SUPPLEMENTARY DATA

Quantitative Analysis of TALE-DNA Interactions Suggests Polarity Effects

Joshua F. Meckler[1†], Mital S. Bhakta[1†], Moon-Soo Kim[1§], Robert Ovadia[2], Chris H. Habrian[2], Artem Zykovich[1], Abigail Yu[1], Sarah H. Lockwood[1], Robert Morbitzer[3], Janett Elsäesser[3], Thomas Lahaye[3], David J. Segal[1]* and Enoch P. Baldwin[2]*

[1]Genome Center and Department of Biochemistry and Molecular Medicine, [2]Department of Molecular and Cellular Biology, University of California, Davis, CA 95616, [3]Department of Biology, Institute of Genetics, Ludwig-Maximilians-University Munich, 82152 Martinsried, Germany

[§]Current address: Department of Chemistry, 1906 College Heights Blvd., Western Kentucky University, Bowling Green, KY 42101.

[†]These authors contributed equally to this work

*Address correspondence to: djsegal@ucdavis.edu and epbaldwin@ucdavis.edu

Table of Contents

Figure S1: Sequences of dTALEs used in this study.

A) N- and C-terminal truncations reported by several studies.

```
this study
Mussolino et al. (1)                Miller et al. (2)
Sun et al. (3)                      PthXo1 structure (4)
Zhang et al. (5)                    dHax3 structure (6)

BamHI
254-this study
|287-Mussolino et al, Sun et al
||
MDPI
RPRRPSPARELLPGPQPDRVQPTADRGVSAPAGSPLDGLPARRTVSRTRL 201
                          207 Sun et al
PSPPAPSPAFSAGSFSDLLRPFDPSLLDTSLLDSMPAVGTPHTAAAPAEW 151
                                     N2-Zhang et al
                            Factor Xa| 111-this study
             153-Mussolino et al    || |        Δ152-Miller et al
                               |    || |        |135-Sun et al
DEAQSALRAADDPPPTVRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVD 101
    N3-Zhang et al              N5-Zhang et al
    | 94 this study             |       N6-Zhang et al
    | |  125-Sun et al          |       |      N7-Zhang et al
    | |  |    PthXo1 Struct     |       |      |
    | |  |    |  N4-Zhang et al |       |      |
LRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALG 051
                          dHax3 Struct
                          |   N8-Zhang et al
                          |   |   50-Sun et al
                          |   |   |49-Mussolino et al
                          |   |
TVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTDAGELRGPPLQ 001

LDTGQ…           --- TALE REPEATS ---       …SRPDP

                +28-Miller et al   47-Mussolino et al
                |  31-Sun et al    | Factor Xa
                |  |C4-Zhang et al | |   42 this study
                |  ||     Factor Xa| |   |
     17-Mussolino et al   |        | |   Factor Xa
     |          |  ||     |        | |   |
ALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRIGERTSHR 050
+63-Miller et al, Sun et al
 |C3-Zhang et al
VADYAQVVRVLEFFQCHSHPAYAFDEAMTQFGMSRNGLVQLFRRVGVTEL 100
     117-Sun et al
EARGGTLPPASQRWDRILQASGMKRAKPSPTSAQTPDQASLHAFADSLER 150
   163-Sun et al (imprecise due to low conservation with AvrXa10)
   |                         180 this study
   |                         |        200-Sun et al
DLDAPSPMHEGDQTRASSRKRSRSDRAVTGPSAQQAVEVRVPEQRDALHL 200
                   BamHI
                   231-Mussolino et al
PLSWRVKRPRTRIWGGLPDPGTPMAADLAASSTVMWEQDADPFAGAADDF 250
PAFNEEELAWLMELLPQ*
```

B) Construct sequences for bacterial expression for *in vitro* assays

Factor Xa site: IEGR^
TEV protease site: *ENLYFQ^S*
XhoI: CTCGAG
BamHI: GGATCC
StuI: AGGCCT
AatII: GACGTC
AgeI: ACCGGT
HindIII: AAGCTT
His6-purification tag: HHHHHH

Construct Schema:
MBP(not shown)-Factor Xa-*TEV*-XhoIBamHI-StuI-**TALE_repeats**-AatII-AgeI-His6-Stop-HindIII

AvrBs3$_{254-180}$ in pMal-TEV

ATCGAGGGAAGGCTC*GAAAATCTTTATTTTCAGTCT*CTCGAGGATCCCATTCGTTCGCGCACACCAAGTCCTGCCCG
CGAGCTTCTGCCCGGACCCCAACCCGATGGGGTTCAGCCGACTGCAGATCGTGGGGTGTCTCCGCCTGCCGGCGGCC
CCCTGGATGGCTTGCCCGCTCGGCGGACGATGTCCCGGACCCGGCTGCCATCTCCCCCTGCCCCCTCACCTGCGTTC
TCGGCGGGCAGCTTCAGTGACCTGTTACGTCAGTTCGATCCGTCACTTTTTAATACATCGCTTTTTGATTCATTGCC
TCCCTTCGGCGCTCACCATACAGAGGCTGCCACAGGCGAGTGGGATGAGGTGCAATCGGGTCTGCGGGCAGCCGACG
CCCCCCCACCCACCATGCGCGTGGCTGTCACTGCCGCGCGGCCGCCGCGCGCCAAGCCGGCGCCGCGACGACGTGCT
GCGCAACCCTCCGACGCTTCGCCGGCCGCGCAGGTGGATCTACGCACGCTCGGCTACAGCCAGCAGCAACAGGAGAA
GATCAAACCGAAGGTTCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTCGGCCATGGGTTTACACACGCGCACA
TCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCAAGTATCAGGACATGATCGCAGCGTTGCCA
GAGGCGACACACGAAGCGATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCTCTGGAGGCCTTGCTCACGGT
GGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGA
CCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCCTGAACCTGACCCCGGAGCAGGTGGTG
GCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCA
TGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGC
TGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAG
CAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGC
CATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATG
GCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTG
TTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCA
GGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCA
TCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGC
CTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTT
GCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGG
CGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATC
GCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCT
GACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGC
CGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGC
CAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGA
CCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCG
GTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCA
GCAATGGCGGCGGCAGGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACC
CCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGT
GCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGG
AGAGCATTGTTGCCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCC
TGCCTCGGCGGACGTCCTGCGCTGGATGCAGTGAAAAAGGGATTGCCGCACGCGCCGGCCTTGATCAAAAGAACCAA
TCGCCGTATTCCCGAACGCACATCCCATCGCGTTGCCGACCACGCGCAAGTGGTTCGCGTGCTGGGTTTTTCCAGT
GCCACTCCCACCCAGCGCAAGCATTTGATGACGCCATGACGCAGTTCGGGATGAGCAGGCACGGGTTGTTACAGCTC

