Supporting information

Figure Legends

Figure S1. RACE products and reconstructed su(Hw) insulator transcripts. A) RACE amplification products for *su(Hw)*, *mod(mdg4)* and *CP190*. All genes had a single transcript for the 5' and 3' RACE reactions, with the exception of *mod(mdg4)* 3' RACE, which had nine groups with varying molecular weights. All groups were cloned and sequenced. Each of the groups gave the following *Anopheles stephensi mod(mdg4)* isoform: group 1: isoform C; group 2: isoform H; group 3: isoform O; group 4: isoforms A, D, L; group 5: isoform G; group 6: isoforms K, N; group 7: isoforms I, M; group 8: isoform E, J; group 9: F, P. B) RACE and walking primer sequencing data from *su(Hw), mod(mdg4)* and *CP190* from *An. stephensi.* The 5' and 3' UTR sequences are highlighted in yellow and bold letters indicate exon-exon junctions.

Figure S2. Alignment of Su(Hw) protein complex. Alignment of predicted amino acid sequences encoding Su(Hw), Mod(mdg4) and CP190 from *An. stephensi* and *D. melanogaster*. The red color indicates highly conserved residues and the blue indicated less conservation as determined by alignment using Cobalt (http://www.st-va.ncbi.nlm.nih.gov/tools/cobalt).

Figure S3. Predicted protein domains encoded by the gene orthologues of the *su(Hw)*-insulator complex in mosquitoes. Protein sequences were aligned using the ClustalW algorithm. Identical and highly conserved residues are highlighted in gray. The zinc-coordinating residues are indicated in red. Abbreviations: STEPHE, *Anopheles stephensi*; GAMBIAE, *Anopheles gambiae*; AEDES, *Aedes aegypti*; CULEX, *Culex quinquefasciatus*.

Tables

Table S1. Oligonucleotide primers.

Primer Name	Sequence (5’ → 3’)	Application	Reference
S7-F	GGTGCACCTGGATAAGAACCA	S7 controls	Brown et al., 2014
S7-R	CGGCCAGTCAGCTTCTTGTAC	S7 controls	Brown et al., 2014
Su(hw) F14	CCAATTTACGGGGTTGTTCGC	Genomic and RT-PCR	This study
Su(hw) R14	ACGCCCGGGCACTTGTATTT	Genomic and RT-PCR	This study
Su(hw) F15	TGATACTGGACATGAAGCGCA	Genomic and RT-PCR	This study
Su(hw) R15	ATGCTTCCGCAGCGCA	Genomic and RT-PCR	This study
Su(hw) F16	CCGTGCGATCTGTGCCC	Genomic and RT-PCR	This study
Su(hw) R16	TCCATCGTGGTGATCTGATAGATG	Genomic and RT-PCR	This study
Su(hw) F17	ACCGAGAAGGAAAAGATGGAAAAG	Genomic and RT-PCR	This study
Su(hw) R17	CCCATCAGCATCTGCAGCA	Genomic and RT-PCR	This study
Su(hw) F18	CCGATCTTCCTGAGCGACG	RT-PCR/3’ RACE PCR	This study
Su(hw) R18	CCTTACGCGTTTATAAGGTCTTTAGAAG	Genomic and RT-PCR	This study
Su(hw) F22	GCGATATCGCCCAAGCGAGAAACG	Probe synthesis	This study
Su(hw) R22	TAGCGGCCGCCTACTTGTTCTTCAGCA	Probe synthesis	This study
Su(hw) R25	GAGTCGCCCAAAACTAGA	5’ RACE PCR	This study
Su(hw) qPCR F1	AGGACGGTCGGTTGATACTG	qRT-PCR	This study
Su(hw) qPCR R1	CGGGCACTTGTATTTGTGTG	qRT-PCR	This study
Mod F12	GACTGAGCGGCGGTAAGCG	Genomic, RT-PCR	This study
Mod R12	TTTGTACGATTGAATGCGGTTACG	Genomic, RT-PCR/5’ RACE PCR	This study
Mod F13	CTACCCGCCTTTATCAGTACCGC	Genomic, RT-PCR/probe creation	This study
Mod R13	AGCTCAATGTATTCCGGCTCAATC	Genomic, RT-PCR	This study
Mod F14	ACGAAAATGTCCATCAGTGCTAG	Genomic, RT-PCR/3’ RACE PCR	This study
Mod R14-1	TCCACGAATCCAGATACGCCAG	Genomic, RT-PCR/probe creation	This study
Mod R14-2	AAGACCGCCAAGGCTGGAG	Genomic, RT-PCR	This study
Mod F15	GAACTGGCACGGAACAATC	Genomic, RT-PCR/3’nested RACE PCR	This study
Mod qPCR F1	AAGCAACTCCCTGCACAAGT	qRT-PCR	This study
Mod qPCR R1	CTGTTCCGTCTCAGCATCAA	qRT-PCR	This study
CP190 14F	CATCCAATTTCGGCAACG	Genomic, RT-PCR	This study
CP190 14R	CAGCACCGGTTGCTTCACAA	Genomic, RT-PCR	This study
CP190 15F	GACATGAACATGACCGTGCT	Genomic, RT-PCR/probe synthesis	This study
CP190 15R	GCCGGGTGATTTCGATTCCT	Genomic, RT-PCR/probe synthesis	This study
CP190 16F	GATTCGCCCGACTACGTGGA	Genomic, RT-PCR	This study
CP190 16R	ATGGTTGTAGAACGCTTTGTAGCT	Genomic, RT-PCR	This study
CP190 17F	GATTCGCCCGACTACGTGGA	Genomic, RT-PCR	This study
CP190 17R	ATGCCGGACGCCACATT	Genomic, RT-PCR	This study
CP190 18F	CTGTCCAACATTACGCTGCC	Genomic, RT-PCR	This study
CP190 18R	TCCGGCTTGTCGGACG	Genomic, RT-PCR	This study
CP190 19F	CATAACGATAGCCAAAAGTCCG	Genomic, RT-PCR	This study
CP190 19R	GTATCCTCCGACCAATCGCC	Genomic, RT-PCR	This study
CP190 20F	ACGATTCGGAAGATCGGCAG	Genomic, RT-PCR	This study

CP190 20R	TTTGTCCTTATCCTCTTCGTCGC	Genomic, RT-PCR	This study
CP190 24R	ATCGGCACTATCACGTCGGC	5’ RACE	This study
CP190 26F	CTAACTACAACAGTCACAGAACTGGC	3’ RACE	This study
CP190 qPCR F2	GGACGAGCTCACAGAGGAAG	qRT-PCR	This study
CP190 qPCR R2	CCTTCACACCGGTCACTTTT	qRT-PCR	This study

All primers were oligomer ssDNA strands synthesized to order by Sigma Aldrich (sigmaaldrich.com). “RT-PCR” refers to the primer’s usage in the primer walking strategy.

Table S2. Detailed analysis of the transcript-specific sequences of the *Anopheles stephensi mod(mdg4)* transcripts.

***Anopheles stephensi* name**	***Anopheles gambiae* homology[1]**	**Intron 3-4 (bp)**	**Exon 4 (bp)**	**Intron 4-5 (bp)**	**Exon 5 (bp)**	**Intron 5-6 (bp)**	**Exon 6 (bp)**	**3’ UTR (bp)**	**cds length (bp)**
Mod A	None	7091	88	33	141	95	200	138	1519
Mod B	None	7903	356	-	-	-	-	47	1446
Mod C	R-A	8442	406	252	22	-	-	243	1518
Mod D	R-N	9558	392	-	-	-	-	102	1482
Mod E	None	13,827	332	-	-	-	-	41	1422
Mod F	None	14,897	218	-	-	-	-	105	1308
Mod G	None	15,368	109	74	259	-	-	208	1458
Mod H	R-K	16,415	458	-	-	-	-	381	1548
Mod I	None	18,799	320	-	-	-	-	100	1410
Mod J	R-R	19,380	356	-	-	-	-	106	1446
Mod K	R-L	22,232	176	-	-	-	-	317	1266
Mod L	None	24,598	371	-	-	-	-	211	1461
Mod M	R-T	27,919	323	-	-	-	-	89	1413
Mod N	None	28,639	143	-	-	-	-	305	1233
Mod O	R-I	35,113+ poly-N[2]	461	-	-	-	-	231	1551
Mod P	003438-RA	44,010+ poly-N[2]	287	-	-	-	-	95	1377

[1] *Anopheles gambiae* homology refers to transcript-specific regions in *An. stephensi* transcripts with a low e-value (<1e-10) hit against the *An. gambiae mod(mdg4)* transcripts.

