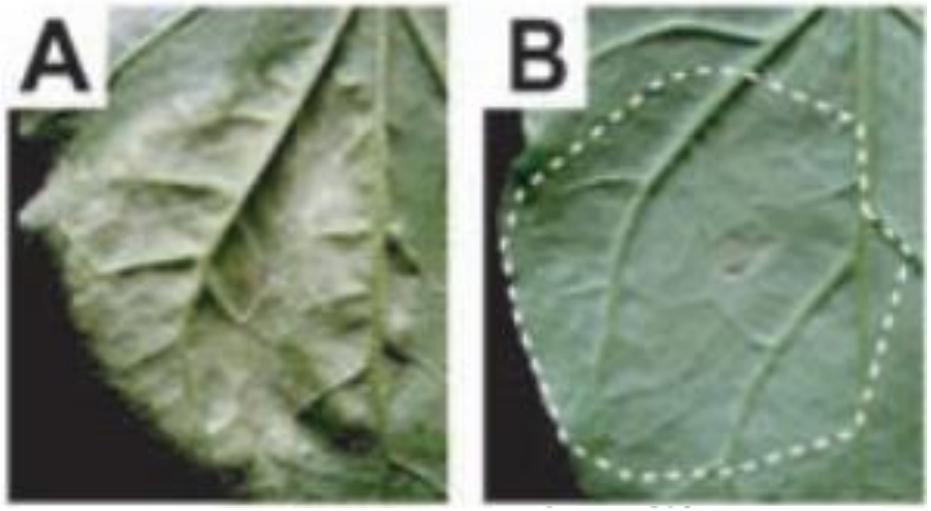
Genome Editing and Engineering

Course No: BT-637

LECTURE-15

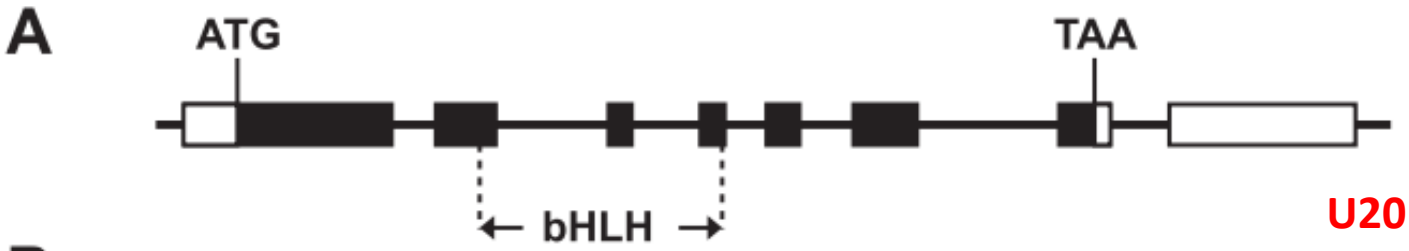
Dr. Kusum K. Singh
Department of Biosciences and Bioengineering
Indian Institute of Technology Guwahati

Novel Type of DNA Binding Domain



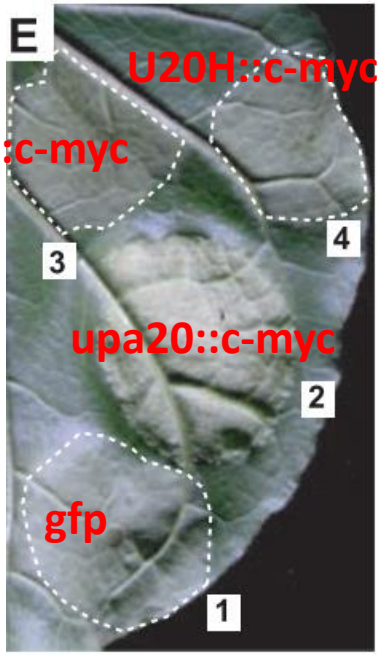
- **Upa20 alone induces hypertrophy**

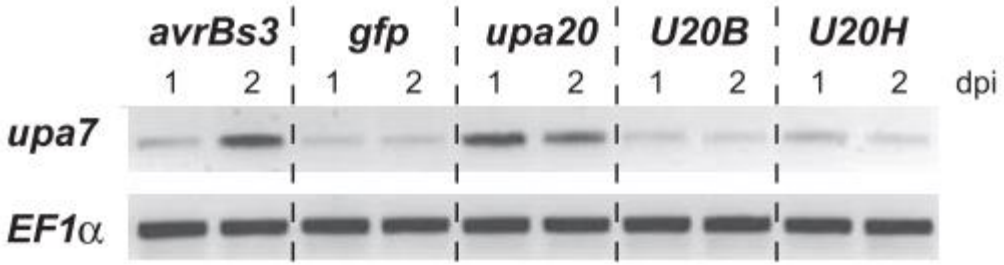
Novel Type of DNA Binding Domain



B

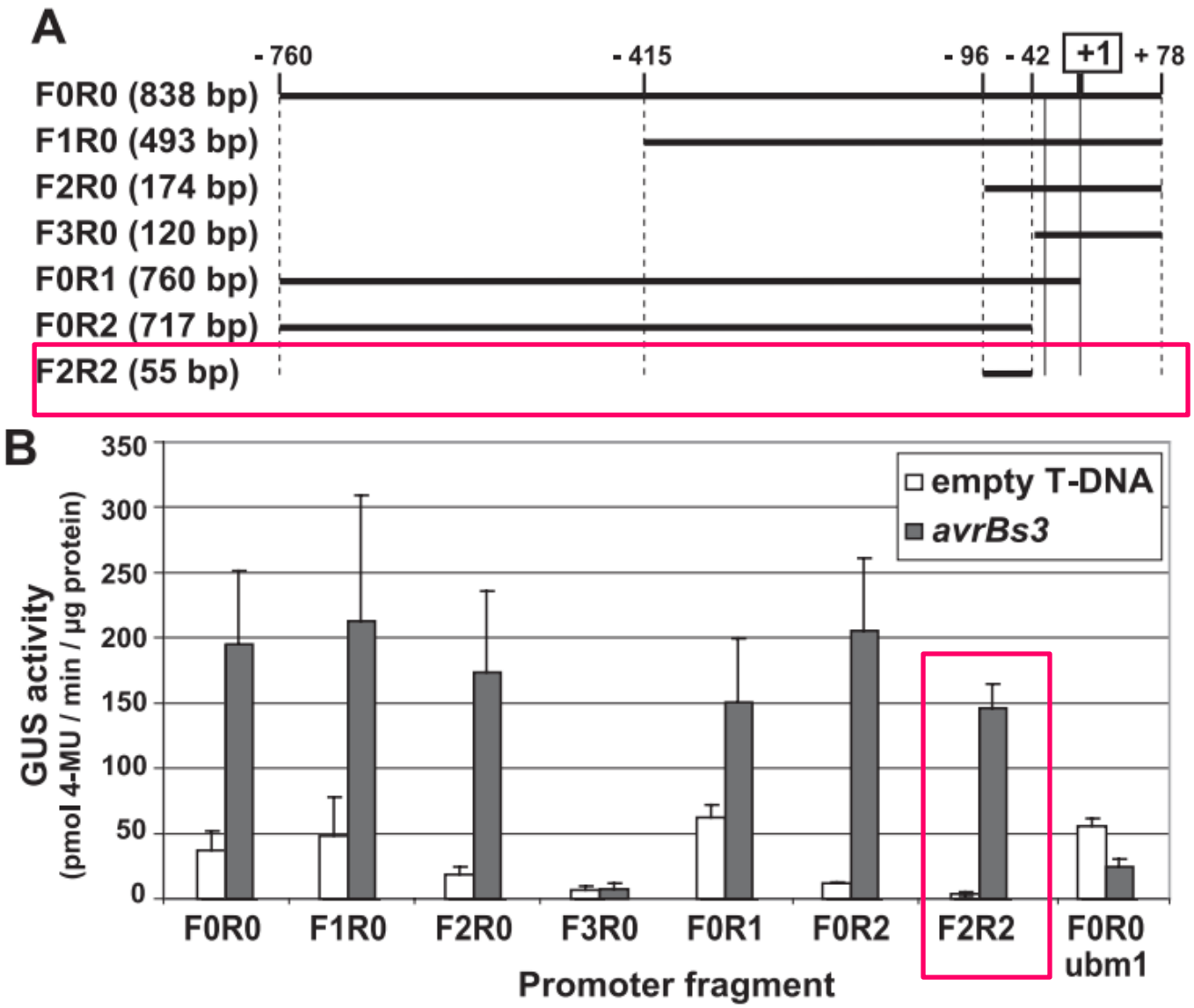
```
Upa20   RRGQATDSHSLAERVRREKISERMKMLQSLVPGCDKVTGKALMLDEIINYVQSLQNQVE
U20B    ............AA.....................................................
U20H    ..........................................AA..............
```