TTTCGCAGAGTGGGCGTCACCGAACTCGAAGCCCGCAGTGGAACGCTCCCCCCAGCCTCGCAGCGTTGGGACCGTAT
CCTCCAGGCATCAGGGATGAAAAGGGCCAAACCGTCCCCTACTTCAACTCAAACGCCGGATCAGGCGTCTTTGCATG
CATTCGCCGATTCGCTGGAGCGTGACCTTGATGCGCCTAGCCCAATGCACGAGGGAGATCAGACGCGGGCAAGCAGC
CGTAAACGGTCCCGATCGGATCGTGCTGTCACCGGTCATCACCATCACCATCACTGAAAGCTT

IEGRL*ENLYFQS*LEDPIRSRTPSPARELLPGPQPDGVQPTADRGVSPPAGGPLDGLPARR
TMSRTRLPSPPAPSPAFSAGSFSDLLRQFDPSLFNTSLFDSLPPFGAHHTEAATGEWDEV
QSGLRAADAPPPTMRVAVTAARPPRAKPAPRRRAAQPSDASPAAQVDLRTLGYSQQQQEK
IKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVKYQDMIAALPEATHEAIVG
VGKQWSGARALEALLTVAGELRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPALDAVKKGLPHAPALIKRTNRRIPER
TSHRVADHAQVVRVLGFFQCHSHPAQAFDDAMTQFGMSRHGLLQLFRRVGVTELEARSGTLPPA
SQRWDRILQASGMKRAKPSPTSTQTPDQASLHAFADSLERDLDAPSPMHEGDQTRASSRK
RSRSDRAVTGHHHHHH*KL

AvrBs3$_{111-42}$ in pMal-TEV

ATCGAGGGAAGGCTC*GAAAATCTTTATTTTCAGTCT*CTCGAGATGGATCCCTCCGACGCTTCGCCGGCCGCGCAGGT
GGATCTACGCACGCTCGGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACAGTGGCGCAGC
ACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGG
ACCGTCGCTGTCACGTATCAGCACATAATCACGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAA
ACAGTGGTCCGGCGCACGCGCCCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGG
ACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCA
CTGACGGGTGCCCCCCTGAACCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCA
GCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACC
CCGCAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGT
GCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGC
AATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCC
GGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGC
TGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAG
ACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGG
AGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATA

GCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAG
CAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTG
CCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGG
TGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGAT
GGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCA
GGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCC
AGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGACGGTG
CAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGG
CGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGG
TGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGAGCATTGTTGCCCAGTTATCTCGCCCTGATCCG
GCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGACGTCCTGCCATGGATGCAGTGAA
AAAGGGACTGCCGCACGCGCCGGAATTGATCAGAAGAGTCAATCGCCGTCCGGATCCTACCGGTCATCACCATCACC
ATCACTGAAAGCTT

IEGRL*ENLYFQS*LEMDPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHG
FTHAHIVALSQHPAALGTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTVAGE
LRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRPDPT
GHHHHHH*KL

AvrBs3$_{94-42}$ in pMal-TEV

ATCGAGGGAAGGCTCGAAAATCTTTATTTTCAGTCTCTCGAGATGGATCCCAGTCAGCAGCAGCAAGAGAAGATCAA
ACCGAAGGTGCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTG
CGCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCACATAATCACGGCGTTGCCAGAGGCG
ACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCCCTGGAGGCCTTGCTCACGGTGGCGGG
AGAGTTGAGAGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAG
TGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCCTGAACCTGACCCCGGAGCAGGTGGTGGCCATC
GCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCT
GACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGC
CGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCG
CTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGC
CAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGA
CCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCG
GTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCT
GGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCA
GCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACC
CCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGT

GCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGC
AATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCC
GGAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGC
TGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAGCGGTGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCA
CGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGG
AGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATG
GCGGCGGCAGGCCGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAG
CAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTG
CCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGAGCA
TTGTTGCCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCCTGCCTC
GGCGGACGTCCTGCCATGGATGCAGTGAAAAAGGGACTGCCGCACGCGCCGGAATTGATCAGAAGAGTCAATCGCCG
TCCGGATCCTACCGGTCATCACCATCACCATCACTGAAAGCTT

IEGRL*ENLYFQS*LEMDPSQQQQEKIKPKVRSTVAQHHEALVGHGFTHAHIVALSQHPAAL
GTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTVAGELRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQALLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNSGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRPDPT
GHHHHHH*KL

II-NGp$_{111-42}$ in pMal-TEV

ATCGAGGGAAGGCTC*GAAAATCTTTATTTTCAGTCT*CTCGAGATGGATCCCTCCGACGCTTCGCCGGCCGCGCAGGT
GGATCTACGCACGCTCGGCTACAGTCAGCAGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACAGTGGCGCAGC
ACCACGAGGCACTGGTGGGCCATGGGTTTACACACGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGG
ACCGTCGCTGTCACGTATCAGCACATAATCACGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAA
ACAGTGGTCCGGCGCACGCGCCCTGGAGGCCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGG
ACACAGGCCAACTTCTCAAGATTGCAAAACGTGGCGGCGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCA
CTGACGGGTGCCCCCCTGAACCTTACGCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCT
GGAGACGGTGCAGCGGCTGCTTCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCA
GCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGATTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACC
CCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACTGTCCAGCGGCTGTTGCCGGT
GCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTTG
AGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGC
AATGGCGGTGGCAAGCAGGCTCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCC
GGAGCAGGTGGTGGCCATCGCCAGCCACGACGGGGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGC

TGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAG
ACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCAAGCAA
TAAGGGTGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGG
AGCAGGTGGTGGCAATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCCAGGCCCATGGCCTGACCCCGCAACAGGTGGTAGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACACCCCAGCAGGTGGTAGCGATCGCCAGCAATG
GCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGCTTCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAG
CAGGTGGTGGCCATCGCCAGCCACGATGGCGGCAAGCAGGCGCTGGAGACGGTGCAGCGATTGTTGCCGGTGCTGTG
CCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACTG
TCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGC
GGTGGCAAGCAGGCGCTTGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCA
GGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCTCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCC
AGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAAGGGGGGCAAGCAGGCGCTGGAGACGGTG
CAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGG
CGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACACCCCAGCAGG
TGGTGGCCATCGCCAGCAATGGCGGCGGCAGGCCGGCGCTGGAGAGCATTGTTGCCCAGTTATCTCGCCCTGATCCG
GCGTTGGCCGCGTTGACCAACGACCACCTCGTCGCCTTGGCCTGCCTCGGCGGACGTCCTGCCATGGATGCAGTGAA
AAAGGGACTGCCGCACGCGCCGGAATTGATCAGAAGAGTCAATCGCCGTCCGGATCCTACCGGTCATCACCATCACC
ATCACTGAAAGCTT

IEGRL*ENLYFQS*LEMDPSDASPAAQVDLRTLGYSQQQQEKIKPKVRSTVAQHHEALVGHG
FTHAHIVALSQHPAALGTVAVTYQHIITALPEATHEDIVGVGKQWSGARALEALLTVAGE
LRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNKGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNKGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQALLPVLCQAHG
LTPQQVVAIASNGGGRPALESIVAQLSRPDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRPDPT
GHHHHHH*KL

C) Construct sequences for eukaryotic expression for reporter assays

HAtag: YPYDVPDYA
NLS: PKKKRKV
XhoI: CTCGAG
StuI: AGGCCT
AatII: GACGTC
AgeI: ACCGGT
VP64: *DALDDFDLDMLDALDDFDLDMLDALDDFDLDMLDALDDFDLDML*
PstI: CTGCAG

Construct Schema:
HAtag-NLS-XhoI-StuI-**TALE_repeats**-AatII-AgeI-VP64-**Stop**-PstI