[2] Poly-N refers to a poorly sequenced region of the *An. stephensi* genome that was an indeterminate length.

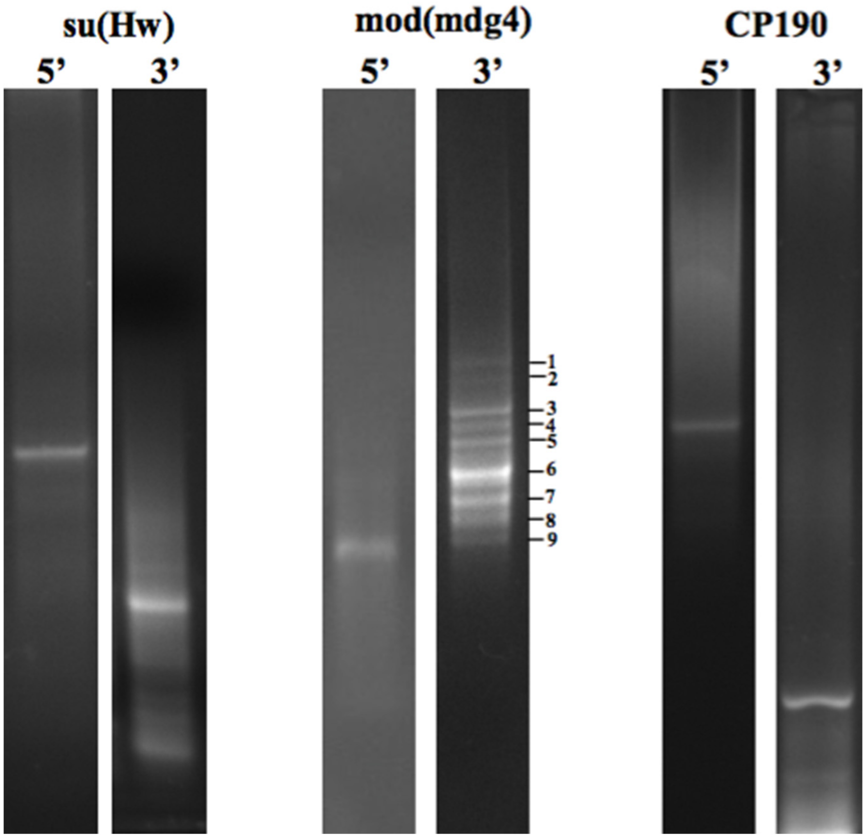
Table S3. Transcript abundance (ΔΔCq) and statistical analyses of differences in *As-su(Hw)*, *As-mod(mdg4)* and *As-CP190* mRNA abundance in embryonic, larval, pupal and adult developmental stages, and in ovaries and midguts of adult females.

Sample	**Su(Hw) [p-value]**	**Mod(mdg4) [p-value]**	**CP190 [p-value]**
Eggs	1.80 [N/A]	3.79 [N/A]	13.07 [N/A]
L4	1.02 [0.980]	1.02 [**0.0006**]	1.03 [0.070]
Male pupae	1.29 [0.999]	1.63 [**0.008**]	2.41 [**0.007**]
Female pupae	1.24 [0.998]	1.32 [**0.002**]	2.43 [**0.007**]
Male	5.54 **[0.001]**	4.60 [0.814]	6.68 [**0.000**]
Female (adult, sugar fed)	1.14 [0.993]	2.00 [**0.042**]	1.52 [**0.036**]
FEM 6 hPBM	1.55 [1.000]	1.69 [**0.011**]	1.61 [**0.031**]
FEM 24 hPBM	2.17 [0.999]	4.98 [0.371]	5.02 [**0.000**]
FEM 48 hPBM	4.50 **[0.032]**	10.26 [**0.000**]	7.75 [**0.000**]
FEM 72 hPBM	5.50 **[0.001]**	10.75 [**0.000**]	10.54 [**0.003**]
FEM POST-OVI	2.30 [0.999]	2.41 [0.589]	3.38 [0.095]
Ovaries	3.84 [0.543]	7.65 [**0.0005**]	4.40 [**0.015**]
OV 6 hPBM	2.75 [0.990]	5.24 [0.526]	2.00 [0.329]
OV 24 hPBM	2.74 [0.990]	7.07 [**0.0032**]	4.03 [**0.036**]
OV 48 hPBM	5.49 **[0.028]**	13.07 [**0.000**]	12.11 [0.431]
OV 72 hPBM	4.58 [0.179]	11.24 [**0.000**]	11.88 [0.186]
Midgut	2.12 [1.000]	1.60 [0.093]	2.74 [0.189]
MG 6 hPBM	1.78 [1.024]	1.30 [**0.038**]	1.88 [0.381]
MG 24 hPBM	7.87 **[0.0001]**	2.11 [0.338]	5.12 [**0.002**]
MG 48 hPBM	2.12 [1.000]	1.01 [0.125]	2.42 [0.297]
MG 72 hPBM	2.04 [1.000]	1.88 [0.542]	1.91 [0.417]

P values calculated with a *t* test. *P* values ≤0.05 were considered significant (shown in bold). Abbreviations: L4: 4th instar larvae, FEM: female, hPBM: hours post blood meal, FEM POST-OVI: females post-laid eggs, OV: ovaries, MG: midgut.

Figure S1 RACE products and reconstructed *su(Hw)* insulator transcripts.

A) RACE amplification products for *su(Hw)*, *mod(mdg4)* and *CP190*.



B) Reconstructed *su(Hw)* insulator transcripts.

su(Hw)

GTGTACAAGTGAATTTGATCGCAAATATCCACCAAAGATGCGCTTCGGTGGTGCCAAAGTGG
TTTGCTAGTGGTTTGCTGACCGATGGTTGTGCATCAAACATTCCTACACACAGTTCTTTCGGT
GCGTGCGATTGGTTTGGCTGTAGGCGCATATTGTGTGACCAGAACGAGCGTGCATGCTAATG
CCCGCCTTCGCGCGCGCTTCGCGCGCGCGGCCGTGTAAGACAGTAAACATCGCGTTCTGGTT
TCAGTTGGCGTTAGCGTTGGTGCTCGTACAAAAAACAGGGGAGAATCCAATTTACGGGGTTG
TTCGCGCTTGGATCAAAGCAGGCGACGTTTCTAGTTTTGGGCGATTTCCCCATGATGGTAGG
CAGTCGGCGATCCGTTGTCGGTGGAGTCCGTAGCGGTGGGGTGAAGAAGGAAGAGATTTTG
GTAAACGTGCAGGTGAACGGTAACGACGAGATTGACCTGGTGCCGGTCAACAAGAGCGCGG
ACTATGGGAGAGGATCGAGCGACGATCGAGACGGTGATGAGGAGGAAGAGGAGGACTACG
AAGATCTGTCCGACAATCCGGACTCGGATCCGGATCTGCTCGAGACGGAGGACGGTCGGTT
GATACTGGACATGAAGCGCAGCAAGTACTGTGGCAGGTGTAAAAAGATTTTCAAAACATCA
GC**GG**AACAGAAAGCACACAAATACAAGTGCCCGGGCGTTGTGGATGAGGTGCAGGACGTC
ATCCTGAAGGAGCGTAAGGCGGCCAAGCATGCGAACCCGGGCGAAGAGTTCCACAAGTACT

