also available; it provides the coordinates of the enzyme (either an actual crystal structure or a homology model) superimposed on I-OnuI or I-CreI bound to its DNA target (I-OnuI if the LHE is a monomer; I-CreI if the LHE is derived from a homodimer). When you open the session file (after downloading and installing Pymol, via http://www.pymol.org) you will see the ribbon diagram of the enzyme, with the N- and C-terminal domains colored red and chocolate. Eight separate groups of amino acids,

corresponding to our first approximation of the 'contact modules' for sets of DNA basepairs distributed across the target, are available to highlight and display using the selection list to the right.



The eight protein modules that we usually choose to consider for engineering of LHE specificity cover a corresponding series of eight separate three-basepair DNA modules: positions -11,-10,-9 ; positions -8,-7,-6; positions-6,-5,-4 ; positions -5,-4,-3; positions +3,+4,+5; positions +4,+5,+6; positions +6,+7,+8; and positions +9,+10,+11. By clicking on one of the modular selections, displaying its side chains and then zooming in, you can examine the contacts made between those positions ion the protein and the corresponding DNA:



ABOVE: The 'contact module' in I-HjeMI for basepairs -6,-5,-4, which consists of Y20, S22, G35, E37, S59,R61, S70 and R72. A combination of designed and/or selected mutations within these residues will likely yield a variant of this LHE that can recognize a DNA target with basepair substitutions at one or more of those three positions.

Not all of the side chains in a contact module that is illustrated in an enzyme's structure or Pymol session file in LAHEDES will necessarily be in direct contact with DNA bases. Some are engaged in networked contacts with contact residues and others are in contact with DNA-bound water molecules. We recommend close examination of contact modules, to make sure you agree with the choice of amino acids to be altered during engineering, before proceeding further.

3. Structure-based engineering substitutions

High resolution crystal structures of the wild-type protein-DNA complexes have been used in various contexts either to enable the targeted redesign of homing endonuclease-DNA contacts at individual residues (thus bypassing selection approaches altogether), or to facilitate more efficient mutational screening of enzyme libraries (by identifying, and thus greatly reducing, the number of protein residues to be randomized). We do not recommend strategies in which mutations at individual contact residues are simply combined in the absence of selections for optimal activity in living cells. However, there are a number of studies that have identified individual, specificity-shifting mutations in several LHE scaffolds, that might be included within such selection experiments (see below for examples). Many of the mutations described in those studies have been entered into the LAHEDES database and are available under 'specificity-changing mutations' in the corresponding entry.

I-CreI:

Seligman, L. M. et al. (2002) "Mutations altering the cleavage specificity of a homing endonuclease" Nucleic Acids Research 30 (17): 3870 - 3879.

Sussman, D., Chadsey, M., Monnat, R. J., **Stoddard, B. L.** and Seligman, L.M. (2004) "Isolation and characterization of new homing endonuclease specificities at individual target site positions" *J. Mol. Biol.* **342** (1): 34 - 41.

Rosen, L.E. et al. (2006) "Homing endonuclease I_CreI derivatives with novel DNA target specificities" Nucleic Acids Research 34 (17): 4791 - 4800.

Ulge, U. Y., Baker, D. A. and Monnat, R. J. Jr. (2011) "Comprehensive computaitonal design of mCreI homing endonuclease cleavage specificity for genome engineering" Nucleic Acids Research 39 (10): 4330 - 4339.

I-MsoI:

Ashworth, J., Havranek, J., Duarte, C., Sussman, D., Monnat, R. J. Jr., **Stoddard, B. L.**, and Baker, D. (2006) "Computational redesign of endonuclease DNA binding and cleavage specificity " *Nature* **441**: 656 - 659.

Ashworth, J., Taylor, G., K., Havranek, J. J., Quadri, S. A., **Stoddard, B. L.** and Baker, D. (2010) "Computational reprogramming of homing endonuclease specificity at multiple adjacent base pairs" *Nuc. Acids Res.* 38 (16) 5601 - 5608.

I-AniI:

Thyme, S., Takeuchi, R., Jarjour, J., Scharenberg, A., **Stoddard, B. L.** and Baker, D. (2009) "Exploitation of homing endonuclease binding energy for catalysis and design" *Nature* **461**: 1300 - 1304.