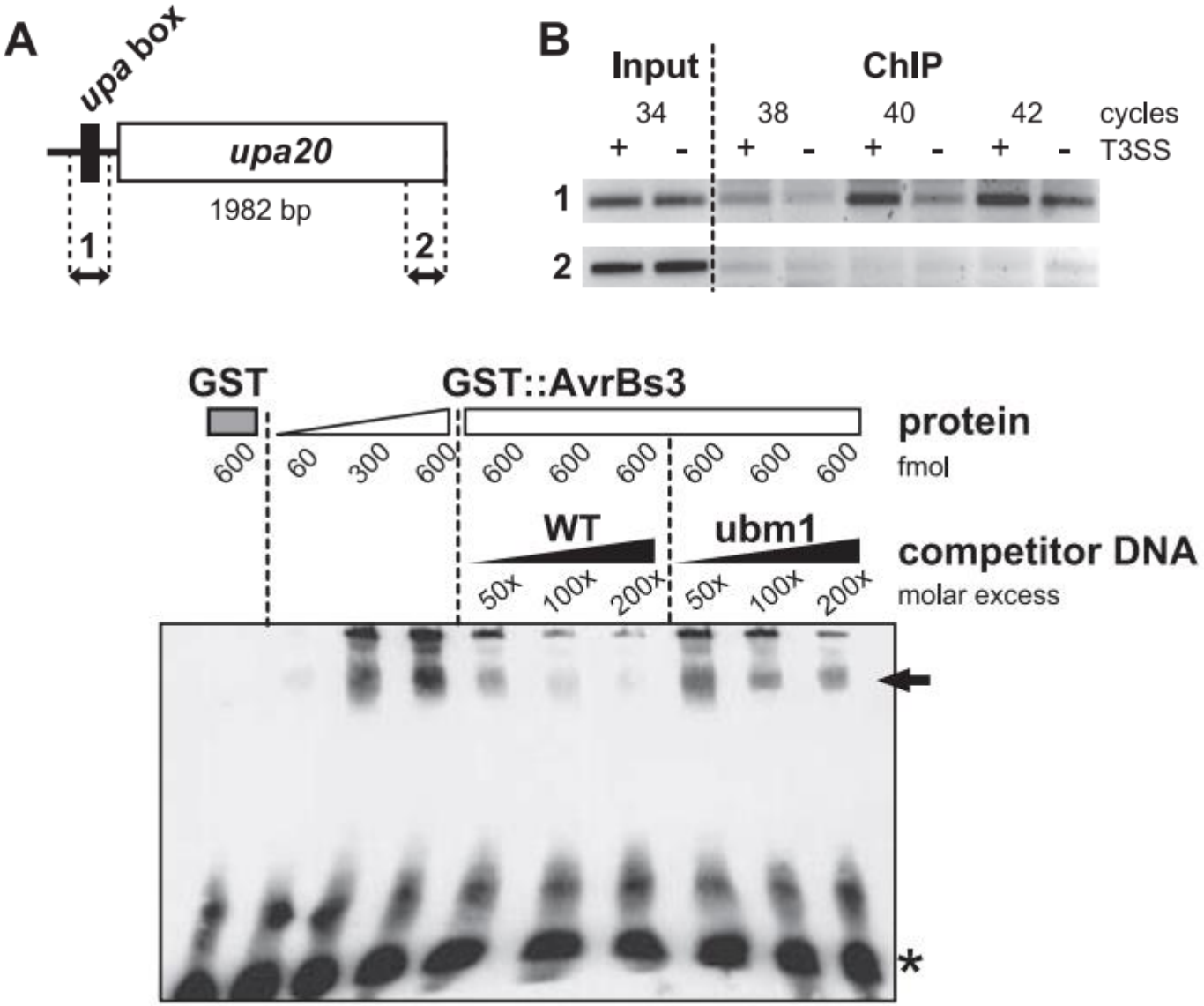


- Visible hypertrophy caused by AvrBs3 targeting *upa20*

Novel Type of DNA Binding Domain



Novel Type of DNA Binding Domain



Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors

Jens Boch,* Heidi Scholze, Sebastian Schornack,† Angelika Landgraf, Simone Hahn, Sabine Kay, Thomas Lahaye, Anja Nickstadt,‡ Ulla Bonas

The pathogenicity of many bacteria depends on the injection of effector proteins via type III secretion into eukaryotic cells in order to manipulate cellular processes. TAL (transcription activator–like) effectors from plant pathogenic *Xanthomonas* are important virulence factors that act as transcriptional activators in the plant cell nucleus, where they directly bind to DNA via a central domain of tandem repeats. Here, we show how target DNA specificity of TAL effectors is encoded. Two hypervariable amino acid residues in each repeat recognize one base pair in the target DNA. Recognition sequences of TAL effectors were predicted and experimentally confirmed. The modular protein architecture enabled the construction of artificial effectors with new specificities. Our study describes the functionality of a distinct type of DNA binding domain and allows the design of DNA binding domains for biotechnology.

Phytopathogenic bacteria of the genus *Xanthomonas* cause severe diseases on many crop plants. Pathogenicity relies on the translocation of effector proteins into the plant cell cytoplasm via the type III secretion system (*1*–*5*). Members of the large transcription activator-like (TAL) effector family are key virulence factors of *Xanthomonas* (*4*–*7*) and reprogram host cells by mimicking eukaryotic transcription factors (*8*–*13*). TAL effector–mediated gene induction leads to plant developmental changes [for example, cell divisions and cell enlargement such as citrus canker and hypertrophy (*4*)], thus contributing to disease symptoms. Although a number of plant targets, including susceptibility genes, have been identified (*8*–*10*, *12*–*14*), the targets of most TAL effectors

peats in a TAL effector determine its specific activity (*17*, *18*). The type member of this effector family, AvrBs3 from *Xanthomonas campestris* pv. *vesicatoria*, contains 17.5 repeats and induces expression of *UPA* (upregulated by AvrBs3) genes, including the *Bs3* resistance gene in pepper plants (*9*, *10*, *14*, *19*). The repeats of AvrBs3 are essential for DNA binding of AvrBs3 and represent a distinct type of DNA binding domain (*9*). How this domain contacts DNA and what determines specificity has remained enigmatic so far.

A model for sequence specificity. The fact that AvrBs3 directly binds to the *UPA* box, a promoter element in induced target genes (*9*, *10*), prompted us to investigate the basis for DNA-sequence specificity. The repeat region of AvrBs3 consists of 34 amino acid repeat units that are nearly identical; however, amino acids 12 and 13 are hypervariable (Fig. 1A) (*11*). The most C-terminal repeat of AvrBs3 shows a sequence similarity to other repeats only in its first 20 amino acids and is therefore referred to as a half repeat. The repeats can be classified into different repeat types on the basis of their hypervariable 12th and 13th amino acids (Fig. 1B). Because the size of the *UPA* box [18 base pairs (bp) (*20*)

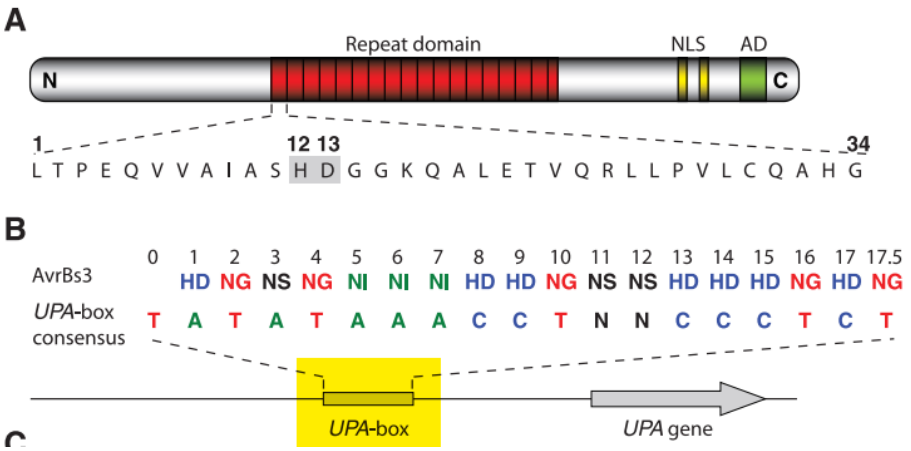


Fig. 1. Model for DNA-target specificity of TAL effectors. (**A**) TAL effectors contain central tandem repeats, NLSs, and an AD. Shown is the amino acid sequence of the first repeat of AvrBs3. Hypervariable amino acids 12 and 13 are shaded in gray. (**B**) Hypervariable amino acids at position 12 and 13 of the 17.5

A Simple Cipher Governs DNA Recognition by TAL Effectors

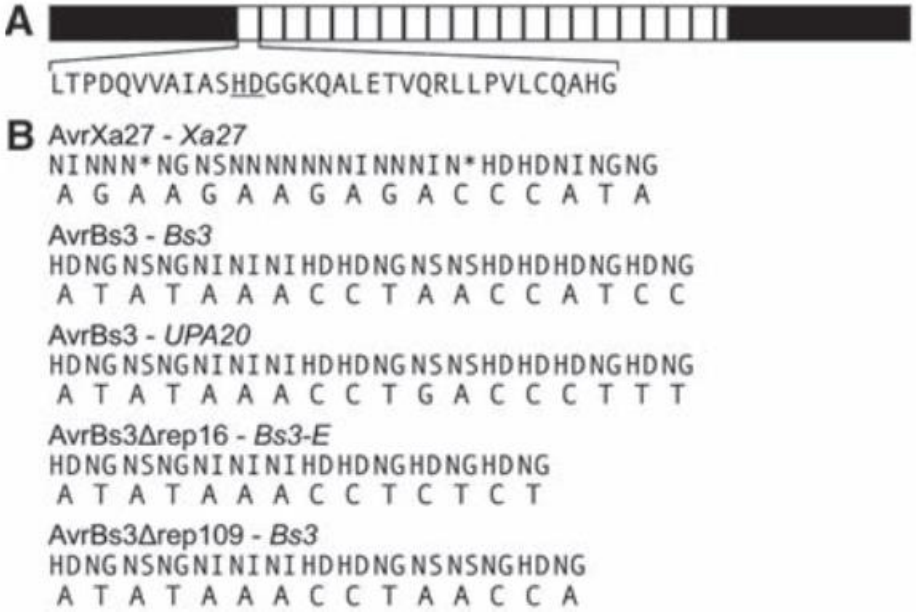
Matthew J. Moscou and Adam J. Bogdanove*

TAL (transcription activator–like) effectors of plant pathogenic bacteria in the genus *Xanthomonas* contribute to disease or trigger defense by binding host DNA and activating effector-

the 54–base pair (bp) *UPA20* promoter fragment that is sufficient and necessary for activation, and it coincided with the *UPA* box common to genes directly activated by AvrBs3 (*3*). For effectors PthXo1

all rice pror
TAL effecto
for which t
during infec
(http://PLE)
T precedes
scored better
stitute a strik