GCAATCCGAACCCGGACAATCCGTGCTACTGCTGCGGTGAGGATGTGAGCACCGCGCACGT
CGGCCACATCCGCTGCAAGTTCTGTCCCAAATCGTTCAAAGCGTACGAGTACATGGAGCGGC
ACCTGAGCTCGATCCACTCGGAGTCGGACGCATTTCCGTGCTGCTACTGCAACGCCAAGTGC
AGCACGCAGGAAATTCTGGACGAGCATCTGAAGACACACGACGAGGGCAAACCGTACGCGT
GTATGGCGTGCGGTAAAGATTTACCCGCCGCTACCATCTGGACCGCCACGAAAAGCACACC
AAGTGCGGGATCGTTCCCAAGGAGGTGGAGGTGCTACCGTGCGAGGTCTGTGGCAAGGAGT
TTACGCGCATTGACAATCTGCGGGAACATTTGCGTTCGCACATGGGGCAAGGGGCACGCAA
GCGGGATTACCAGTGTCCGTACTGCTCGAAGTCATTCTACGGCTCATCGTTGCTGAACATAC
ACATACGCACGCACACGGGCGAGAAACCGTTTCCGTGCGATCTGTGCCCGATGTCGTTCCCG
TCGACCGGTGCGCTGCGGAAGCATCGGCGCAAGCATACGGGCGAACGGCCGTACCGCTGCG
ACGAGTGTTCAGCAACGTTTGCCGCGCGTGAAACGCTTAACCGGCACCGGAAAACGCACAC
CGGCGAGAAGCCGCACGTGTGTCAGGAGTGCGGCAAACGGTTCATCCAAGCGACACAGCTC
CGGTCCCACATGCTCAACCACAGTGGCGACAGTGCGTTCCAGTGCAGCGAGTGCGATGCGA
CGTTCAATCGCAAAGCCCGGCTCACGGAGCACATCAAGTTCACGCACAAGGGAGAGCAACC
GTTCGAGTGCGAGAAATGTTCCAAAACGTTCCTGCGCAAGGATGATCTCGCACGCCACGAGC
TGATCCATTCGGACGTTAAA**CC**CTACGAATGTTTGGTGTGTAGGAAATTTTTCACCACCAAA
AACGCTATCCTGTTGCACCAGCGGACGCATCTGGAGGAAAATCCGGTCCAGTGTAAGGTTTG
CTCCGGCACCTTCAAACGGTCCGACTGTCTGATGCGGCACATGCGTACAAAACACC**GC**GAC
TTTTTGGACAAAATTATCGACGATACCGAGAAGGAAAAGATGGAAAAGCTGACCACCACCA
AGCGAGAAACGGAAAAGGAAGCGGTGGTGGAGATAGACGGTGCCATCTATCAGATCACCAC
GATGGACGATGCACCCGCGAAGGGGACCGCTTACACCGGCAACAACCAAACGGTGGAGGTG
TTCGAGGTGGTAGAGCTCCCGGCAAGCGCTTCGGACGCGATACCAAACACACTGCCGACCA
CCTCGGTACCGTACGAGCTGGTACCGGTTGTGGATGATACGGAAGTAATACTGGGCGACAG
CAGTCCGGTAGAGTTTGTCGGTAAAGCAAACACTCCAATCGACGTGGTTGTGAAGAAGGAG
GAAGATTCGTTCAAAAAGGTAAAGTCCCCGCCCAAAGCGACTGTCGTTTCAAAATCAAGCG
AACCCGTTGCACCGAAGGTTGTTGGTAAAGCTAGCGAAGGAGAACCGTCGCGCAAGAGCGT
TCCAAGCACTTCAAAAACAACGGCCACGATAAAACCGAAGACGGAACATGTAGTAGCTGTG
CGTGCGAGCAATGAACCACCACAGAAAACGCCACCGAAGCTAGCGAAAGAGATTGCGCGCG
AAAAGGAGGAACCCAGCACCAAGCTGGCAAAGAAGCCGGCTACATTACCAACCGCCCGACC
GAAAGAGGCCAAGAGCAGGGCGGTGGAGGAAGCGAAGCTGAAGCGCAAACGTGAACCCAA
ACCGGCCACCCCGAAGCGAATGCGCCCGGACGAGGACAGCATTCCGATCTTCCTGAGCGAC
GATATGCTCGAGCAAAAGATTTCCGAACTGCTGCAGATGCTGATGGGGGAGGATATGTTGC
GCAGCTTCGGTTGGCCGAACGCACCGGTCGAGGTGGTGCTCGGGCGCGTCATCCAGGGCTGC
GGTCATCAGCCGGCAGACTAGGAGGCCGGTGACCATACGACGCGCATGCGAGAAAACACCA
AGATACTGTTTTCCGTCACGATGGACGATGACGACATTAAGGCTTTGCTGAACAATCACACG
GTCGACGAGGTGATCATGCACGTGCTGAAGAACAAGTAGTTCGCGCGCGCGGGGCTAATCT
TCAAACGATGATGGTACACGGTGCAATGGGTTCCAGATTGATCGTACAGTTGCAGTCTACCC
TTCCTACATCCTGTAACGCAATGTGGCCTTCTAAAGACCTTATAAACGCGTAAGGTTATAGA
GTATTGCTTTTTTTTTATTATTCAATGAAGATTTATCAATGTTTTTATTAAACACAGTATGCAA
TGTATACAATAAACAAGAACTGAACTAACTCCACAC

mod(mdg4)

Mod-RA

CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT

CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GC**GGACGAAAACTTGCTGTTCGGTCTGGACAAAAAGCGCACC
ACGCGAACGGCGGTGGTGTGCGGCGGCAAACACTTGTTCCTCACCA**GC**CACTTTCT
GTACCGCTCGAACCTGCGCCGGCAGGGCCGCAATGGCGACGTCATCTACTGGGAGT
GTATCTACAACCGGGGCCAGAAGTGTCGCGGACGGCTGAAAACGATAGGCAACAAC
ATCTACATTTCGAACGGCA**AA**GTGGAGCACAACCATCCCAACGACCGTAAGCGTAT
CGTGAATGCGACACTGTGCCGAACCATCATGATGAAGGACATAATGTCTCTGTAAA
AGGACATCCGCCGACTCTTACAGCAATGTACAACGGTCGTTTGGGCTCATTTCGAAA
TAAAGCAGTTTTATGTTTTATGAAAAATGAAGCAAAAGGCTCTCTTTCTAGTTATAT
CTTTACGGTTTTCCCATTTACAAACATCCTTTTTTTACAAAGTATGATGAATGTTGGT
GAAGCTGTACATCGTTTAAAGTATGGCCCGAGGAATAGGGGAATTTTTACTAAGAAT
AAAAGCAATGGGTATAC

Mod-RB
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG

GAACAATCGACACAA**GC**ACGGACGAAATATCGCCATGGCAGACAGTTAGCATCACCG
ACGGCGGTAAAGCGCGATCAGGATGAAGGACCGCGCACCGACAGGATGATGTTTGT
CCGCAGTTCGTGGGGCCGTAAGCATCTGCTGTACGATGGGTACACGTACGCGTCGAA
CAAGCGGGTCGAGTCGACCAACACCACCTACTGGCGCTGCACGTTGTTCCATCGGCG
AAATGTGCGCGTACCGTGCGGGGTGCGCTGTACGGTGCGCGACGGGCAGATCATCA
AAATGTACGGCAAGCATAACCATCCGCCGGAAGTGCACAGCTTCGATGGCAAGCAG
GTCGAAACCGATTACTGTATACATGACCCGTGAAGAAATAAAGTTTACGAGGGCTTT
TCGGTAAAGGAAGGTGTAACTAC