4. Engineering methods and strategies

A. Modular randomization of the LHE DNA binding region

The core methodology for the creation of gene-targeting reagents from LHEs is the introduction of mutations into one more positions on the protein, located within or near the DNA interface. In certain fortunate situations, a single isolated basepair substitution may be accommodated with a previously identified amino acid substitution. However, up to to three sequential DNA substitutions can and must be accommodated via the identification of an enzyme variant harboring multiple substitutions at anywhere from 5 to 10 amino acids.

Obviously this requires the randomization of a physical cluster of amino acids, often presented across discontinuous regions of the LHE protein, followed by high-throughput selection or screens for desired cleavage activities. There are a large number of protocols available for saturation mutagenesis using PCR amplification with degenerate oligonucleotide pools, most of which can be tailored for randomization of LHE DNA binding regions. Recent methodological manuscripts that we recommend, for the development of your own strategy, are:

Chonopoulou, E. G. and Labrou, N. E. (2011) "Site-saturation mutagenesis: a powerful tool for structure-based design of combinatorial mutation libraries" *Current Protocols in Protein Science* 26.6.1 - 26.6.10.

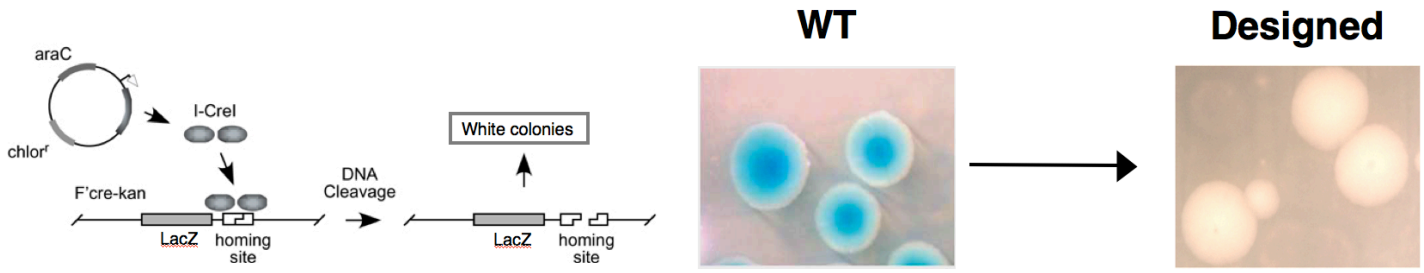
Tang, L. et al. (2011) "Construction of "small-intelligent" focused mutagenesis libraries using well-designed combinatorial degenerate primers" *BioTechniques* 52 (3): 149 - 155.

We also have relied heavily upon the *Agilent* "GeneMorph" kits and instructions for regional and ORF-wide mutagenesis: http://www.chem.agilent.com/Library/usermanuals/Public/200552.pdf

B. Screening and selection strategies

At least five different strategies have been described in the literature for the selection of homing endonucleases, as well as one additional strategy (IVC) that has recently been adapted by our lab for homing endonucleases. Below we provide a short illustrated summary of each, with relevant citations.

1. A bacterial selection strategy based on the cleavage and elimination of a reporter gene resulting in cells being converted from lac + to lac-. This method allows selection of desired activities based on a simple blue-white colony phenotype screen. Undesirable activity (for example, cleavage of the original wild-type site) can be suppressed through a secondary 'negative' screen for elimination of an essential reporter (such as an antibiotic resistance marker). Using this method, endonuclease mutants with single or double amino acid substitutions, at positions predicted to make base-specific DNA contacts, are assayed against appropriate DNA target site mutants.