There is
associations
ones may pr
effects. Analy



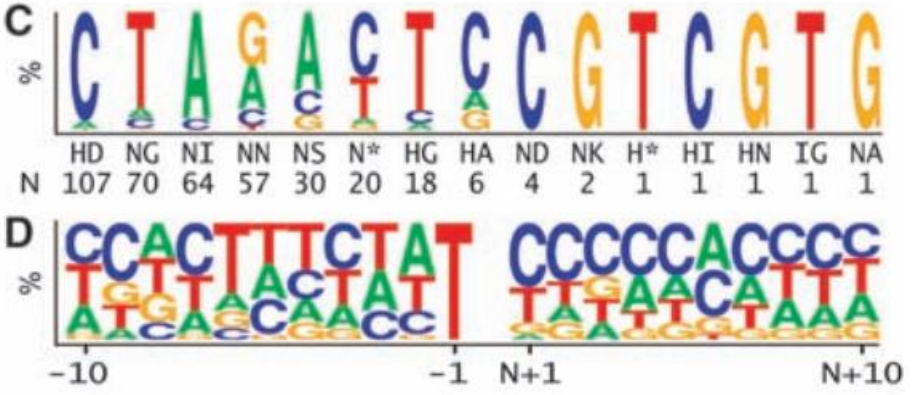
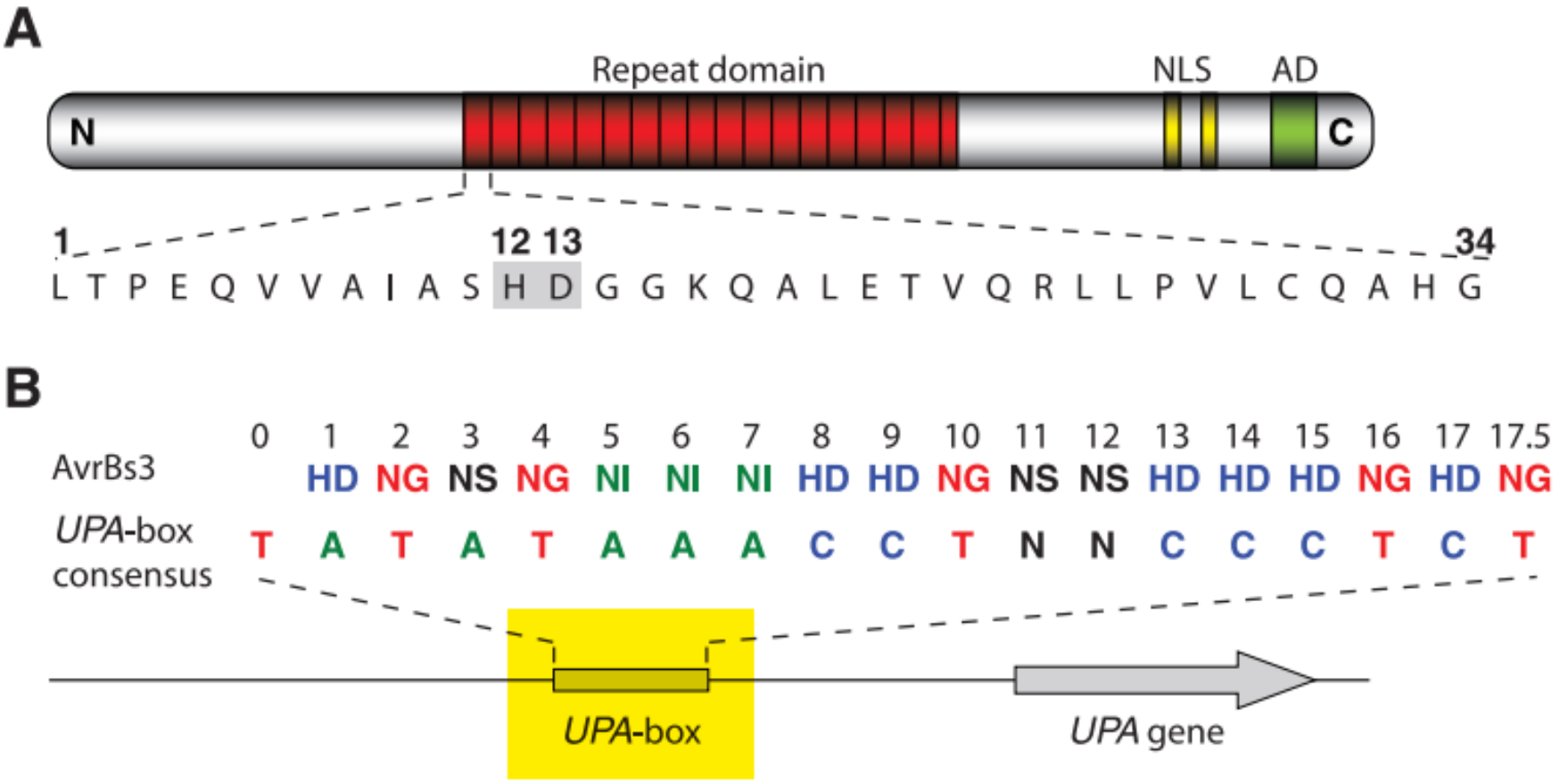
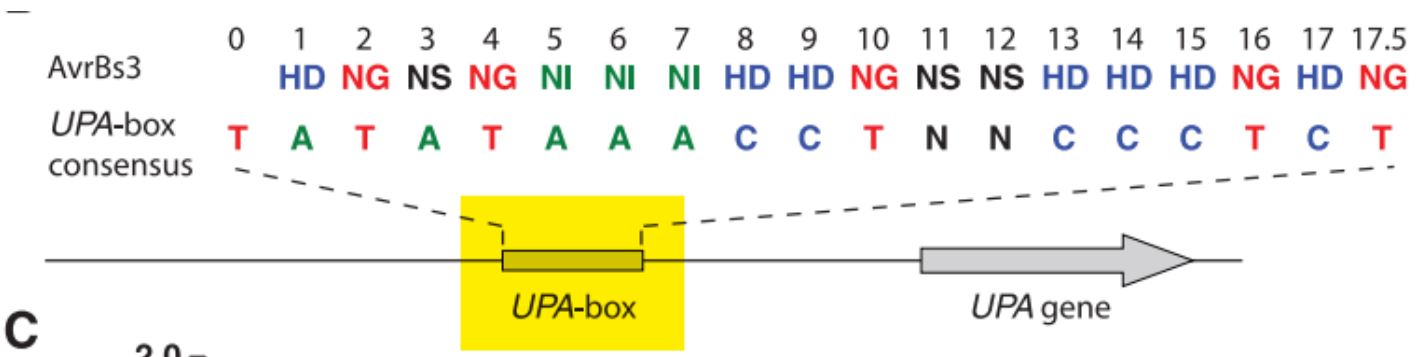


Fig. 1. The TAL effector–DNA recognition cipher. (**A**) A generic TAL effector showing the repeat region (open boxes) and a representative repeat sequence with the RVD underlined. (**B**) Best pattern matches (low-entropy alignments) for

Novel Type of DNA Binding Domain



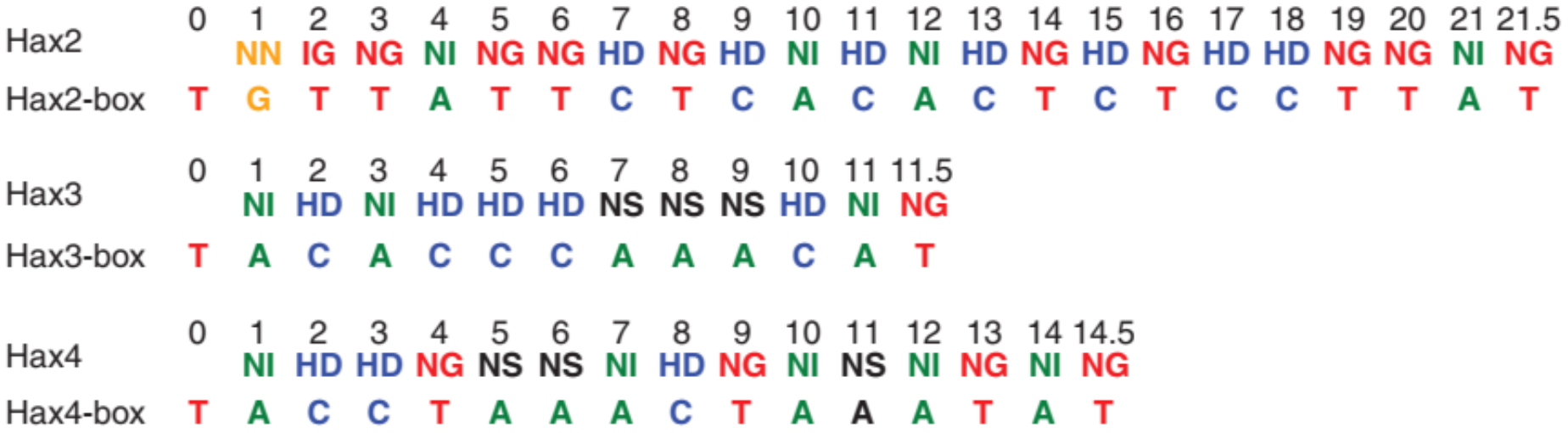
Novel Type of DNA Binding Domain



C

	HD	NG	NI	NS	NN	N*	HG	H*	IG
nucleotide frequency A	19	7	**55**	**20**	**7**	1	0	0	0
C	**69**	12	5	10	2	**7**	**2**	0	0
G	0	0	0	4	**7**	0	0	0	0
T	5	**50**	0	3	1	1	**2**	**1**	**1**