Mod-RC
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GC**TGCTCCGACTTTGGAAACGAACAGCAACGGACCACCACCG
TCCGGTTGTCTGACCTCACGGACGGCTGGCACTAACGTTGAGCTGAAGTTCATCAAG
AGCCCCTGGTCCACGCCGTGCCTGGTGTTGAACGACTTTCTGTACAATGCGCACAGT
ACGCGGGGCGAAATTGGCTACTGGCGGTGCCACAACTACTCGCGCAAGGTAAAGGA
AGAACGGTGCCGGGCGCGGTGCGTGATCAAGAATGGGCGGCTCAGTGCGCTGACCG
GTGCCCAGCACAATCATCCACCGCACACCGAAAAAATCGAGCGCATCGTGCGCCGC
AATCTGGCGGACGAGCTACAGGAGCGGGAAATGGCACGCAACCAGCAGCACCAGG
AGGCGGAGGCTCTGCAGTGTTATCAGC**CA**GTGGCCCAACAGTTCCCATAGCCCCCA
ATAGAATGTGTGGCGCGGGTGGGGTTAGCCAGCAGCAGGAGCAATAATCATTGAAA
ATTTGTGCAATAAAACAATGTGAAACGGTGTGCAATATTGTGGTGCATTCTGACGGG
GAGCACTAGGGCCAACACTTGGGCCAACACTGAGCAGTATTTACCTCTGAGCCAAG
ATCCTATCTTAAGAATAAAGCTAATCTATTGTTTCGTGTGCGCGCAGGAAGGTTGTG
CCTCATCAGC

Mod-RD
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT

TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GA**TTGTTCAGTGTTGCATCCGGTGGCAACGGAATCCAGCGTG
CAGGGAGGCCACATCTACATACGCAGCCAGCGCGGGTTGCCACTGTTGGCGAAGAA
TGACTACATCTATCGGTGCGAGCGCGCGCAACCAACGTTCGTACTGGCTGTGCAT
CCGGTACAAGACGCACCGGTGCAAGGGCCGGCTGATCTGTCTGAAGAACACGGTGC
TGAAGGAAACCAAGCACTGCCATCCGGACGATCGGCGCCGGCTGGCGATGTCGGAC
AAGTGCGTGATCGATCTGTCCGGCATCAATATCAACGCGTGGGTTAAGTCGGTTGGG
CCACGGAAGCCGGTGCGCCGCGCCAGCGCCAGTGCGCCGGCGGACCTTGTTACAGG
CTTGTGCTGTTGATGTCAAAAGGGGGGCAAAAACAGGAAATGATTAGCGATAAGA
TTCGTTTGAAGAGCATTTTTGCGGGCTATTTGTGAAATAAATAAAACCACTTAATGA
CG

Mod-RE
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA

AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GT**GGACAAACAGCTGACGAAGCGTAAGGGCACCGAGCTAAT
CTCGCTGATCGACAACCAGGATGTAGCGTTCGGCGAAACGGACAACGGTGGCAAGC
AGCTGACGTACCGCAATTACATCTACCATCGGAACCTAACGACGGGCAGCACCGTG
TACTGGCGCTGCTCGAAAGCGATGCGCCTCAAGTGTAAGGCGACGATCGTGACCAG
GGGCGATCTGATGCGTGTCAACAACGTCGAGCACAATCACGTCCCGATGCGCCGAC
TGACGTACGGATTGGGTGTGAACGAGAAGCGCAAGCTGAAGAAACGCAACGATCCG
TGGACTCGATGAGCGCGCAATAAAATTGGTTTAAACCTCTCGCAACAACCGAA

Mod-RF
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GA**ACCAGCACGCTACATTGTTGGAGTGCGGGGTAGCCGGAAG
CTGAAGGTGGGCGACTATACCTTCACCAGGAACAAGGAGTGTTCGGACAAAACGTA
CTGGTCCTGTGCCCGGGCGGGCATGCATCGGTGCAAAGCGCGCGTCCTCACCTACAA
CAACAAGAACGGCGAACAGACGTACATACTGCGCAACGGCTTCCACAATCACGAGC
CGTTCTAGTGCAGCATTGCTTTGCTTGTGTGTACAATAAAGCTTTTTCGGTCGGGCGG
GCATGATTGCTTTATGAAAACCAATAAAACGATCTTAAGACACACATTTTACCAC

Mod-RG
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT

CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GG**TGGTGTACATAGTCGGGCAGCGTGGCAGCATACTGCTCTC
GGTGAATGGGCATCGTTACGTGAAGAACCGCAAGGGCCACACGAAAACCTACTGGA
TTTGCGCGAAA**AA**GCGTTGGATGCCGGGCGCGCCTTACAACGACCGTGTCGGAAAA
CGATCACCGCACAAGCGCCGAAGCACCGAAAATCATTCTCATGTCCGGCATTCACA
ATCACGCCCAGCCGACCAACTTCCGGCGCGTTAATGAGGGCAGGCAACCGACCACC
GTACCACAGCCGATTCGAGCCGGGCGGATACGGCCGCGGAAGGGAGGGCGTCCA
ATCAGCAGCCCCCCATACTGTCCACCACACCACCGCAGCGGATGCTGATGAAGGCG
AGAATGATTTCCCTGCAGTTGTCCTGTGGCAGAGAAGACGGGGGCCGGAAAGAGTG
AATCACCTGCCCTGCTGAAATGTCTTCTTTCCCCTGGAGAAAGGCACCCTGGCAGCC
ATGGCAGCCGGAAGAGGTGTTGTTTGCTTTAATAAGGCCCAATAAAATGTAATAGG
GCGCGCTCGTGCGTGCGTGCTATACTGAACC

Mod-RH
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG

GAACAATCGACACAA**GA**TCGCGCGCGGGCAACTAAACCAACGGTGCTGAACGGCAA
GCAAAACTCCGGTGCCGTTGTGACGACGTCGGTGAAGCAGCCGCGCATTCCGAAAC
ACGTAAAGTCACCCAAAATGAAGCGCACCATCCGATCCCAGGGGGGAGAGGAGGA
GGAGGAGCAGGAACCGTCGACACTTGCATTCCGTGTGCGCAACACGCCGGTCGCGG
TGGACGAGACCAAGATCCAATACACTACCGGGCGCGGTAACAACGTGCTCGTGTAC
GACGGCCATCGGTACATCAAGAACAACTGCTACGGTGGGAAGGTGTACTGGAAGTG
CAGCAAGTGGCACACGATGTGCATGGCGCGCGCAATAACCTCGATCCTGCATCCGA
ACCAGTGCGTGCTGAAGAACTCGCACAACCACGGCCCGCTGAAGGAAGACATCACG
CTGGATGTGCCGCAGGACTGGCGATAGATTTCGTTAACCGAGCTAAAAGGACAAAA
CCGCTGGCCCTTTTTTTGTTGCTTCTTTGCTTGCTTTTGCTTTATACACAAACCGATCC
GGTGGCCGAGGCGATAGCGTCGCCGGTCTTCGCTCGGCAGGAAGGGCCGGGATTCA
AATCCGATCCAGACCCCCGTCTCCCGGTACGCAGGGCTCTGACTACTTTACTACGGG
TAAAATTAAGTCACAGAAAGCCAGAAATGGCTGGCAGGCCGGAGACTTCTCGAGGT
TGTGTAGTGCCAAGGAAGAAGAAGAAGAAGATCACTCTGTAATGATTATTGGC
GTGTAGGTTAAGTGTGAGAGGAGGGGTATGGTCAAATTCCGACCTACAATAAATCC
TACATACAGGCT

Mod-RI
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GA**AGTGATTAAGCAGGACGCCACCGCCTCCACAACGATCTGC
TTCCGGAAGAGCCAAAAGGGCAAGCTGCAGCTATCGCACGACGGGTACTACTACTG
TATGGAGAAACGGATACAGCGCAAAGAGTACTGGCGCTGCATTTACTACACGACGA
AGATCAAGTGCCACGGGCGGCTGCACCGGGACGACAGCAAGGTGCGCTACATGGGC
GCACACAACCACGCGCCCAAGCTGTTCAAACGGGCCGACTACAAACGGCTGTCCGA
GATTGTAGGGCTGTCCGCCGAGGACCAACCAACGACCGGCAAGACGCACACTTGAA
CACCAGCTGAACAAGGATTACTTTCTGAGAGAGAGATAGAGTATAAAGCGAAACAA
TAAAGCGTTACACAATAAACGATCATGCCGAACGTCCCACATC