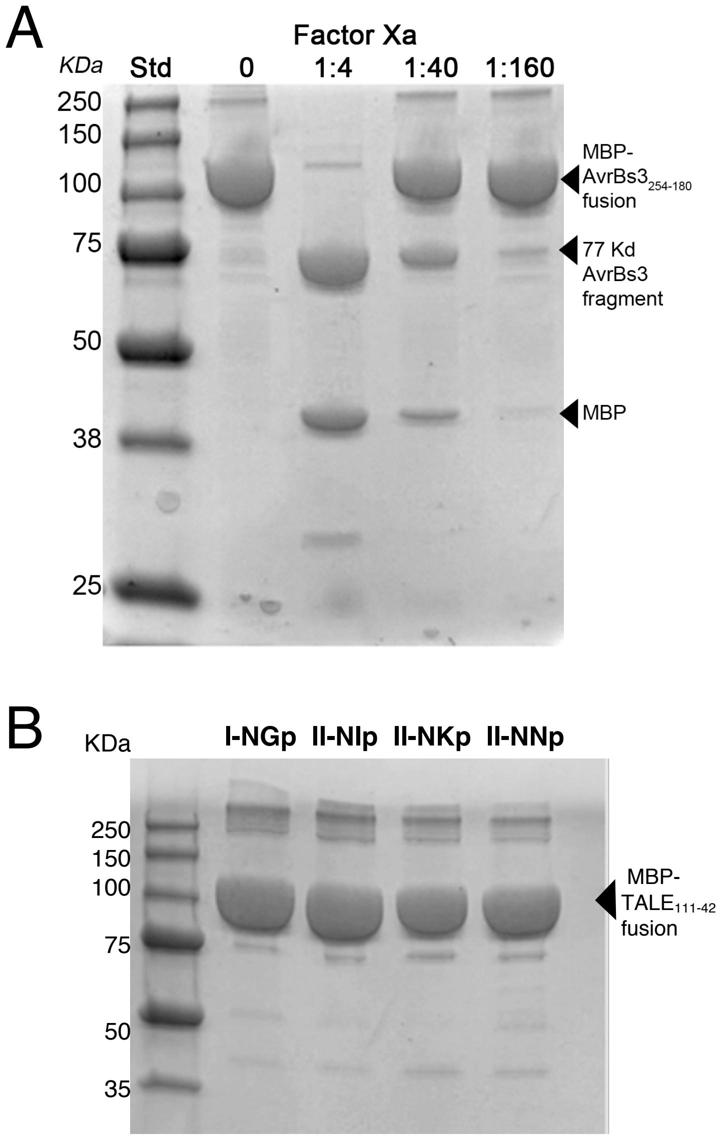
I-NGp$_{111-42}$ in pPGK-VP64

ATGTACCCATACGATGTCCCAGACTACGCGAATTCCCCGGGGATCCCAGGCATGGGGCCCAAAAAGAAACGCAAAGT
TGGGCGCCTCGAGATGGATCCCTCCGACGCTTCGCCGGCCGCGCAGGTGGATCTACGCACGCTCGGCTACAGTCAGC
AGCAGCAAGAGAAGATCAAACCGAAGGTGCGTTCGACAGTGGCGCAGCACCACGAGGCACTGGTGGGCCATGGGTTT
ACACACGCGCACATCGTTGCGCTCAGCCAACACCCGGCAGCGTTAGGGACCGTCGCTGTCACGTATCAGCACATAAT
CACGGCGTTGCCAGAGGCGACACACGAAGACATCGTTGGCGTCGGCAAACAGTGGTCCGGCGCACGCGCCCTGGAGG
CCTTGCTCACGGTGGCGGGAGAGTTGAGAGGTCCACCGTTACAGTTGGACACAGGCCAACTTCTCAAGATTGCAAAA
CGTGGCGGCGTGACCGCAGTGGAGGCAGTGCATGCATGGCGCAATGCACTGACGGGTGCCCCCCTGAACCTTACGCC
GCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGCTTCCGGTGC
TGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAG
ACGGTGCAGCGATTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAA
TGGCGGTGGCAAGCAGGCGCTGGAGACTGTCCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGG
AGCAGGTGGTGGCCATCGCCAGCCACGACGGTGGCAAGCAGGCGCTTGAGACGGTGCAGCGGCTGTTGCCGGTGCTG
TGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCTCTGGAGAC
GGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATA
AGGGGGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAG
CAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGGCGCTGTTGCCGGTGCTGTG
CCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCAAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGG
TGCAGGCGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCAATCGCCAGCAATGGC
GGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAACA
GGTGGTAGCCATCGCCAGCAATATTGGTGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCC
AGGCCCATGGCCTGACACCCCAGCAGGTGGTAGCGATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACGGTG
CAGCGGCTGCTTCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGG
TGGCAAGCAGGCGCTGGAGACGGTGCAGCGATTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGG
TGGTGGCCATCGCCAGCAATGGCGGTGGCAAGCAGGCGCTGGAGACTGTCCAGCGGCTGTTGCCGGTGCTGTGCCAG
GCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATAAGGGTGGCAAGCAGGCGCTTGAGACGGTGCA
GCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTGGTGGCCATCGCCAGCAATGGCGGTG
GCAAGCAGGCTCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACCCCGGAGCAGGTG
GTGGCCATCGCCAGCCACGACGGGGGCAAGCAGGCGCTGGAGACGGTGCAGCGGCTGTTGCCGGTGCTGTGCCAGGC
CCATGGCCTGACCCCGCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGCAAGCAGGCGCTGGAGACGGTGCAGG
CGCTGTTGCCGGTGCTGTGCCAGGCCCATGGCCTGACACCCCAGCAGGTGGTGGCCATCGCCAGCAATGGCGGCGGC
AGGCCGGCGCTGGAGAGCATTGTTGCCCAGTTATCTCGCCCTGATCCGGCGTTGGCCGCGTTGACCAACGACCACCT
CGTCGCCTTGGCCTGCCTCGGCGGACGTCCTGCCATGGATGCAGTGAAAAAGGGACTGCCGCACGCGCCGGAATTGA
TCAGAAGAGTCAATCGCCGTCCGGATCCTACCGGTGCGGCCGCC*GACGCTTTGGATGACTTTGATTTGGACATGCTG*
GATGCTCTAGATGACTTCGACCTGGATATGCTGGACGCACTTGACGACTTTGACCTCGACATGCTAGACGCTCTGGA
*CGACTTCGATCTAGACATGCTC***TAA**GTCGACCTGCAG

MYPYDVPDYANSPGIPGMGPKKKRKVGRLEMDPSDASPAAQVDLRTLGYSQQQQEKIKPK
VRSTVAQHHEALVGHGFTHAHIVALSQHPAALGTVAVTYQHIITALPEATHEDIVGVGKQ
WSGARALEALLTVAGELRGPPLQ
LDTGQLLKIAKRGGVTAVEAVHAWRNALTGAPLN
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG

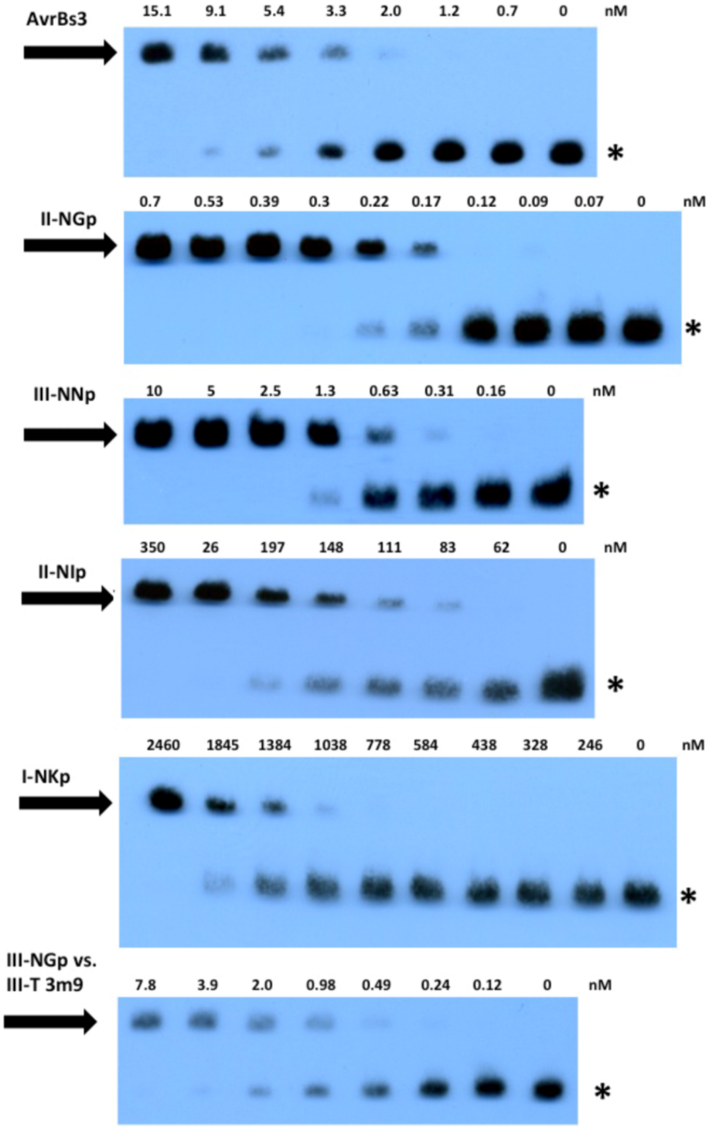
LTPEQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNKGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQALLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNIGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNKGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASNGGGKQALETVQRLLPVLCQAHG
LTPEQVVAIASHDGGKQALETVQRLLPVLCQAHG
LTPQQVVAIASNGGGKQALETVQALLPVLCQAHGLTPQQVVAIASNGGGRPALESIVAQ
LSRPDPALAALTNDHLVALACLGGRPAMDAVKKGLPHAPELIRRVNRRPDPTGAAA_DALD_
_DFDLDMLDALDDFDLDMLDALDDFDLDMLDALDDFDLDML_*VDLQ

Figure S2: Coomassie-stained protein gel of representative dTALEs.