Novel Type of DNA Binding Domain

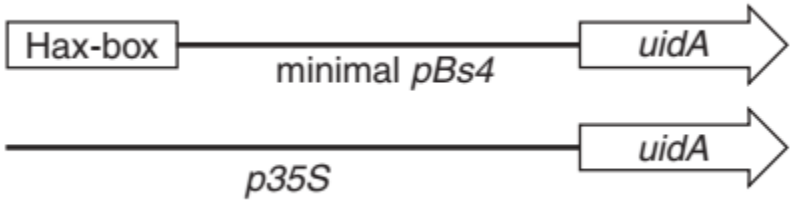


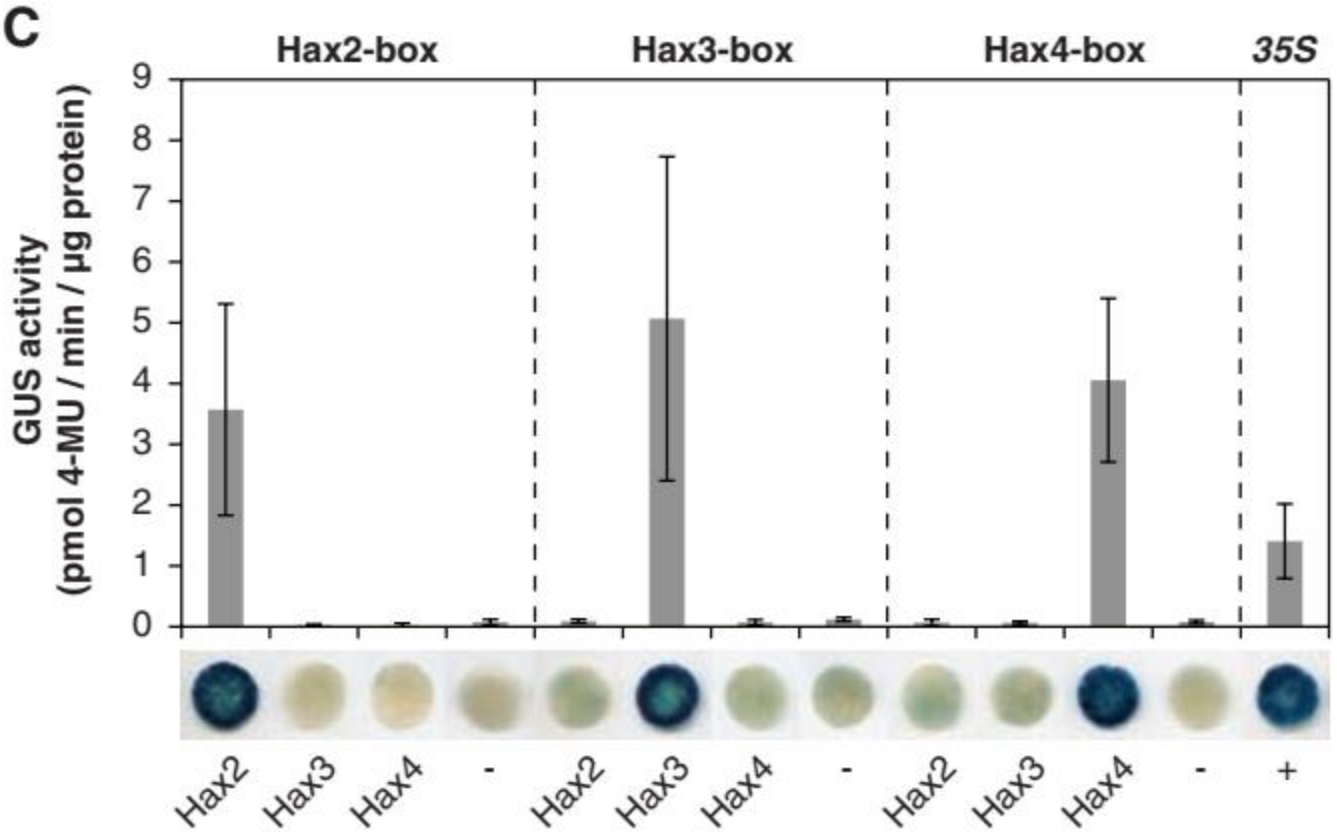
Novel Type of DNA Binding Domain

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	21.5
Hax2		NN	IG	NG	NI	NG	NG	HD	NG	HD	NI	HD	NI	HD	NG	HD	NG	HD	HD	NG	NG	NI	NG
Hax2-box	T	G	T	T	A	T	T	C	T	C	A	C	A	C	T	C	T	C	C	T	T	A	T

	0	1	2	3	4	5	6	7	8	9	10	11	11.5
Hax3		NI	HD	NI	HD	HD	HD	NS	NS	NS	HD	NI	NG
Hax3-box	T	A	C	A	C	C	C	A	A	A	C	A	T

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	14.5
Hax4		NI	HD	HD	NG	NS	NS	NI	HD	NG	NI	NS	NI	NG	NI	NG
Hax4-box	T	A	C	C	T	A	A	A	C	T	A	A	A	T	A	T