Mod-RJ
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GC**CGCCATTCTCACCGCAACGCCTGAACCGCACTCGTCGTCG
TGGATCGCGGTGGAAAAGTCCGTGCTGCAGTACGCGACAACGCAGCGCGGCCGAAC
GATGCTCATCTTCAGCGGCTACAAGTTTGTGGAAAATCGGCAGAGCAAGCGGAACA
TCTTCTGGCGCTGCGCCCGCTACGTGAAGCACGGTTGCCGGGCCGCTTGCGTGACGT
CGAAAAATGGTGCCGGCACCCAATCGATACGGTTGACCGGGCTGGCCCATAGCCAC
CCGCCGGAAGATGGACGAACCACGGACAGGGCGAGCGATCAACCGTCCACCACCAC
CACCACCACCATGGTGTCGGTCGTTGGACAGTAGATCGAGCAATAATGGTAGGAAA
GGGCGAGCCAACGTGCGTGTGAGTGCATTTATTTTCCGCAGCCGATTTTCAGCGCTT
TTTTTAAAATAAATAATTCAAAGTACG

Mod-RK
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA

AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GT**ACGAGCAGGATTCCCGCTTGCTGACGCCCGAGTTTTGTGT
CACCCAGCGCGGTCATCCATTGCTAGTGCACGGCGGCTACAGCTTCACGCGCAATGG
CGAGTTTTCCGACACGATCAACTGGCGCTGCTCGATGTACCGGAAAAGGAAGTGCA
AGGCCAAAGCGATCACCGTTAAAACCGATGGGCAGGATTCGGTGAAGCTGACCCAC
GCGATACACAACCATCCCCCCGGGCTTCCACGGCAGAAACACATCAAACACTGCTG
GACCGAAATGCTCACACTGGCACCGACGACGACGACGACGGCGACGGCAACGACG
GGACCGCTCGCACAGCTGGACGATGCATTCGGCGACATTATGGGACTGGCGGATGA
ACATGACTTTTTCATCACCATTAAGGACGAATCGCTCCACATGGATTGCATTGAAGA
TTCACTGTAAGGTGGTTGTCGCAGAGATGAAAGAGATTAAAGATTCAAACTTGTGA
ACG

Mod-RL
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GA**TCCCGAAGCACCACGCAACATGGTGCTGAACGTTAAGGCC
GAAGGGATGAAACAGCAGCCGCCAAAATCGGGCACCGCCACCAGCCTGGGCCTGA
AGAAGGTGTACAAAACCGATGACGATGGGGAAACGATCGATGTGGAGCACAAGTTT
GCACTCGATTTGCGGCTCGAGACGGGTTGCAAGGGCCGCCCGAAGCTCATCATGGG
GGGGTACGCGTTTTTCCGCAACAACAGCAGCAACAACAAGACCTACTGGCTGTGTTC
CAAGAGTCGGCTGCTGAAGTGTCGCGCCCGCATCATCACCTTGGACGGGTGCTCGG
GGATGATACTCAAGAATCAGCTCCACAATCATCCACCCACCGAGCTATAATACGAC
GGGCCGGCGGTGGCTCTGTTGCTATACTTTTGTTTTTTGAATTCCCTCTTAACGAACT
CTGTTTTGTAGTTTATATCTCGCAGGGGGCCATATATTTGAAAAGTATTTTTCGTCGC

CCAGTTTATGCAATACAGTGAATTATGTTTAGTAATAGTATGATCATCATCGGTTCAT
GAAATAAAAGCGTACGTACTACTGCAATCAC

Mod-RM
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**GG**TGTTGATGTGTCCAGATCCATTACGCTATACCGTTCCCGGC
TTACGTTCATACGAGGACAGCGCGATCACAAGCTGCTGGTGGTTGGCGGGTACACCT
ACGCACGGAACAATTCTGCCAAGAACACCCCCATCTACTGGGCCTGCCGCACGTCGT
ACGCTAGCGTGCGGTGTAATTCGCGCGTCGTGACTACGCACATGGCGGACGGCACG
TACCGCATCGTGATCACCAACCCACGGCACAACCATCCGCGTCGGGTCGCACTGAC
GAACGAATATTTGCACAGCTTAAGCCTGCTGAACCGTCCATTCAATCGGAAGTGAAC
GGTTGTGGTAGGTGAATGACGCCGAACTACTCGCGAGAACGTAGCTTAAACTGGGT
CAGCTTTCGGAAACGGAAAAAGAGAATAAAC

Mod-RN
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC

AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**G**

Mod-RO
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA
CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**G**

Mod-RP
CTAATACGACTCACTATAGGGCAAGCAGTGGTATCAACGCAGAGTACATGGGGGCGAAGTA
ACTGTGCTAGGAGCTGGATGCCGTCGCTCCGCTGTCCGTGTTTGTGAAAACTATTTTATTCTT
TAGACTTAGTGTTTAGTTACACTAAAATTCAGTGTATTGCGTGAATTTAAGCGTCATTGACTG
AGCGGCGGTAAGCGGAAGCGTCGTCCAAGATGGCGGACGATGAGCAGTTCTCTCTATGTTG
GAACAATTTCAATTCCAACTTATCGGCGGGATTTCATGAATCCCTGCAGCGTGGTGATTTGG
TGGACGTGACGCTGGCTGCTGAAGGTCAGCTAGTTCAGGCACATCGCCTCGTTTTATCAGTA
TGTTCCCCGTATTTCCGGAAAATGTTCAACCAAATGCCTAAGAACCAACATGCTTTC**GT**TTT
CCTGAAAGATGTTACCCACTCGGCGCTCCAGGATCTCATACAGTTTATGTACTGTGGCGAGG
TGAATGTGAAGCAAGATGCCCTACCCGCCTTTATCAGTACCGCCGAAGCATTGCAAATCAAA
GGATTGACAGAAA**CG**GGCGATAGCGCACCAGCACAACAGTCGCCTGTCAAAGAAGACCCG
CCGGCGGTCCCTGTGGCGACAGCCAGTATTTCCACCACCGTCCCCGTCACTCCCCAACGAAC
AAAGGTTCAACGTAACCGCATTCAATCGTACAAATTAGAATCGGAAGAGAGTGGCGATGAC
AAAATAGTGCATTTGCAAGCCAGCACATCTCATCACGTGACGGCCCAGCAGCTTCAGGCCCA
GACAAGCGCTAGCTCTCAGAAACGTGCCATGCCACAGCGTGGCGTTCAGAGTCATGCAAGC
AAGCGGACGAAAATGTCCATCAGTGCTAGCAGCGACGGATTGGACACATCGGACACGGTAA

CCACACCGACCCAGGCACAGACCATACAGACAGTCCAGATAGTGAAGCAACTCCCTGCACA
AGTGATTGAGCCGGAATACATTGAGCTACCGATGGAATCTGTCAATCCTAAGGCTGAACCAG
ACTACACGGACGAGGCAGCCGAAATCGAAACGGTTGATGCTGAGACGGAACAGGAACACA
AGCTATCGGAACATGAGCAAGGCGATGCGGACGATGATGGCAACTACGTTGAGGATGAAAC
GTACGGAGACATGGGCATGGGCAAGTACGAAGAATCATACCTAACGGAAGGCGAGGAAGG
AGCGAAACCTGGCGTATCTGGATTCGTGGATTCGTACACATCGGACGGAGGAACTGGCACG
GAACAATCGACACAA**G**