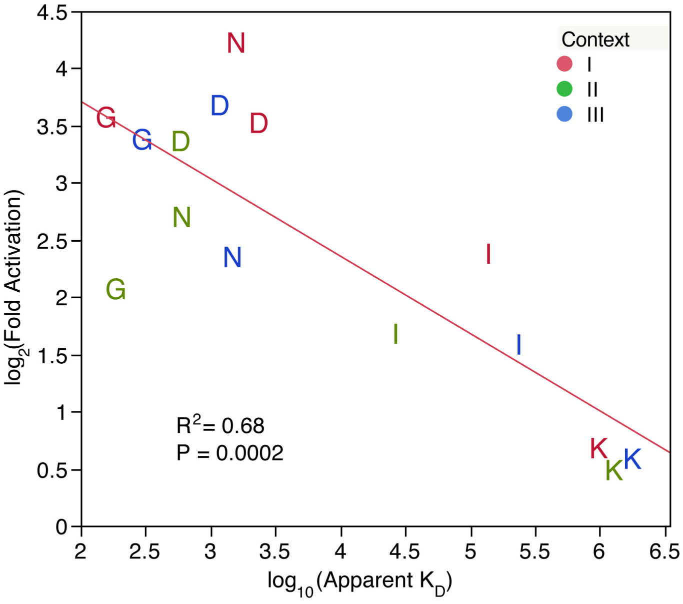
A) Coomassie-stained SDS PAGE gel shows purified **AvrBs3$_{254-180}$** after treatment with various concentrations of Factor Xa protease. B) Purified dTALE$_{111-42}$ proteins run at the expected 100 kilodalton size. Proteins shown are diluted 1:5 from stock. Typical protein yields were between 10 and 30 μM.

Figure S3. Representative EMSA data.



Sample data showing the EMSA dynamic range of 10^4. The TALEs used are shown and unless stated, the cognate DNA targets are used. The shifted TALE-DNA band is indicated with an arrow while the unshifted DNA is indicated with a “*”. The bottom panel shows representative data for a mixed site target used in polarity experiments. The sharpness of the transition (apparent nH = 3-4) is consistent over 4 logs of protein concentration, and therefore is not due to lack of equilibration or protein concentration-dependent aggregation. We do not know its origin.

Figure S4: Correlation of affinity with reporter gene activation.



The relationship between apparent K_D and fold activation of the reporter gene was not linear. The best fit model of these data was a log-log relationship. The correlation of $\log_2$(Fold Activation) versus $\log_{10}$(Apparent K_D (pM)) had an $R^2 = 0.68$ and a $P = 0.0002$. Host context is indicated by color, guest RVD by letter; G, N<u>G</u>; N, N<u>N</u>; D, H<u>D</u>; I, N<u>I</u>; K, N<u>K</u>.

Table S1: Primers used to create N- and C-terminal truncations.

EcoRI: GAATTC
5'SfiI: GGCCGCTAAGGCC
3’SfiI: GGCCAAGCTGGCC
XhoI: CTCGAG
BamHI: GGATCC
SmaI: CCCGGG
AgeI: ACCGGT

>PthXoI_5’111-f
5’-AATAGGAGGTGCACCGAATTCGTGGCCGCTAAGGCCCTCGAGATGGATCCTCCGACGCTTCGCCGGCCGCGCAGGT

>PthXoI_5’94-f
5’-AATAGGAGGTGCACCGAATTCGTGGCCGCTAAGGCCCTCGAGATGGATCCAGTCAGCAGCAGCAAGAGAAGATCAAACC

>PthXoI_5’end-r
5’-CGACGAGGTGGTCGTTGGTCAACG**CCCGGG**CTGTAACGGCGGACCTCTCAACTC

>PthXoI_3’end-f
5’-GAGTTGAGAGGTCCGCCGTTACAG**CCCGGG**CGTTGACCAACGACCACCTCGTCG

>PthXoI_3’42-r
5’-CAAGAAAGCTGGGTCGAATTCGGCCAAGCTGGCCTTACCGGTAGGATCCGGACGGCGATTGACTCTTCTGATCAATTC

Table S2: Primers used in EMSA assays.

Biotinylated forward primer	/5Biosg/CCTCTTCGCTATTACGCCAGC
Reverse primer	CACCCTGACTCGAGTACGATCGAACGTTC

For RVD affinity studies

I-A	CCTCTTCGCTATTACGCCAGC	TAAACAGATAAATAGACAA	GAACGTTCGATCGTACTCGAGTCAGGGTG
I-G	CCTCTTCGCTATTACGCCAGC	TGAGCGGGTGAGTGGGCGG	GAACGTTCGATCGTACTCGAGTCAGGGTG
I-C	CCTCTTCGCTATTACGCCAGC	TCACCCGCTCACTCGCCCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
I-T	CCTCTTCGCTATTACGCCAGC	TTATCTGTTTATTTGTCTT	GAACGTTCGATCGTACTCGAGTCAGGGTG
II-A	CCTCTTCGCTATTACGCCAGC	TAAATACAGAAACATAGAA	GAACGTTCGATCGTACTCGAGTCAGGGTG
II-G	CCTCTTCGCTATTACGCCAGC	TGAGTGCGGGAGCGTGGGG	GAACGTTCGATCGTACTCGAGTCAGGGTG
II-C	CCTCTTCGCTATTACGCCAGC	TCACTCCCGCACCCTCGCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
II-T	CCTCTTCGCTATTACGCCAGC	TTATTTCTGTATCTTTGTT	GAACGTTCGATCGTACTCGAGTCAGGGTG
III-A	CCTCTTCGCTATTACGCCAGC	TACAAAGACATAGAAATAA	GAACGTTCGATCGTACTCGAGTCAGGGTG
III-G	CCTCTTCGCTATTACGCCAGC	TGCGAGGGCGTGGGAGTGG	GAACGTTCGATCGTACTCGAGTCAGGGTG
III-C	CCTCTTCGCTATTACGCCAGC	TCCCACGCCCTCGCACTCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
III-T	CCTCTTCGCTATTACGCCAGC	TTCTATGTCTTTGTATTTT	GAACGTTCGATCGTACTCGAGTCAGGGTG