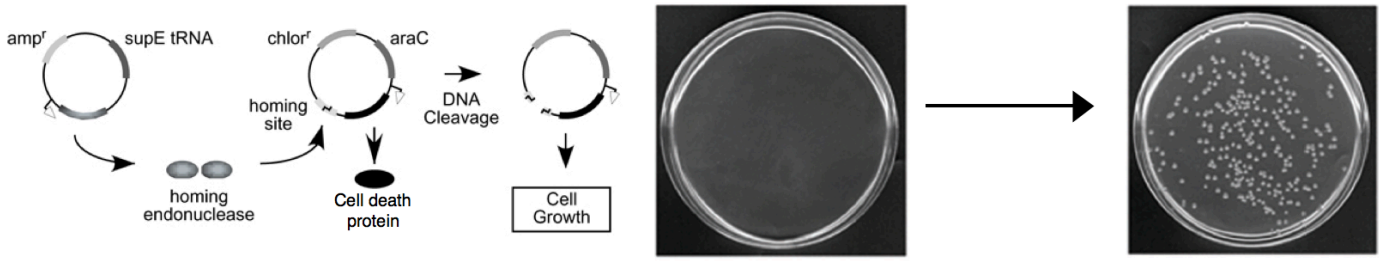
Above: Screen for homing endonuclease based on cleavage and gene elimination (Seligman et al. 2002; 2004). Plasmid-encoded I-CreI derivatives that bind and cleave an I-CreI homing site located on an F' lead to its elimination. Concomitant loss of an adjacent kanamycin antibiotic marker yields kanamycin-sensitive cells.

Seligman, L. M. et al. (2002) "Mutations altering the cleavage specificity of a homing endonuclease" Nucleic Acids Research 30 (17): 3870 - 3879.

Sussman, D., Chadsey, M., Monnat, R. J., **Stoddard, B. L.** and Seligman, L.M. (2004) "Isolation and characterization of new homing endonuclease specificities at individual target site positions" *J. Mol. Biol.* **342** (1): 34 - 41.

Rosen, L.E. et al. (2006) "Homing endonuclease I_CreI derivatives with novel DNA target specificities" Nucleic Acids Research 34 (17): 4791 - 4800.

2. A 2-plasmid bacterial selection strategy in which elimination of a toxic gene product via homing endonuclease cleavage within the toxin expression vector results in cell survival. Two versions of this screen have been described: the first utilizing the nonspecific ribonuclease barnase (Gruen et al., 2002); and the latter describing a more easily controlled system based on the 'control of cell death B' (CcdB) protein (Doyon et al., 2006). In the latter screening strategy, CcdB expression leads to very low rates of background survival, without requiring additional gene expression control elements such as those required to use the more toxic barnase protein. As with the bacterial selection described in the preceding paragraph, the CcdB system can be tailored as a positive selection for cleavage of a desired target, or as a negative selection, to strongly disfavor the recognition of a non-cognate DNA target site sequence. This system, when optimized, can lead to nearly 100% survival of cells expressing an active homing endonuclease, against a background survival of < 1 in 5 x 10^4 with an inactive enzyme. This method has been used as the sole strategy for the creation of a highly active gene-targeting enzyme that was used to knockout the endogenous human MAO-B gene (Takeuchi et al. 2011).



Above: Selection for homing endonuclease activity based on cleavage and elimination of a toxic gene product. Co-existence of two plasmids kills bacterial cells because one expresses a toxic gene product. However, when a homing endonuclease is expressed cleaves the homing site on the toxin expression plasmid, the cells survive.

Gruen, M., Chang, K., Serbanescu, I., Liu, D. R. (2002) "An in vivo selection system for homing endonuclease activity" Nucleic Acids Research. 30 (7): e29.

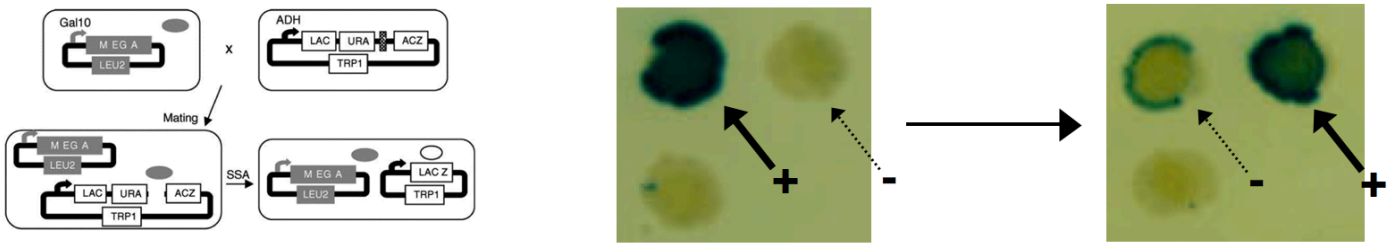
Doyon, J. B., Pattanayak, V., Meyer, C. B., Liu, D. R. (2006) "Directed evolution and substrate specificity profile of homing endonuclease I-SceI" J. Am. Chem. Soc. 128 (7): 2477 - 2484.

and an example of the use of this methodology for the complete redesign of an LHE to cleave an endogenous chromosomal target in human cells:

Takeuchi, R., et al. (2011) "Tapping natural reservoirs of homing endonucleases for targeted gene modification" PNAS USA 108 (32): 13077 - 13082.