CP190

NCCATGCTTTCCCNTAAGCAGTGNNATCAACGNNCAGAGTACATGGGAAACAGCAAATTTC
GTTCGTTTTTAAGTCAACTTTTGTGCGAGTTCTCGAAACAAAGCACAGTTTTGATTTGGTGAA
GTGCTGGCAGAGNTGAAAAAAGCAGTTAAAAAGGACGCGCAGTGAGCACAGTGCATGTGA
ATGATGTAATTAAAAGATGACGCGTTACGCGGAGGCCAAGTCGGTGAAGGTGGACAACTGG
GGTGTGTTTTTCCTGCAGAAGCTGCAAAACTTCTTCAACAAAACCGACCACTGCGATCTGAC
GCTACAGTTCAACGACAACTCGCAGCTGAAGGTGCACCGGCTAGTGCTGTCCGCCTGTACGG
ACTACTTTAACCATCTGGAAAACGTGTGCGAGATGTACGACGACGTACTGATCATGCCGATG
GATCTGCAGGCCGACGTGATAGTGCCGATTGTGAACTTTATGTACACCGGCACGCTCGAGTT
CCAGTACAATATGTACGACCGGCTGCTGAAAACGGCCACCGACATGAACATGACCGTGCTG
CTGAAGCTGCTCGAGGCCCATCGGCAAACGTCGTCGCGAATTGTGAAGCAACCGGTGCTGTT
GAACAAGCAAGCGCCACGTCCTCGCGGTGCTCCAGTGTACCAAGGCGCTGGGCCGCGCGTT
ATGCCCGCACAGAGAGGCGCGAACGTGGGACCGAGACCGCGATTACCGGTGGCACGCCAAC
CGATGCATATGTACGGCCACAAAGGGGCCGGTCACAGCACAGCAAAGGACGTCAAGCCAGG
TCCTTCGCGGTTCGATGACGGCGAAATGGGCGAAGGTTTCGAGGGATCGTTCGATGGCATT
ACGTACGAAACCAAACCACTGCTAACAGCCGACCAGGTGAAAAAGGAAGAAGAGACATCT
CCGTTCGAAAAGCTGCGCAAAGGCTACACCAACGCAAACGTGGTGAAGCGATCGTCCAGC
GGTTCGCTCACATCTCCGCCGGCCAAGAAGCCGAACCTGGATGAAGTGAAAGAATTCACC
GAAGCGGCCCGCATGCGAAGCCAGCTAGCGACGGCGAATGATGATGATTCGCCCGACTAC
GTGGATGACGATACCCACTTTAACGACGACGATGATGAGGATTATCAGCCACCGGCTTCG
GTAAAAGCCGCTTTGGCAAAGTCCGCCAGTCACCACGCGGGAACGGTCAAGCACGGTACC
ATCAAGCAGGAATCGAAATCACCCGGCTCGGGTGGCAATTCGAGCGCGATGAAACAGATC
ACGGTAAAGGACGATTCCGGCAGCGTCGATCATGCGAAGATCATATCGGAGGTGCTGAAG
AAATACCCCCATCTGGTGAAGAAGAACAAAAACATCAAGCTAAAGATCATGCAGAAACCG
TCGCCGAACAATGCGAACGCAACCAACGCCAGCAGCGCCTCGTCCAGTGCGGTAGCGACA
CCGAAGATGGAAGTGCGTACCGCAACGCCTTTAAAGCAGGACCTCTCGATGGTGCGCCGC
GCGACCGCAAACATACGCAGCGAAACAGCGAAACCGTCAGCCGGGTCCGCGACGTCCACC
GGTAAACCGACCACCGGCACAACCAAAACGATCGACGCGAAAACGATGCACGCGCTGATC
GCGAAGGGTGCGGAAAACACCACCGGCCCTTGGTTGTGTTTGCGGTGCGGCATCAACGGC
CGTCCGATCAGCATTCCGAGCTACAAAGCGTTCTACAACCATTTGATCCACAAGCACAAG
GAGCGCATCGATGTGCGCATCTGCGAGCACTGCGGGTTCAAGGCGCAGCAGCGCAATCCG
CACCTGATCTACCACTGGTTCACCGAGCACAAGGTGAAGCCCACGATCAAGTTCCCGCGG
TGCGACGAGTGCGATCACGTCGCAATGACGGCGGACGATTTGCAGAAGCACGAGGAGGAG
GCGCATTCCCGGGGCGACCTGCAGCAGTGCATCTACTGCAACAAGGTGTTTGCGAAGGAG
ATGGAGCTGTACGATCACATGAAGGAGCAGCACAAGGCGCGTGCTATCGCGGACGGTGTG
CTGGAGTTTACGGACGATGAAGACTACCAATCGCAGGAGGAGGAAGTATCGTCCCCCCAG
GGGGGCAAAACGTACTCCGCCCAAGGGGCTAGCGGAGAAGGTAAGATCAAAATCTTGTCC
AACATTACGCTGCCTCCGGCGAAGGGATCGAGCTCGGCGGCGGCTGGTGGTTCGGTGCGT
CACATCACGCTCGAACCATCGTCCGAAGCGGAAGCACTGAGCAATGTGGCGTCCGGCATC
GCCACCAGTCTGACGCTTGTCGGTGACAATGGGGTCGTGATCGACGACCCCAACTACCAG
AATCAGTTCATCGAAGCGGAGCTGGCGAGCGTGCATGGGGAGAATGCGAGCACGGGCACG

GAAACAACCATCCCGAAGCTGGTGACTGCCGATGGGACCGAGCTGCAGCTGACTCAATCG
CAGAAGGATGAAATAATTTCTCAGCTACAAAACACTGGCGGT**GG**CAACGACGTGGTGATG
GTGTTGAACGAGGACAACTTTTCGGTCGGATCCACGATACAGCTGGACGAGGGGGTGCAG
ATCGTAGACACGTCCTCGCAAAACATTATGGTAATGTACAGCCAGCCGGAGGGTGGCGAA
TCGGAACATAACGATAGCCAAAAGTCCGACGCGAAGGTGGCCGCCGCCGATGTGACGACG
GCATCGGAAGACGACAATTCGCAAGCATCCGTCTCGCTGTCCGAGGACAAGGCAGCATCG
TCGATCGCGGACAAGCAACCGTCGCTTGAGGGTGGCGATGAGAGCCAGGAATCGCGCCGT
ACGGACGCGTCCGACAAGCCGGAAGCCGAAGACGAAGGCAACACCACCATGGACGATGAC
GTGGACATGTCCGGTGGCGATGTAACGTCGGAAACGAACGAAACGCTGCAGCAAGCCGTC
AAGGACGAGCTGATGGATCTGGAGGACGAGTTGCGGGCCCCCGACACGTCGGCCGACGGA
ACGAAGGAGGACGAGCTGAACGATTCGGAAGATCGGCAGCAATCGGACAAGCTGAAGCTG
ATTTCCGAGCTGGAAGGCGATTGGTCGGAGGATACGGCGGAAGAACAACCCGGCACACCG
GCTGTCGCGGACGTTTCGGAGAAGGCTGCCACGAAAAAGGTGACAGCAACGGCAACGGTT
GATAAGAAAGCAACCGCATCAACTCCTGCCGCCAAAACAGCGTCAAAGCCGACCGGCGGA
GAGAAAAAGCCCGTAGCGGAAGCGGTGGCGACAGGGAAGAAGAGGCCTCTGGGTGAAGCG
CCAGCAGCAGCAGCAGCAGCACCCCCTTCCAAGGAGCGTAAAAAGTCAGGAGGCGAGGAG
AAACTTTCCGAAAAGGAAATCGATAAGCTGCTCGATGATTGGAACAACGAGGCAAAGCTG
GACGAGCTCACAGAGGAAGCGGATCTCGAGAAGGAAATAGCAAAAGTGACCGGTGTGAAG
GCTGCTGAGACCGGCGAGAGTGATAGTAAAAAGCCGAACGAGGTGAAAAAGGACACTGCG
GCTGCGGATGCGACGAAACACGACAAACCGGACACAACGGACGCGGACGCGGACAAGCCA
ACGGTGGAGGAAGAGGAGCAGGAGGAGGATGTGGAGGAAGAGGAAACGGTTACACCTGCG
GAAGAAGGAAAGGAAAGCGTCTCGGTGAAAGGGAAGAAAAATGGTACAGCCTCCGCCACT
GAGACCAGCCGGAGCAAGCTGGAGGACACCAGTGAACTGGACGAAGAGCAGGAAGAGGAC
GACGATGATGGCGACGAAGAGGATAAGGACAAAGCGAACGCTAAGTCGAAGGAGGTTAGC
TCACTGTTGGGCGAGTGGGATGAGGACGATGATCTGTGAAGGGATCAAAGTTGACCACACA
TGTAGTAGCTTACAAAGTTTTACATTTACATTACGTAGTCAGTTTTTACACCTGCCCGGGTCC
GTAGGAAANGGTCGCCAAGCCNGGGGCGTGAGANNTTTTGANAGTTGACCACAGAGGGATT
AACGTTAAGATACGTTGTACGCAAACGATTCCGCTTTTGAGTAAACCTGCGGTAAGGTAGTT
ATAAAGGTAATAAGCTATTATAAAGGTAATAAGCTANNTTAACNGGCTACATCCTATATCGA
AAGGTTNTGA