Zif268	CCTCTTCGCTATTACGCCAGC	GCGTGGGCGT	GAACGTTCGATCGTACTCGAGTCAGGGTG

For 5’T studies

Bs3	CCTCTTCGCTATTACGCCAGC	TATATAAACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5’A	CCTCTTCGCTATTACGCCAGC	AATATAAACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5’C	CCTCTTCGCTATTACGCCAGC	CATATAAACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5’G	CCTCTTCGCTATTACGCCAGC	GATATAAACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG

For polarity studies (B = C,G,T; D = A,G,T; H = A,C,T; V = A,C,G)

Bs3 variants

3m3	CCTCTTCGCTATTACGCCAGC	TATATAAACCTAACCAVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m6	CCTCTTCGCTATTACGCCAGC	TATATAAACCTAADDBVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m9	CCTCTTCGCTATTACGCCAGC	TATATAAACCVBBDDBVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m3	CCTCTTCGCTATTACGCCAGC	TBVBTAAACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m6	CCTCTTCGCTATTACGCCAGC	TBVBVBBACCTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m9	CCTCTTCGCTATTACGCCAGC	TBVBVBBBDDTAACCATCC	GAACGTTCGATCGTACTCGAGTCAGGGTG

III-C variants

3m3	CCTCTTCGCTATTACGCCAGC	TCCCACGCCCTCGCACVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m6	CCTCTTCGCTATTACGCCAGC	TCCCACGCCCTCGDBDVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m9	CCTCTTCGCTATTACGCCAGC	TCCCACGCCCVDHDBDVDD	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m3	CCTCTTCGCTATTACGCCAGC	TDDDACGCCCTCGCACTCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m6	CCTCTTCGCTATTACGCCAGC	TDDDBDHCCCTCGCACTCC	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m9	CCTCTTCGCTATTACGCCAGC	TDDDBDHDDDTCGCACTCC	GAACGTTCGATCGTACTCGAGTCAGGGTG

III-T variants

3m3	CCTCTTCGCTATTACGCCAGC	TTCTATGTCTTTGTATVVV	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m6	CCTCTTCGCTATTACGCCAGC	TTCTATGTCTTTGVBVVVV	GAACGTTCGATCGTACTCGAGTCAGGGTG
3m9	CCTCTTCGCTATTACGCCAGC	TTCTATGTCTVVHVBVVVV	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m3	CCTCTTCGCTATTACGCCAGC	TVDVATGTCTTTGTATTTT	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m6	CCTCTTCGCTATTACGCCAGC	TVDVBVHTCTTTGTATTTT	GAACGTTCGATCGTACTCGAGTCAGGGTG
5m9	CCTCTTCGCTATTACGCCAGC	TVDVBVHVDVTTGTATTTT	GAACGTTCGATCGTACTCGAGTCAGGGTG

Table S3: Primers for construction of ATF reporter plasmids.

XhoI: CTCGAG
NotI: GCGGCCGC

For RVD studies

>pGL3-control-Not-F
5′- GAGGAGGCGGCCGCAATAAAATATCTTTATTTTC

>PGL3-I-A-R
5′- CTCCTCCTCGAGTTGTCTATTTATCTGTTTACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-I-G-R
5′- CTCCTCCTCGAGCCGCCCACTCACCCGCTCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-I-C-R
5′- CTCCTCCTCGAGGGGGCGAGTGAGCGGGTGACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-I-T-R
5′- CTCCTCCTCGAGAAGACAAATAAACAGATAACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-II-A-R
5′- CTCCTCCTCGAGTTCTATGTTTCTGTATTTACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-II-G-R
5′- CTCCTCCTCGAGCCCCACGCTCCCGCACTCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-II-C-R
5′- CTCCTCCTCGAGGGCGAGGGTGCGGGAGTGACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-II-T-R
5′- CTCCTCCTCGAGAACAAAGATACAGAAATAACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-A-R
5′- CTCCTCCTCGAGTTATTTCTATGTCTTTGTACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-G-R
5′- CTCCTCCTCGAGCCACTCCCACGCCCTCGCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-C-R
5′- CTCCTCCTCGAGGGAGTGCGAGGGCGTGGGACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-R
5′- CTCCTCCTCGAGAAAATACAAAGACATAGAACCCGGGCTAGCACGCGTAAGAGCTC -3′

For AvrBs3 variant studies

>PGL3-Bs3-wt-R
5′- CTCCTCCTCGAGGGATGGTTAGGTTTATATACCCGGGCTAGCACGCGTAAGAGCTC -3′

>pGL3-Bs3-UPA-R
5′- CTCCTCCTCGAGAGAGGGTTAGGTTTATATACCCGGGCTAGCACGCGTAAGAGCTC -3′

For AvrBs3 5′T studies

>PGL3-Bs3-5A-R
5′- CTCCTCCTCGAGGGATGGTTAGGTTTATATTCCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-5C-R
5′- CTCCTCCTCGAGGGATGGTTAGGTTTATATGCCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-5G-R
5′- CTCCTCCTCGAGGGATGGTTAGGTTTATATCCCCGGGCTAGCACGCGTAAGAGCTC -3′

For polarity studies

>PGL3-Bs3-3m3-R
5′- CTCCTCCTCGAGCCCTGGTTAGGTTTATATACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-3m6-R
5′- CTCCTCCTCGAGCCCCCCTTAGGTTTATATACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-3m9-R
5′- CTCCTCCTCGAGCCCCCCAACGGTTTATATACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-5m3-R
5′- CTCCTCCTCGAGGGATGGTTAGGTTTAACCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-5m6-R
5′- CTCCTCCTCGAGGGATGGTTAGGTAACACCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-Bs3-5m9-R
5′- CTCCTCCTCGAGGGATGGTTACCAAACACCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-3m3-R
5′- CTCCTCCTCGAGCCCATACAAAGACATAGAACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-3m6-R
5′- CTCCTCCTCGAGCCCCACCAAAGACATAGAACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-3m9-R
5′- CTCCTCCTCGAGCCCCACGCCAGACATAGAACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-5m3-R
5′- CTCCTCCTCGAGAAAATACAAAGACATCCCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-5m6-R
5′- CTCCTCCTCGAGAAAATACAAAGAGCACCCACCCGGGCTAGCACGCGTAAGAGCTC -3′