3. A Yeast-based Recombination Reporter. A commercial group (Cellectis, Inc.) has developed a powerful eukaryotic assay system that reports on the generation of double-strand break-induced homologous recombination, rather than only DNA cleavage. In this assay, the function of a gene required for growth or for another easily scored phenotype is restored through the action of a homing endonuclease. Prior to expression and action of the HE, the gene sequence is interrupted with an insert containing a desired HE cut site flanked by two direct repeats. In the most recently described version of this screen, the endonuclease expression construct and the 'reporter' construct are located in separate yeast strains, allowing the investigator to introduce the HE (or a library of HE variants) to a targets site by mating. The mating of the two required yeast strains can be automated, and thus done as a high-throughput assay. Thus the same library of endonuclease variants can be efficiently screened, against multiple DNA target site variants, in parallel or rapid sequential experiments.

Using this method, the DNA recognition specificity of I-CreI LHE has been thoroughly analyzed. Small endonuclease mutant libraries resulting from the randomization of 2 to 4 amino acids (corresponding to the 'nearest protein neighbors' of individual DNA basepairs) were individually screened against all potential variant cognate sequences. This analysis identified individual mutations in the I-CreI protein scaffold that were associated with shifts in specificity at individual target site base pair positions. From these studies, hundreds of endonuclease mutants with altered specificities were identified, catalogued and archived. Many of these variants displayed cleavage activities and levels of site discrimination that were at least equivalent to the wild-type endonuclease, thus providing a starting point for more ambitious endonuclease redesign to physiological targets in eukaryotic, mammalian and even human genomes.



Above: Yeast gene homologous recombination and gene complementation screen (Arnould et al. 2006). A strain harboring the expression vector encoding endonuclease variants is mated with a strain harboring a reporter plasmid, which contains a LacZ gene interrupted with the desired endonuclease target, flanked by two direct repeats. Endonuclease cleavage at this site induces restoration of the functional LacZ gene by single-strand annealing between the two repeats.

Using the approach summarized above, derivatives of the I-CreI LHE have been generated that display sequence-specific cleavage and recombination activity against the human *RAG1* gene (which is associated with severe combined immunodeficiency disease (or SCID) phenotypes). I-CreI variants directed at the human *XPC* (Xeroderma Pigmentosum complementation group C) gene were also generated. XPC, when mutant, confers an extreme UV or sun- sensitivity phenotype together with a predisposition to sunlight-induced skin cancer. Both of these human disease genes are candidates for corrective genetic therapies. Furthermore, in the latter experiments the modified I-CreI derivatives cleaving sequences from the XPC gene were found to induce a high level of gene targeting, similar to levels observed with the wild-type I-CreI or I-SceI scaffolds. This was the first time an engineered homing endonuclease has been used in mammalian cells to target and modify a chromosomal target locus.

Development of the assay and screen:

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

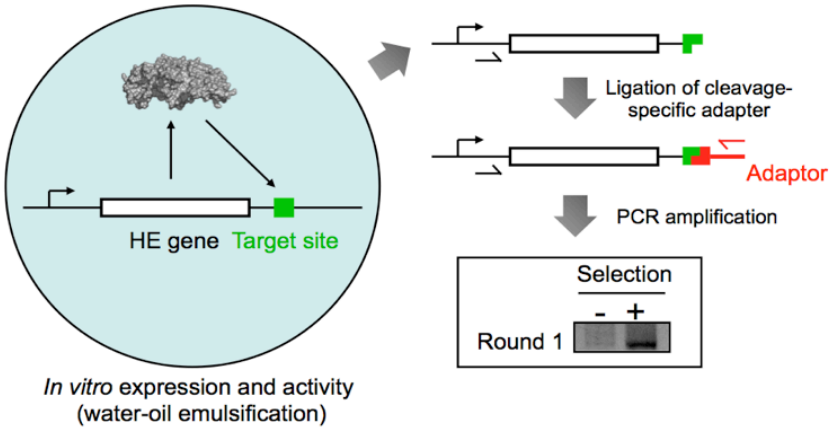
Arnould, S., Chames, P., Perez, C., Lacroix, E., Duclert, A., Epinat, J. C., Stricher, F., Petit, A. S., Patin, A., Guillier, S., *et al.* (2006). "Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets" J Mol Biol *355*, 443-458.