Figure S2 Amino Acid alignments

SU(HW)

```
DROSO  1    MSASKEGKEKKGKLLGVENISPPKDKRPATRMKLLNDVGAGEDSEASTTTTTSRTPSNKQEKRGSVAGSRIKILNEEILG  80
STEPH  1    MVGS----------------------------------------------------------RRSVVGG-----------  11

DROSO  81   TPKTEKRGATKSTAPAASTVKILNEKKTPSATVTAVETTKIKTSPSKRKKMEHYVLQAVKSENTKADTTVTVVTEEDDTI  160
STEPH  12   ----------------------------------------------------------VRSGGVKKEEILVNVQVNGN--  31

DROSO  161  DFILADDEEVVPGRIENNNGQEIVVTEDDEDLGEDGDEDGEDSSGKGNSS----QTKIKEIV-----EHVCGKCYKTFRR  231
STEPH  32   -----DEIDLVPVNKSADYGRG---SSDDRDGDEEEEEDYEDLSDNPDSDPDLLETEDGRLILDMKRSKYCGRCKKIFKT  103

DROSO  232  VQSLKKHLEFCRYDSGYHLRKADMLKNLEKIEKDAVVMEK---------KDICFCCSESYDTFHLGHINCPDCPKSFKTQ  302
STEPH  104  SAEQKAHKYKC---PGVVDEVQDVILKERKAAKHANPGEEFHKYCNPNPDNPCYCCGEDVSTAHVGHIRCKFCPKSFKAY  180

DROSO  303  TSYERHIFITHSEFSDFPCSICNANLRSEALLALHEEQHKSRGKPYACKICGKDFTRSYHLKRHQKYSSCS--SNETDTM  380
STEPH  181  EYMERHLSSIHSESDAFPCCYCNAKCSTQEILDEHLKTH-DEGKPYACMACGKDFTRRYHLDRHEKHTKCGIVPKEVEVL  259

DROSO  381  SCKVCDRVFYRLDNLRSHLKQHLGTQVVKKPEYMCHTCKNCFYSLSTLNIHIRTHTGEKPFDCDLCDKKFSALVALKKHR  460
STEPH  260  PCEVCGKEFTRIDNLREHLRSHMG-QGARKRDYQCPYCSKSFYGSSLLNIHIRTHTGEKPFPCDLCPMSFPSTGALRKHR  338

DROSO  461  RYHTGEKPYSCTVCNQAFAVKEVLNRHMKRHTGERPHKCDECGKSFIQATQLRTHSKTHI--RPFPCEQCDEKFKTEKQL  538
STEPH  339  RKHTGERPYRCDECSATFAARETLNRHRKTHTGEKPHVCQECGKRFIQATQLRSHMLNHSGDSAFQCSECDATFNRKARL  418

DROSO  539  ERHVKTHSRTKRPVFSCAECKRNF-RTPALLKEHMDEGKHSPKQ-----QRSSMRSAVKIMERT-------DCAICDKNF  605
STEPH  419  TEHIKFTHKGEQP-FECEKCSKTFLRKDDLARHELIHSDVKPYECLVCRKFFTTKNAILLHQRTHLEENPVQCKVCSGTF  497

DROSO  606  DSSDTLRRHIRTVHECDPDDIFG-VEPHPSKR---AKKDIESEEVVPVALNTSAGSLISSQTD---------GNGVVVRE  672
STEPH  498  KRSDCLMRHMRTKHRDFLDKIIDDTEKEKMEKLTTTKRETEKEAVVEI---DGAIYQITTMDDAPAKGTAYTGNNQTVEV  574

DROSO  673  FLVDE-----GDGAAQTITLENETYTILPLDGAIEGEQLTDEAGVKPEAKKEEAQVSPVVKKEQ----------------  731
STEPH  575  FEVVELPASASDAIPNTLPTTSVPYELVPV--VDDTEVILGDSSPVEFVGKANTPIDVVVKKEEDSFKKVKSPPKATVVS  652

DROSO  732  --------------------RKSLAAS------------------------------LAAAIADNLEESCSE--------  753
STEPH  653  KSSEPVAPKVVGKASEGEPSRKSVPSTSKTTATIKPKTEHVVAVRASNEPPQKTPPKLAKEIAREKEEPSTKLAKKPATL  732

DROSO  754  ----------------------------------DDFSGEILTEEDIKLKENVGKLIDMLVDPPILKKYGWPNAPEETVL  799
```

```
STEPH   733   PTARPKEAKSRAVEEAKLKRKREPKPATPKRMRPDEDSIPIFLSDDM-LEQKISELLQMLMGEDMLRSFGWPNAPVEVVL   811

DROSO   800   CKVIENCGHDLTKGGENYAELDYGSRMREYCKLLFTVVIHNDSIKSLLNNFPIDDVIEYVLGDEDQDEGGLDKDNESHSG   879
STEPH   812   GRVIQGCGHQPAKGEEAG---DHTTRMRENTKILFSVTMDDDDIKALLNNHTVDEVIMHVL-------------------   869

DROSO   880   DEEAVSVTGETKTNEIREKPEKKEVSAKSEKKEIVGKAVDKDNSEEVVRENKKKPVGEQEKA   941
STEPH   870   -------------------------------------------------KNK----------   872
```

MOD(mdg4)

```
DROSO   1     MADDEQFSLCWNNFNTNLSAGFHESLCRGDLVDVSLAAEGQIVKAHRLVLSVCSPFFRKMFTQMPSNTHAIVFLNNVSHS   80
STEPH   1     MADDEQFSLCWNNFNSNLSAGFHESLQRGDLVDVTLAAEGQLVQAHRPVLSVCSPYFRKMFNQMPKNQHAFVFLKDVTHS   80

DROSO   81    ALKDLIQFMYCGEVNVKQDALPAFISTAESLQIKGLTDND---PAPQPPQESSPP--PAAPHVQQQQIPA--QRVQRQQP   153
STEPH   81    ALQDLIQFMYCGEVNVKQDALPAFISTAEALQIKGLTETGDSAPAQQSPVKEDPPAVPVATASISTTVPVTPQRTKVQRN   160

DROSO   154   RASARYKIETVDDGLGDEKQSTTQIVIQTTAAPQATIVQQQQPQQAAQQIQSQQLQTGTTTTATLVSTNKRSAQRSSLTP   233
STEPH   161   RIQS-YKLESEESG--DDKI----VHLQASASHHVT----------AQQLQAQ----------TSASSQKRAMPQRGVQS   213

DROSO   234   ASSSAGVKRSKTSTSANVMDPLDSTTETGATTTAQLVPQQITVQTSVVSAAEAKLHQQSPQQVRQEEAEYIDLPME-LPT   312
STEPH   214   HAS----KRTKMSISAS-SDGLDTSDTVTTPTQAQ------TIQT-------VQIVKQLPAQVI--EPEYIELPMESVNP   273

DROSO   313   KSEPDYSED-------------------HGDAAGDAEGTYVEDDTYGDM---RYDDSYFTENEDAGNQTAAN-----TSG   365
STEPH   274   KAEPDYTDEAAEIETVDAETEQEHKLSEHEQGDADDDGNYVEDETYGDMGMGKYEESYLTEGEEGAKPGVSGFVDSYTSD   353

DROSO   366   GGVTATTSKAVVKQQSQNYSESSFVDTSGDQGNTEAQDGPSKDTA---IPKPAEHPRKPVTDSV--QKSPRDADAIPLFD   440
STEPH   354   GG-TGT------EQSTQD---------RGSLRSTGVDDVPFSNSTRMDLLKTIKKEVDPIDKDVLLEQLPFPLRSINPNE   417

DROSO   441   GSRVFVSKVALAKAYIPMPMIYTCRVMDLVIGKDKLVRIAQHEETTDKDLIQDIITHVCKVFALRGNQLTP---SAVQEF   517
STEPH   418   ASTSETNYTNADATHKYLRLVATRKSMRGVSGKKGETIVCQERPK----LAADDRSTMSKFHELKKSNGSPAHRSRCVSC   493


DROSO   518   IDHKLSTLKLMPIKEGK-----------------------------------   534
STEPH   494   FNRTVLESERMPKPKTKWVTTFCGDCPGKPFLCILCFSRIHCNNRIGIGEVL   545
```