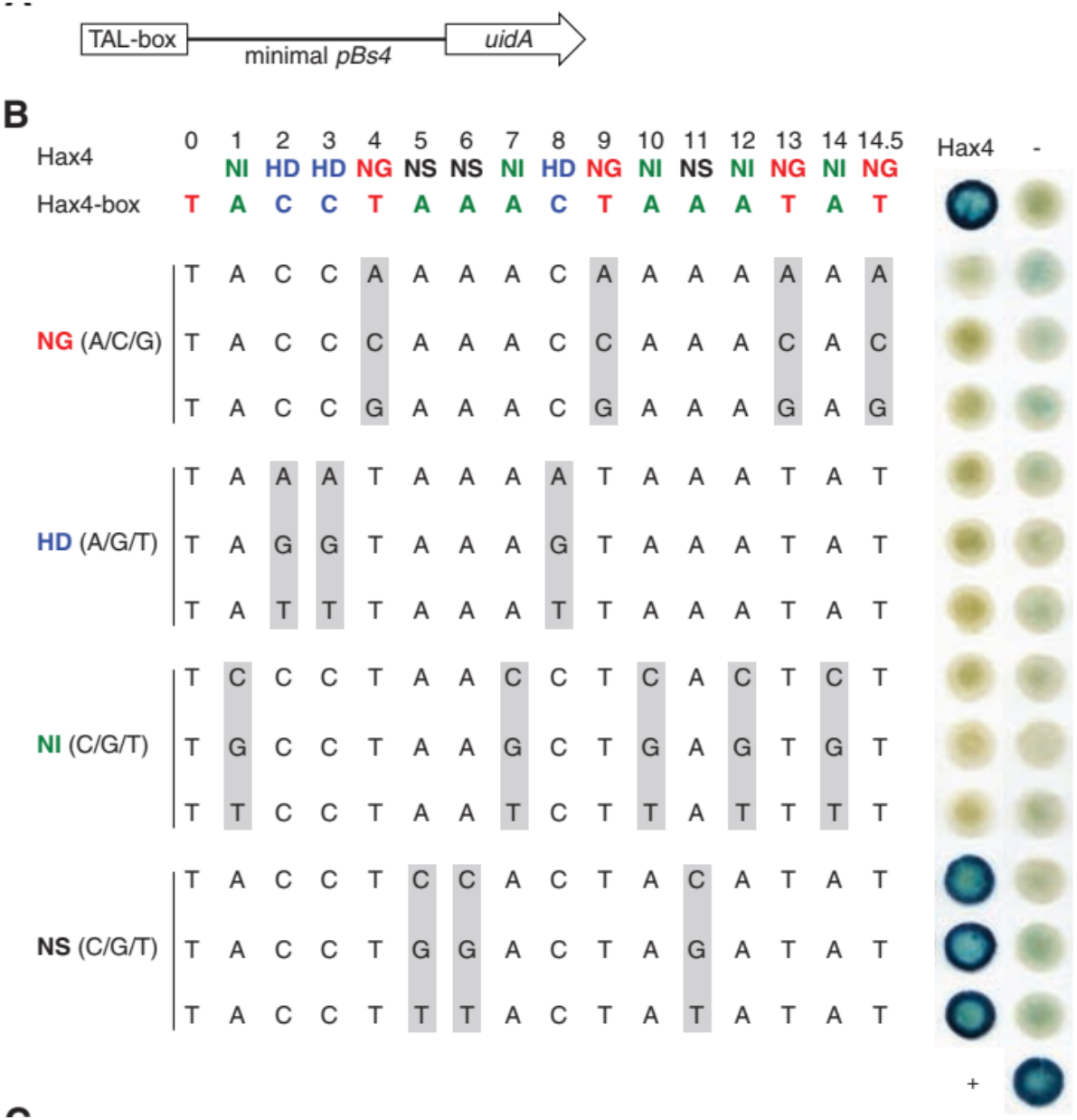




Promoter-box-GUS

Effector genes

Novel Type of DNA Binding Domain

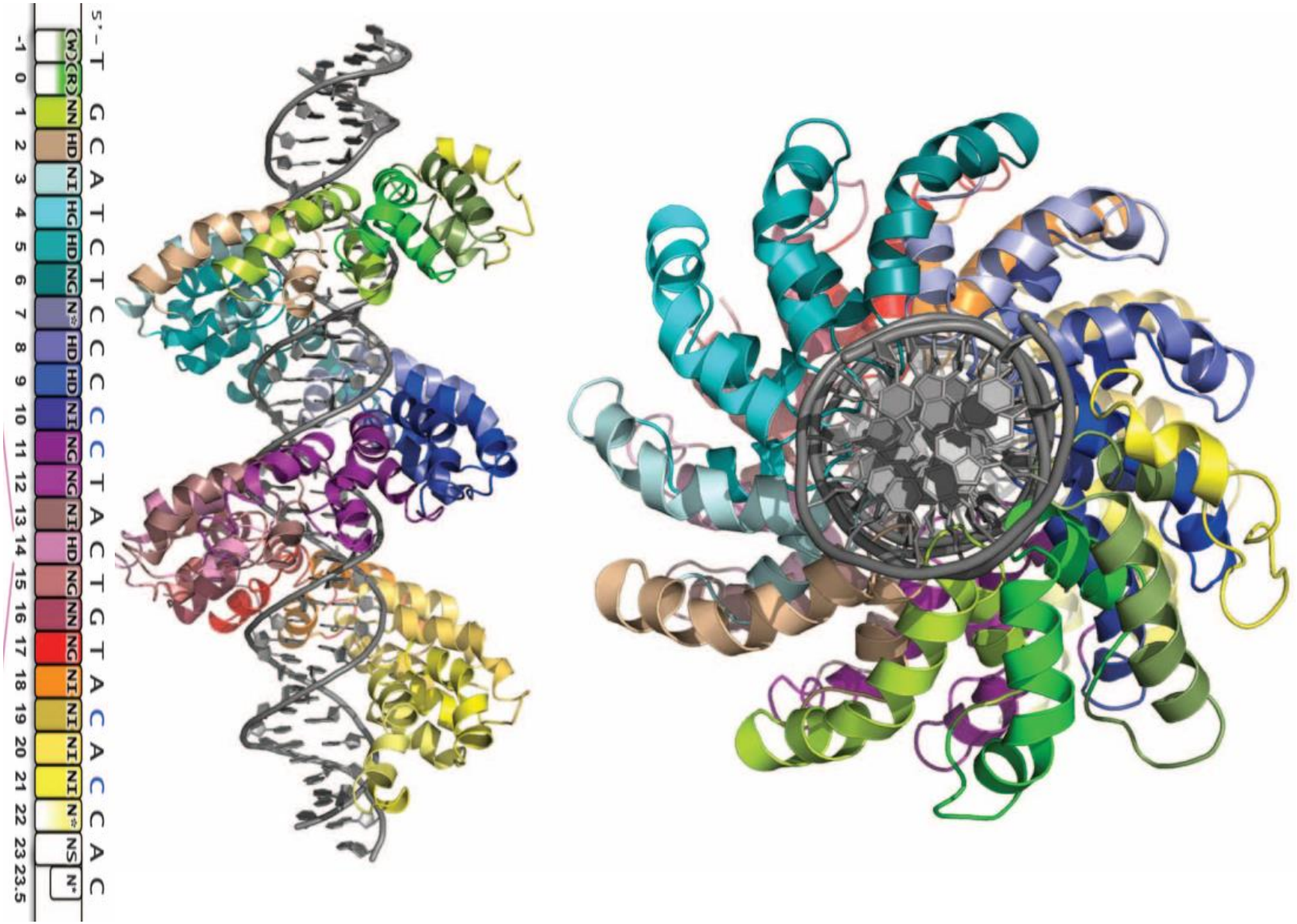


Novel Type of DNA Binding Domain

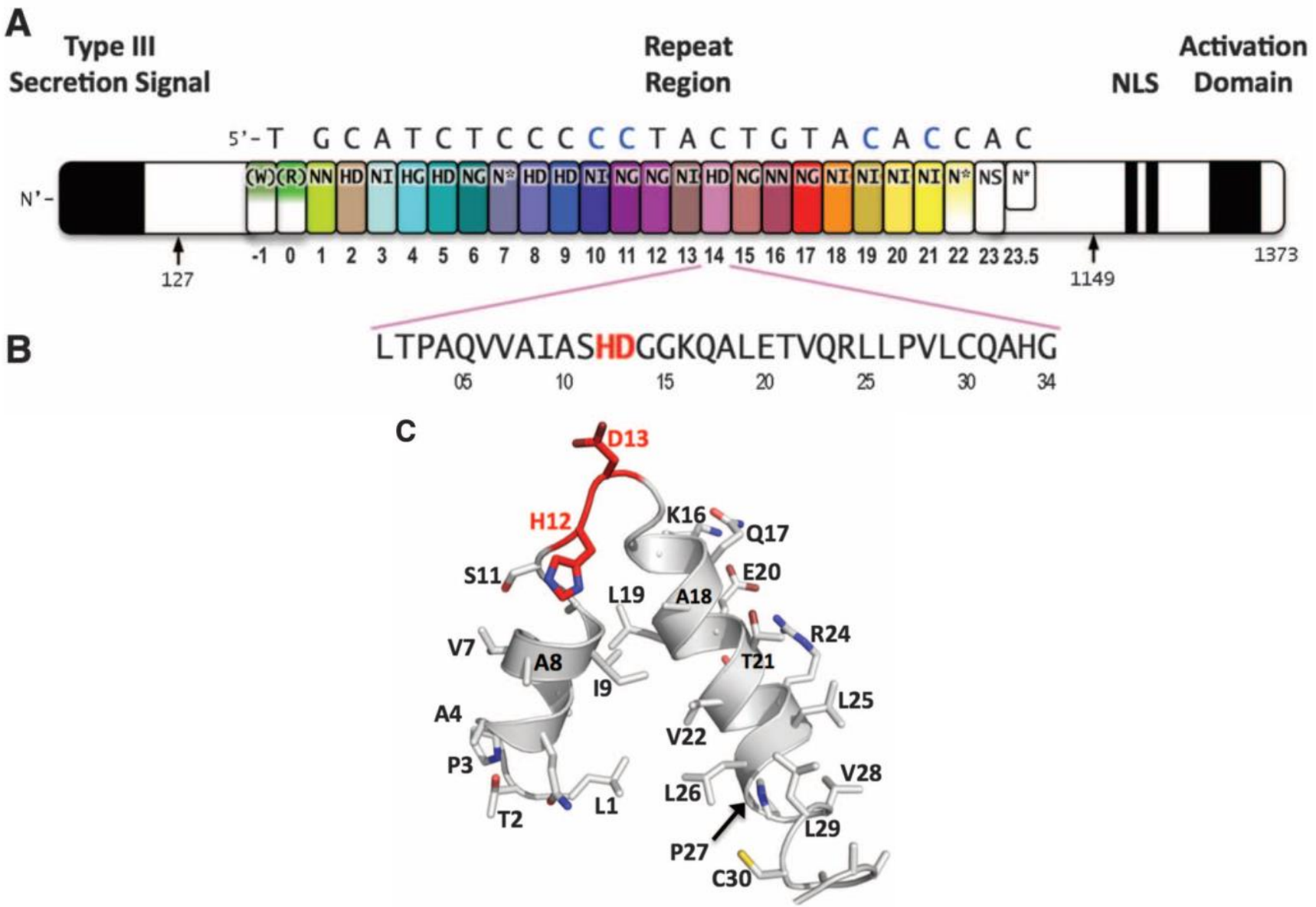
- UPA box was found to be one base pair longer than the AvrBs3 repeats.
- T is conserved at the 5’ end of the repeat (repeat 0).
- Simple code: one repeat = one base pair.

  (NI = A; NG = T; HD = C; NN/NK/NH = G)
- Simple to create artificial effectors with novel targets.

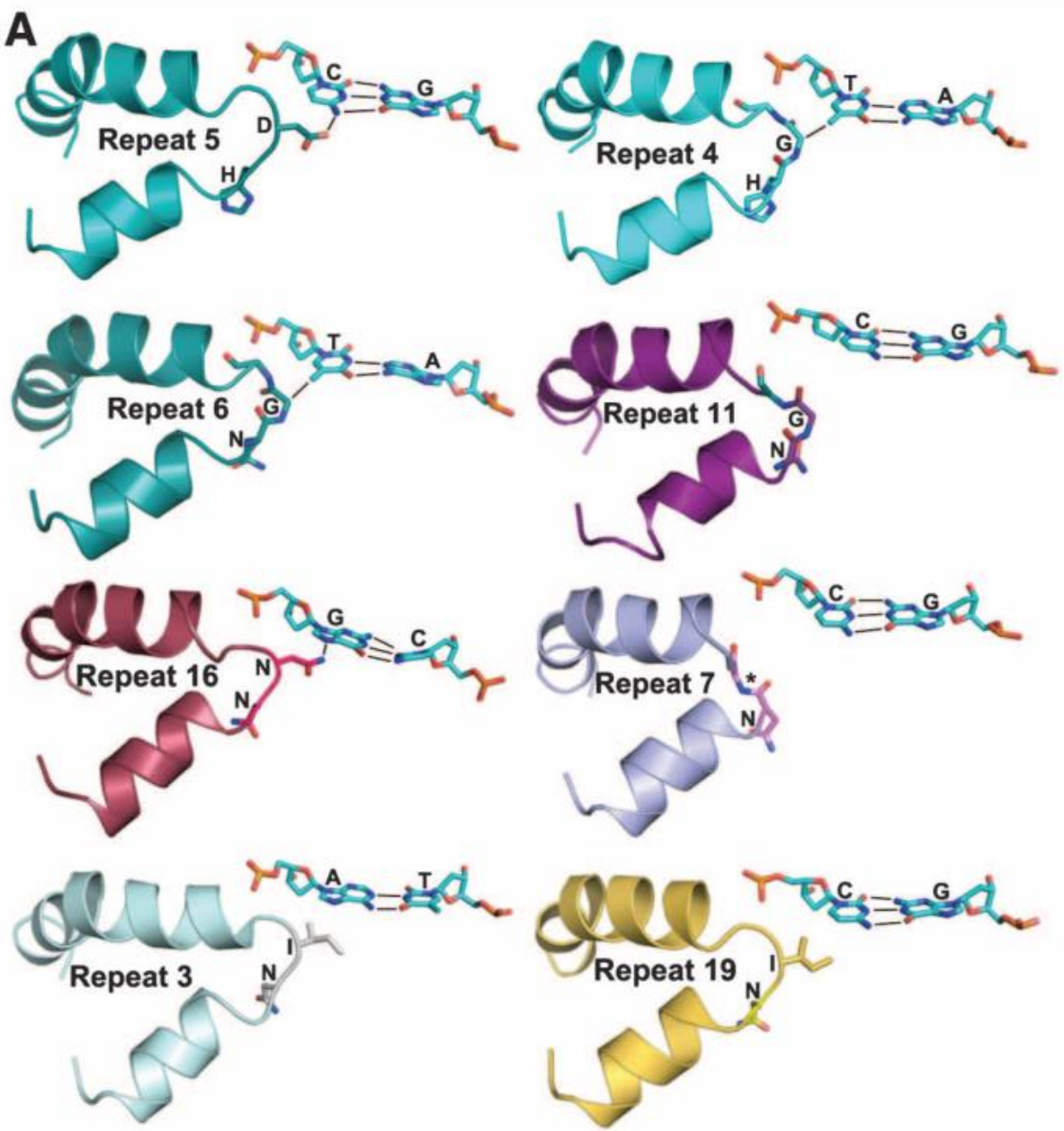
Transcription Activator-Like Effector Structure