>PGL3-III-T-5m9-R
5′- CTCCTCCTCGAGAAAATACAACCCGCACCCACCCGGGCTAGCACGCGTAAGAGCTC -3′

Alignment

>PGL3-*Bs3*-UPA-R	TATATAAACCTAACCCTCT
>PGL3-*Bs3*-wt-R	TATATAAACCTAACCATCC
>PGL3-*Bs3*-5A-R	AATATAAACCTAACCATCC
>PGL3-*Bs3*-5C-R	CATATAAACCTAACCATCC

>PGL3-*Bs3*-5G-R	GATATAAACCTAACCATCC

>PGL3-I-A-R	TAAACAGATAAATAGACAA
>PGL3-I-G-R	TGAGCGGGTGAGTGGGCGG
>PGL3-I-C-R	TCACCCGCTCACTCGCCCC
>PGL3-I-T-R	TTATCTGTTTATTTGTCTT
>PGL3-II-A-R	TAAATACAGAAACATAGAA
>PGL3-II-G-R	TGAGTGCGGGAGCGTGGGG
>PGL3-II-C-R	TCACTCCCGCACCCTCGCC
>PGL3-II-T-R	TTATTTCTGTATCTTTGTT
>PGL3-III-A-R	TACAAAGACATAGAAATAA
>PGL3-III-G-R	TGCGAGGGCGTGGGAGTGG
>PGL3-III-C-R	TCCCACGCCCTCGCACTCC
>PGL3-III-T-R	TTCTATGTCTTTGTATTTT

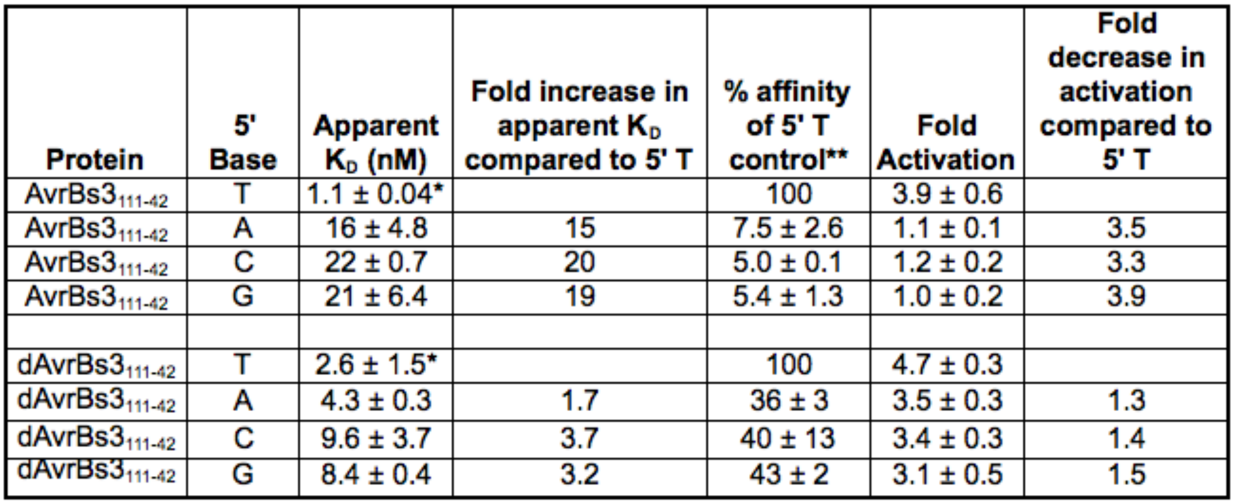
>PGL3-*Bs3*-wt-R	TATATAAACCTAACCATCC
>PGL3-*Bs3*-3m3-R	TATATAAACCTAACCAGGG
>PGL3-*Bs3*-3m6-R	TATATAAACCTAAGGGGGG
>PGL3-*Bs3*-3m9-R	TATATAAACCGTTGGGGGG
>PGL3-*Bs3*-5m3-R	TGGTTAAACCTAACCATCC
>PGL3-*Bs3*-5m6-R	TGGTGTTACCTAACCATCC
>PGL3-*Bs3*-5m9-R	TGGTGTTTGGTAACCATCC

>PGL3-III-T-wt-R	TTCTATGTCTTTGTATTTT
>PGL3-III-T-3m3-R	TTCTATGTCTTTGTATGGG
>PGL3-III-T-3m6-R	TTCTATGTCTTTGGTGGGG
>PGL3-III-T-3m9-R	TTCTATGTCTGGCGTGGGG
>PGL3-III-T-5m3-R	TGGGATGTCTTTGTATTTT
>PGL3-III-T-5m6-R	TGGGTGCTCTTTGTATTTT
>PGL3-III-T-5m9-R	TGGGTGCGGGTTGTATTTT

Table S4: EMSA and ATF data for G/A specificity of the NN RVD.

dTALE	Target Site	Apparent K_D (nM)	Fold increase in apparent K_D over G guest	Fold Activation	Fold decrease in activation over G guest
I-NNp	I-G	1.3 ± 0.3		5.1 ± 0.4	
I-NNp	I-A	64 ± 5	49	0.86 ± 0.17	6
II-NNp	II-G	1.3 ± 0.3		19 ± 2	
II-NNp	II-A	53 ± 11	41	2.4 ± 0.1	8
III-NNp	III-G	0.6 ± 0.2		6.5 ± 0.9	
III-NNp	III-A	10 ± 4	17	0.53 ± 0.03	12

Table S5: EMSA and ATF data probing the requirement of a 5'T.