Application to genomic targets:

Arnould, S., Perez, C., Cabaniols, J.-P., Smith, J., Gouble, A., Grizot, S., Epinat, J.-C., Duclert, A., Duchateau, P., and 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.

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

4. In vitro compartmentalization. In this protocol (which was first described for the selection of restriction endonucleases and methyltransferases; see references below) libraries of randomized enzymes are encoded, expressed and screened in oil-emulsified water droplets. IVC allows library sizes of up to 10^{10} endonuclease variants to be panned in each selection step. Rounds of filtering selection at elevated temperatures can effectively filter out destabilized constructs prior to final stages of optimization, and the substrate concentration in the droplets can be precisely controlled at subnanomolar levels to maximize DNA binding affinity. The generation of an active meganuclease is a function of simple ligation and PCR steps that do not require time-consuming cell growth (which is a staple of all other endonuclease engineering methods).

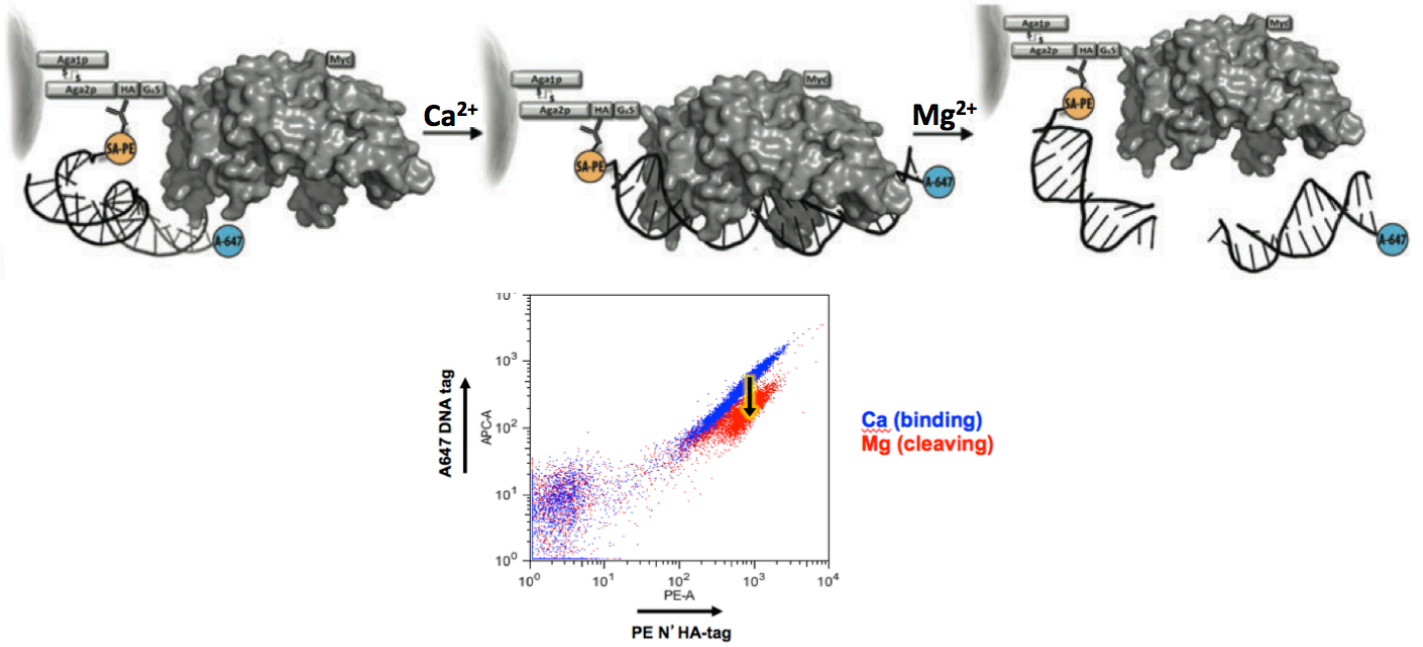


Left: In vitro compartmentalization allows large libraries (up to 10^{11}) of individual meganucleases to be expressed under conditions where (i) the coding sequence of the endonuclease is directly coupled to activity against a desired DNA target site; (ii) elevated temperatures can be employed to provide pressure minimizing the amplification of destabilized enzymes; (iii) the substrate concentration can be precisely defined and kept sufficiently low (0.2 nM is typical) to provide pressure for high affinity DNA binding, and (iv) amplification and recovery of active clones is rapidly achieved via ligation and PCR steps after the emulsification is harvested and broken.

Miller, et al. (2006) "Directed evolution by in vitro compartmentalization" Nature Methods 3 (7): 561 - 570.

Zheng and Roberts (2007) "Selection of restriction endonucleases using artificial cells". *Nucleic Acids Res*. **35** (11) e83.

5. Surface display. We have also developed a system in which yeast surface display of an LHE is used for analyses and selections for the binding and/or cleavage of a desired DNA target. Quantitative flow cytometry using oligonucleotide substrates facilitated a complete profiling and/or selection of specificity, both for DNA-binding and cleavage, with single base pair resolution.