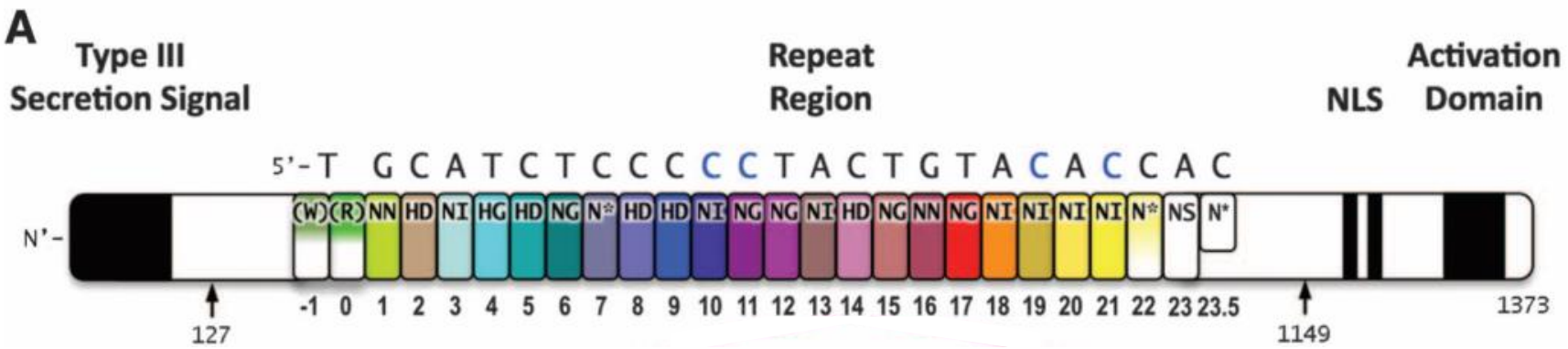
Transcription Activator-Like Effector Structure

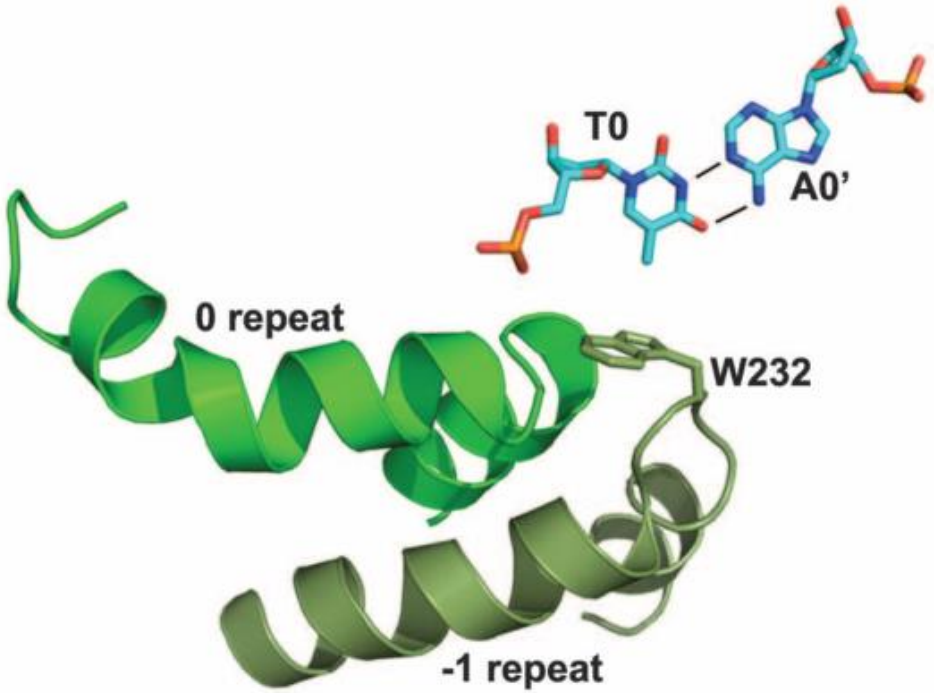


Transcription Activator-Like Effector Structure

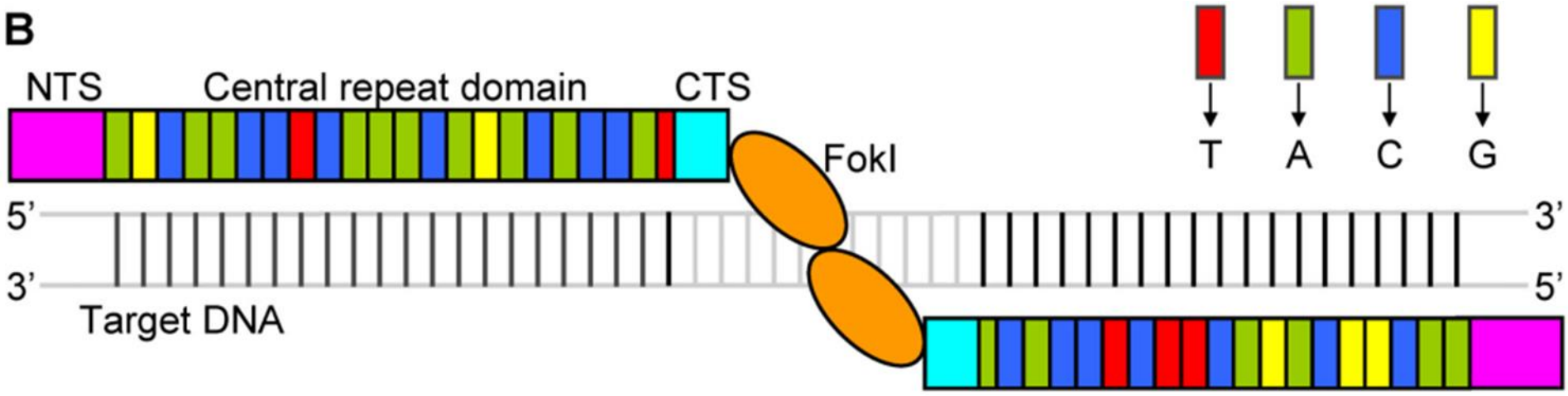


Transcription Activator-Like Effector Structure





Transcription Activator-Like Effector Nucleases (TALENs)



Scaffold Optimization

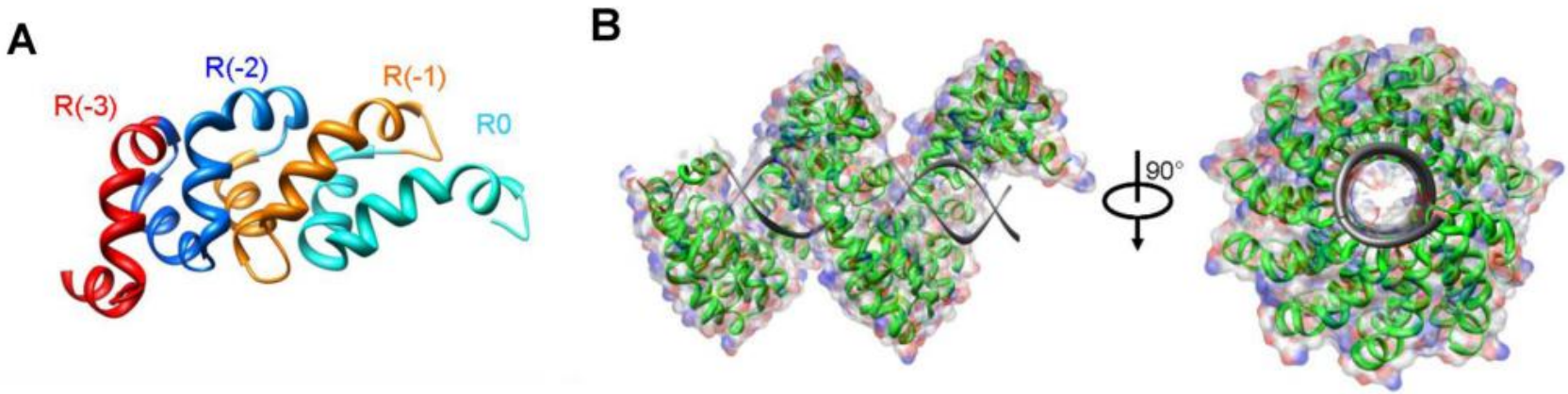
Table II. Engineered TALEN scaffolds with different NTSs and CTSs.