Protein	5' Base	Apparent K_D (nM)	Fold increase in apparent K_D compared to 5' T	% affinity of 5' T control**	Fold Activation	Fold decrease in activation compared to 5' T
$\text{AvrBs3}_{111\text{-}42}$	T	1.1 ± 0.04*		100	3.9 ± 0.6	
$\text{AvrBs3}_{111\text{-}42}$	A	16 ± 4.8	15	7.5 ± 2.6	1.1 ± 0.1	3.5
$\text{AvrBs3}_{111\text{-}42}$	C	22 ± 0.7	20	5.0 ± 0.1	1.2 ± 0.2	3.3
$\text{AvrBs3}_{111\text{-}42}$	G	21 ± 6.4	19	5.4 ± 1.3	1.0 ± 0.2	3.9
$\text{dAvrBs3}_{111\text{-}42}$	T	2.6 ± 1.5*		100	4.7 ± 0.3	
$\text{dAvrBs3}_{111\text{-}42}$	A	4.3 ± 0.3	1.7	36 ± 3	3.5 ± 0.3	1.3
$\text{dAvrBs3}_{111\text{-}42}$	C	9.6 ± 3.7	3.7	40 ± 13	3.4 ± 0.3	1.4
$\text{dAvrBs3}_{111\text{-}42}$	G	8.4 ± 0.4	3.2	43 ± 2	3.1 ± 0.5	1.5

* These values of **$\text{AvrBs3}_{111\text{-}42}$** and **$\text{dAvrBs3}_{111\text{-}42}$** binding to the *Bs3* targets represent the average of two independent EMSA experiments performed on the same days as other individual determinations in the table. The values differ from that in Table 1, but were included in the computation of the Table 1 value.

** Percentage affinities of 5' A, C and G were calculated using the values for **$\text{AvrBs3}_{111\text{-}42}$** or **$\text{dAvrBs3}_{111\text{-}42}$** binding to *Bs3* (5' T) targets, determined on the same day, as a reference.

Table S6: EMSA and ATF data for polarity effects.

Protein	Target	Apparent K_D (nM)	% wild type K_A	Fold Activation
III-HDp	wt	0.6 ± 0.1	100	n.d.
III-HDp	5m3	35 ± 3.7	1.7 ± 0.3	n.d.
III-HDp	5m6	530 ± 134	0.12 ± 0.02	n.d.
III-HDp	5m9	>1650	<0.04	n.d.
III-HDp	3m3	1.7 ± 1.0	44 ± 21	n.d.
III-HDp	3m6	39 ± 28	2.4 ± 1.9	n.d.
III-HDp	3m9	68 ± 23	0.91 ± 0.16	n.d.
III-NGp	wt	0.3 ± 0.1	100	4.3 ± 0.45
III-NGp	5m3	4.2 ± 0.7	6.5 ± 2.9	0.35 ± 0.09
III-NGp	5m6	100 ± 10	0.27 ± 0.13	0.18 ± 0.02
III-NGp	5m9	320 ± 76	0.08 ± 0.03	0.14 ± 0.04
III-NGp	3m3	0.3 ± 0.1	95 ± 8.8	3.1 ± 1.4
III-NGp	3m6	0.5 ± 0.2	49 ± 11	5.3 ± 1.4
III-NGp	3m9	2.5 ± 1.1	12 ± 6	1.6 ± 0.33
AvrBs3$_{111-42}$	wt	5.4 ± 2.6	100	3.9 ± 0.32
AvrBs3$_{111-42}$	5m3	115 ± 11	3.4 ± 1.2	1.2 ± 0.07
AvrBs3$_{111-42}$	5m6	>1500	<0.25	1.1 ± 0.15
AvrBs3$_{111-42}$	5m9	>1500	<0.25	1.2 ± 0.17
AvrBs3$_{111-42}$	3m3	140 ± 78	4.4 ± 2.1	1.3 ± 0.03
AvrBs3$_{111-42}$	3m6	490 ± 48	0.77 ± 0.18	1.0 ± 0.22
AvrBs3$_{111-42}$	3m9	970 ± 64	0.58 ± 0.29	1.1 ± 0.08
dAvrBs3$_{111-42}$	wt	3.2 ± 0.7	100	4.7 ± 0.35
dAvrBs3$_{111-42}$	5m3	1100 ± 85	0.29 ± 0.04	0.52 ± 0.05
dAvrBs3$_{111-42}$	5m6	>1500	<0.22	0.55 ± 0.04
dAvrBs3$_{111-42}$	5m9	>1500	<0.22	0.60 ± 0.04
dAvrBs3$_{111-42}$	3m3	46 ± 12	7.4 ± 2.7	1.4 ± 0.06
dAvrBs3$_{111-42}$	3m6	120 ± 23	2.7 ± 0.59	1.2 ± 0.15
dAvrBs3$_{111-42}$	3m9	93 ± 5.3	3.5 ± 0.66	1.0 ± 0.08
n.d., not done				

References

1. Mussolino, C. and Cathomen, T. (2012) TALE nucleases: tailored genome engineering made easy. *Curr Opin Biotechnol*.
2. Miller, J.C., Tan, S., Qiao, G., Barlow, K.A., Wang, J., Xia, D.F., Meng, X., Paschon, D.E., Leung, E., Hinkley, S.J. *et al.* (2011) A TALE nuclease architecture for efficient genome editing. *Nat Biotechnol*, **29**, 143-148.
3. Sun, N., Liang, J., Abil, Z. and Zhao, H. (2012) Optimized TAL effector nucleases (TALENs) for use in treatment of sickle cell disease. *Mol Biosyst*, **8**, 1255-1263.
4. Mak, A.N., Bradley, P., Cernadas, R.A., Bogdanove, A.J. and Stoddard, B.L. (2012) The crystal structure of TAL effector PthXo1 bound to its DNA target. *Science*, **335**, 716-719.
5. Zhang, F., Cong, L., Lodato, S., Kosuri, S., Church, G.M. and Arlotta, P. (2011) Efficient construction of sequence-specific TAL effectors for modulating mammalian transcription. *Nat Biotechnol*, **29**, 149-153.
6. Deng, D., Yan, C., Pan, X., Mahfouz, M., Wang, J., Zhu, J.K., Shi, Y. and Yan, N. (2012) Structural basis for sequence-specific recognition of DNA by TAL effectors. *Science*, **335**, 720-723.