Above: Schematic representation of a flow cytometric DNA cleavage assay. The homing endonuclease of interest is expressed on the surface of yeast through fusion with the yeast Aga2P protein. Aga2P forms a disulfide linkage with Aga1P, which is co-induced in the EBY100 strain upon Galactose induction. An N-terminal hemagglutinin (HA) tag enables cis-tethering of the double-stranded DNA substrate to the enzyme: a streptavidin-PE 'bridge' connects the bound anti-HA-biotin with biotinylated and A647-labeled DNA substrate. This tethered target substrate results in the characteristic co-linear PE and A647 fluorescence profile observed on the flow cytometer. With the addition of calcium, the enzyme may bind the DNA substrate, but it cannot cleave it; with the addition of magnesium, the enzyme is able to bind and cleave the substrate provided a productive interaction is formed. Cleavage of the tethered substrate leads to a loss of fluorescence in the A647 channel. This loss in A647 fluorescence can be used to quantify cleavage by gating on a population of yeast normalized for their PE signal and comparing the median A647 fluorescence intensities of the corresponding Ca^{2+} and Mg^{2+} samples.

Jarjour, J. et al. (2009) "High-resolution profiling of homing endonuclease binding and catalytic specificity using yeast surface display" *Nuc. Acids Res.* **37** (20): 6871 - 6880.

Baxter, S. et al. (2012) " Flow cytometric assays for interrogating LAGLIDADG homing endonuclease DNA binding and cleavage properties" *Methods* in press.

C. Parallel engineering and fusion of N- and C-terminal domains.

The modular structure of a LAGLIDADG homing endonuclease indicates that in many cases, the N- and C-terminal domains of these enzymes (which are largely responsible for recognition and interaction with the corresponding left and right DNA half-sites of the desired target) can be individually engineered in parallel experiments, and then combined into the final desired gene targeting reagent. As well, in some cases N- and C-terminal domains from different parental LHEs can also be combined using similar methods.f

The success rate of LHE domain recombination is under active investigation, but is clearly less than 100%, even for combinations of engineered domains from the same parental enzyme. The reason for this is not clear, but appears to be related to the bending of the full-length DNA target by the LHE during the cleavage reaction, which may in some cases result in noncompatible sets of redesign solutions for the individual domains. Therefore, we recommend selection and redesign experiments be carried out on the individual domains in parallel, but that investigators be prepared to extend these protocols across the protein interface, in a single series of sequential selection experiments, rather

than counting on the ability to put N- and C-terminal domains together at the end of the engineering protocol.

Baxter, S. et al. (2012) "Engineering domain fusion chimeras from I-OnuI family LAGLIDADG homing endonucleases" Nuc. Acids Res. epub ahead of print (PubMed ID 22684507).