NTS (aa)	CTS (aa)	Spacer (bp)	Reporter system	Refs.
288	295	16–31	β-galactosidase assay in yeast	Li et al. (2011a)
288	285	16[a]	Transient expression assay in tobacco leaves	Mahfouz et al. (2011)
287	231	13–30	β-galactosidase assay in yeast	Christian et al. (2010)
287	63	15[a]	Mutagenesis in medaka embryos	Ansai et al. (2013)
207	63	14–32	β-galactosidase assay in yeast	Sun et al. (2012b)
207	31	10–16	β-galactosidase assay in yeast	Sun et al. (2012b)
153	47	12–21	dsEGFP assay in HEK293	Mussolino et al. (2011)
153	17	12	dsEGFP assay in HEK293	Mussolino et al. (2011)
136	63	12–20	Surveyor nuclease assay in K562	Miller et al. (2011)
136	28	12–13	Surveyor nuclease assay in K562	Miller et al. (2011)
136	18	13–16	β-galactosidase assay in yeast	Christian et al. (2012)

Scaffold Optimization

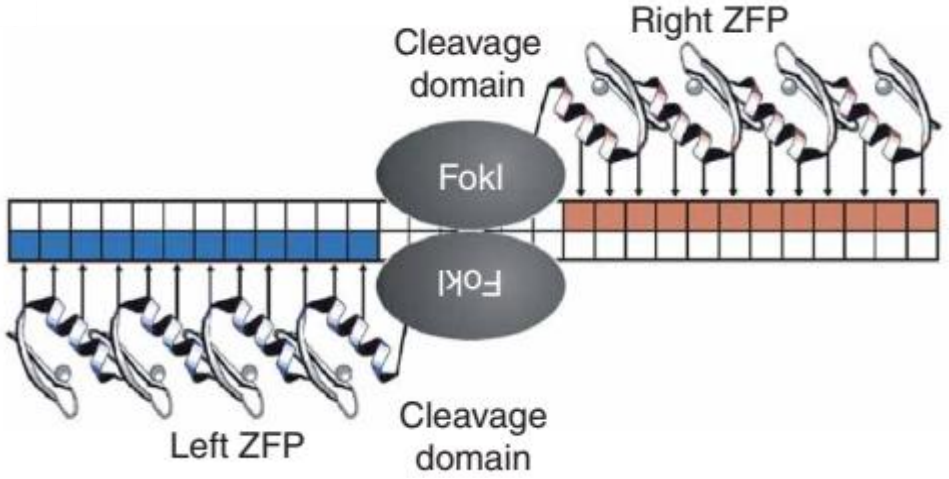
Table II. Engineered TALEN scaffolds with different NTSs and CTSs.

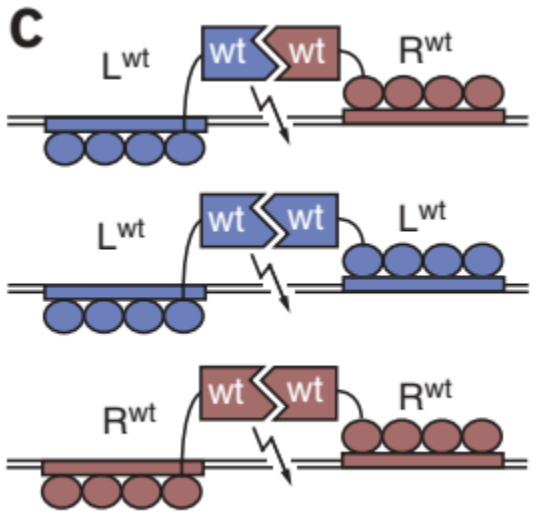
NTS (aa)	CTS (aa)	Spacer (bp)	Reporter system	Refs.
288	295	16–31	β-galactosidase assay in yeast	Li et al. (2011a)
288				
287				
287				
207				
207				
153				
153				
136				
136				
136				



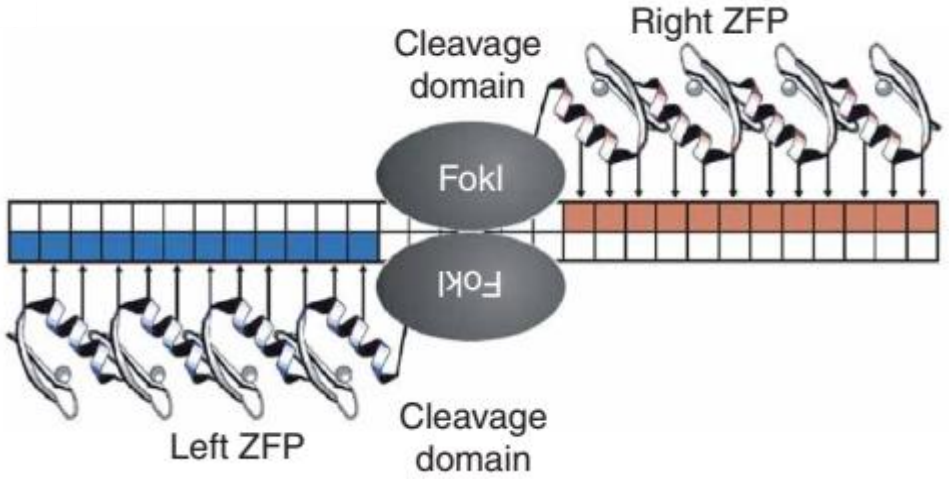
- **All effective TALENs have at least 127 aa preceding the central repeat units**

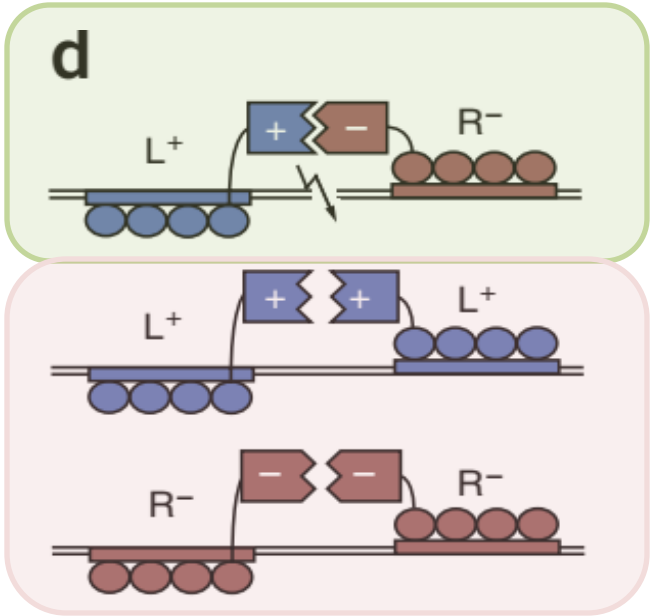
Heterodimeric Nuclease



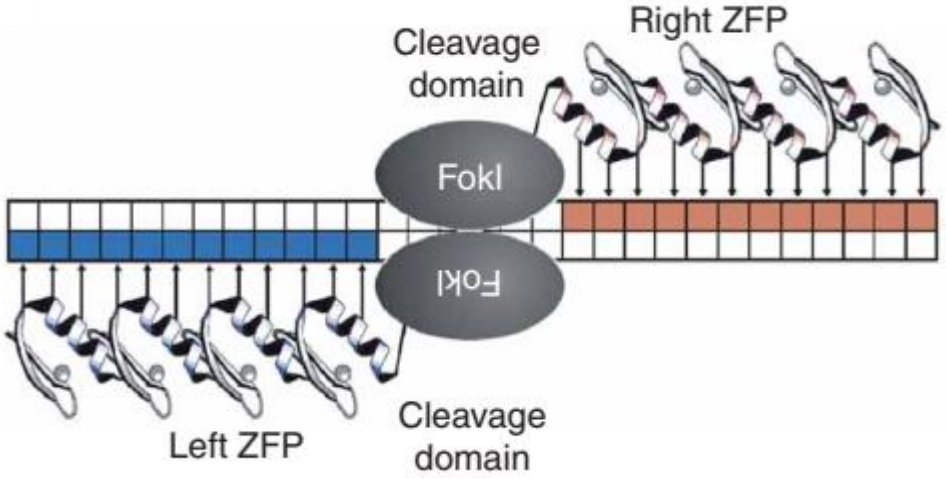


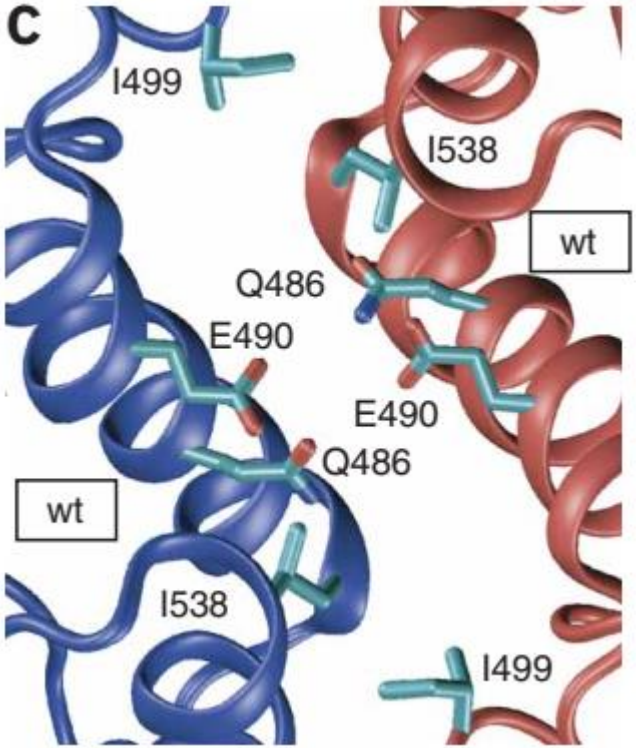
Heterodimeric Nuclease



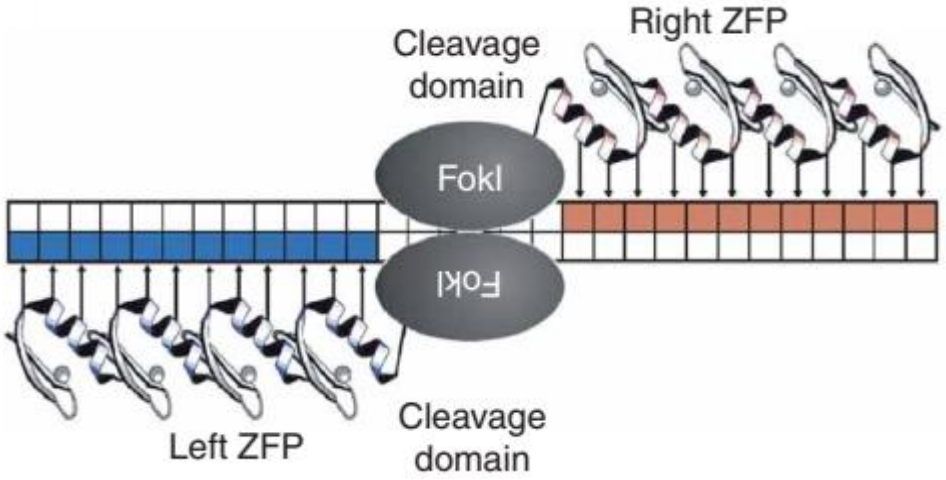


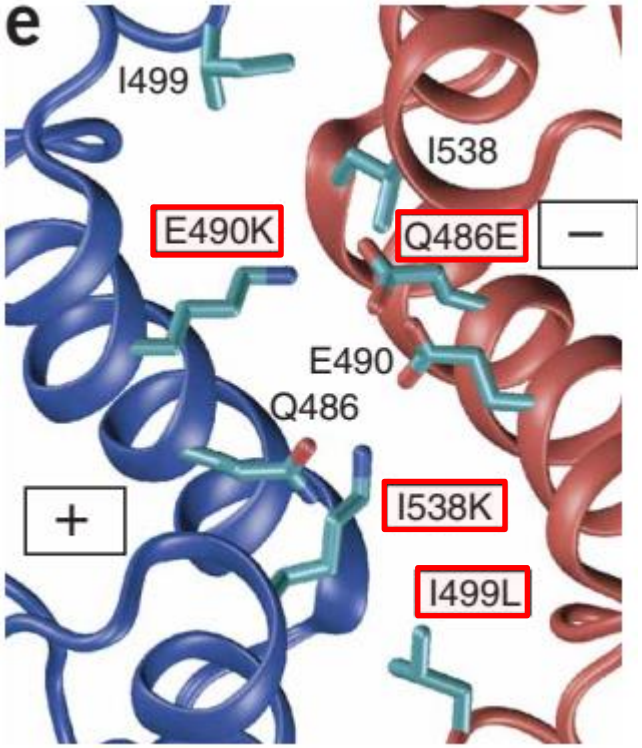
Heterodimeric Nuclease



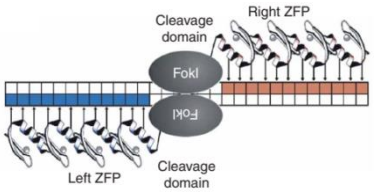


Heterodimeric Nuclease





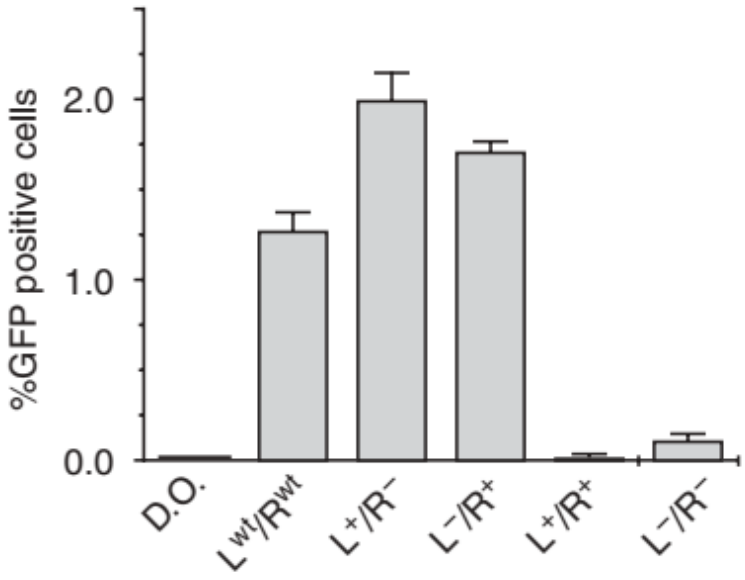
Heterodimeric Nuclease



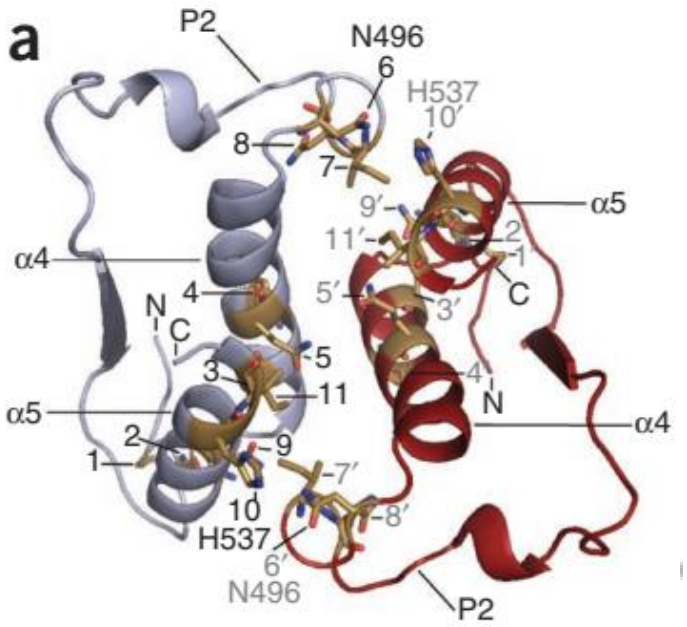
d

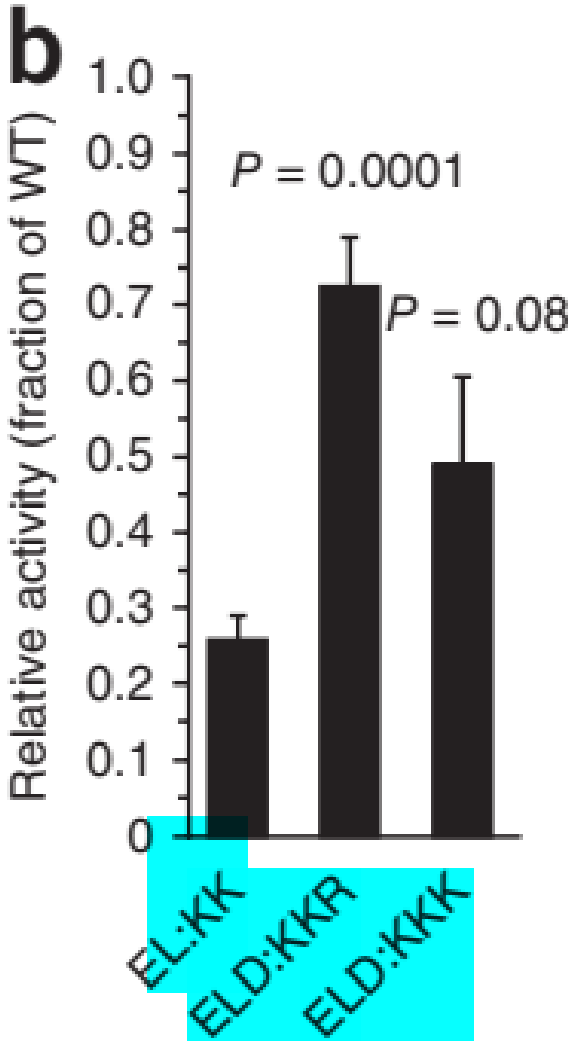
Variants a	Variants b	Relative gene correction rate: Heterodimer a/b	Homodimers a/a	Homodimers b/b
wt	wt	1.00	1.00	1.00
E490K	wt	0.80	0.42	1.00
E490K	Q486E	0.65	0.42	0.19
E490K I538K	Q486E	0.77	0.01	0.19
E490K I538K	Q486E I499L	1.56*	0.02*	0.09*
(+)	(–)			

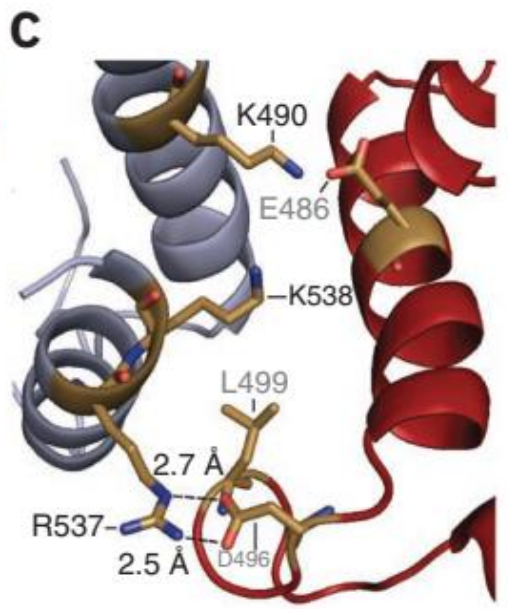
f



Heterodimeric Nuclease







Conclusions

- TALE (AvrBs3) protein binds DNA in a novel way.
- Length of UPA Box corresponds to repeats.
- 13^{th} aa defines the specificity.
- Minimum NTS is 127 aa.
- Obligate-heterodimers (E490K:I538K ; Q486E:I499L); ELD:KKK/ELD:KKR; reduces off-targets.

Questions??

Thank You!