CP190

```
DROSO  1      M---GEVKSVKVDNWGVFFLQKLQNFFNKTDYCDLTLQFRDNSQLKVHRLVLSACTDYFNVLEQTCEIVDDALIMPNEFQ   77
STEPH  1      MTRYAEAKSVKVDNWGVFFLQKLQNFFNKTDHCDLTLQFNDNSQLKVHRLVLSACTDYFNHLENVCEMYDDVLIMPMDLQ   80

DROSO  78     ADVVVPIVNFMYTGTLEFELKMYGKLLRTAKEMNMTVLLKLLEAHRRTMENVNRQ-----QRPPSPKGIRRRTVGQPSSG  152
STEPH 81      ADVIVPIVNFMYTGTLEFQYNMYDRLLKTATDMNMTVLLKLLEAHRQTSSRIVKQPVLLNKQAPRPRGAPVYQGAGPRV-  159

DROSO  153    LPQQRVLGPSPQSRNVATPIAQRANTQRGSTG---NTMSRTSGGSNRSPYGDSSN----------------------VKQ  207
STEPH  160    MPAQRGANVGPRPR---LPVARQPMHMYGHKGAGHSTAKDVKPGPSRFDDGEMGEGFEGSFDGITYETKPLLTADQVKKE  236

DROSO  208    EPTSPFEQLRKGYNNN---KRPAQTSLLSPPSKKPSLEEVKEFAEQQRMRKQIAAEYGDNDPEYDGGMLYDDVHAGDDDD  284
STEPH  237    EETSPFEKLRKGYTNANVVKRSSSGSLTSPPAKKPNLDEVKEFTEAARMRSQLATANDDDSPDY----VDDDTHFNDDDD  312

DROSO  285    DDMPPQPSTSKQQSPQGTQTQ--LEHGSTTIILKQDSP------SQTPTIIVKDSSNAKLNHTKIIAEVLRQYPHIVKGH  356
STEPH  313    EDYQPPASVKAALAKSASHHAGTVKHG--TIKQESKSPGSGGNSSAMKQITVKDDSGS-VDHAKIISEVLKKYPHLVKKN  389

DROSO  357    KNIKLKIMPNTPAAPTEKSAPATVKPPANQSSATTSPHKKLHVSFKADKSTPLI---TAQQKAASSQQKSGTSQTTGNQG  433
STEPH  390    KNIKLKIMQK----PSPNNANATNASSAS-SSAVATPKMEVRTATPLKQDLSMVRRATANIRSETAKPSAGSATSTGKPT  464

DROSO  434    TGANPPANTAAAQKRRIDSKTMHALIAQGAENTTGPWLCLRCGVNGRPISIPSYRGFRRHLINTHKETIDPALCEHCGWR  513
STEPH  465    TGTT----------KTIDAKTMHALIAKGAENTTGPWLCLRCGINGRPISIPSYKAFYNHLIHKHKERIDVRICEHCGFK  534

DROSO  514    SVN-NRELHFHMYMEHQTKSLLYTFAECALCNQSYRTKGELEAHINEVHTDDNKQQCIYCNKVFEQELQLYRHMKSYHKE  592
STEPH  535    AQQRNPHLIYHWFTEHKVKPTI-KFPRCDECDHVAMTADDLQKHEEEAHSRGDLQQCIYCNKVFAKEMELYDHMKEQHKA  613

DROSO  593    QALEDGIIDETDEEFLGSQDEEEEA-EGDEE---QEPEQTGKVRILSDISLPATSA------------ITVQQAQ-----  651
STEPH  614    RAIADGVLEFTDDEDYQSQEEEVSSPQGGKTYSAQGASGEGKIKILSNITLPPAKGSSSAAAGGSVRHITLEPSSEAEAL  693

DROSO  652    --------------------------QEQLQEEDVEQVQQE----------VKFVGADGNEVELTDEQRKEILSQLNQQQ  695
STEPH  694    SNVASGIATSLTLVGDNGVVIDDPNYQNQFIEAELASVHGENASTGTETTIPKLVTADGTELQLTQSQKDEIISQLQN--  771

DROSO  696    AGATAGG--VVMVLSEP-----------EAEHVKQETDEKSLA------GTEEEYDDSQIYSELGAADSVESAKKNIADE  756
STEPH  772    ---TGGGNDVVMVLNEDNFSVGSTIQLDEGVQIVDTSSQNIMVMYSQPEGGESEHNDSQKSDAKVAAADVTTASEDDNSQ  848

DROSO  757    SKESIDNLEWAENLIAESEEQSNKEPKSDKPRDDISEKLKELTGDWTEDEN----DDDVDDKPATAELASELANKDPEPT  832
STEPH  849    ASVSLSEDKAASSIADKQPSLEGGDESQESRRTDASDKPE------AEDEGNTTMDDDVD--MSGGDVTSE-TNETLQQA  919
```

```
DROSO  833   VHEEEDDIDLALQSLHKGPEEATEEKASEESVTSADDAVDAVPNINSQ------------PEKMDVDSEAA---------  891
STEPH  920   VKDELMDLEDELRAPDTSADGTKEDELNDSEDRQQSDKLKLISELEGDWSEDTAEEQPGTPAVADVSEKAATKKVTATAT  999

DROSO  892   -DEKASKAEVQIKKEAELENDQEEFIKEDSPIPHSDSVAELREAVTA---------SEGEDDVHLEADNIRKELLDELIA  961
STEPH  1000  VDKKATASTPAAKTASKPTGGEKKPVAEAVATGKKRPLGEAPAAAAAAPPSKERKKSGGEEK--LSEKEIDK-LLDDWNN  1076

DROSO  962   EAEKPDQEKDIVQSEENATTEALDRSVTDEDDLVPPTQVSTEQMEIDEPAAEKAAENNEDTRTADEKEAVEDKPNQTQDV  1041
STEPH  1077  EAKLDELTEEADLEKEIAKVTGVKAAETGESDSKKPNEVKK-----DTAAADATKHDKPDTTDADADKPTVEEEEQEEDV  1151

DROSO  1042  TTAEKPTLESAKAGDEATS------GEAASVDKVKSLISEWGDDDEDEDENG----------VSAAAKE-----------  1094
STEPH  1152  E--EEETVTPAEEGKESVSVKGKKNGTASATETSRSKLEDTSELDEEQEEDDDDGDEEDKDKANAKSKEVSSLLGEWDED  1229

DROSO  1095  -EL  1096
STEPH  1230  DDL  1232
```

Figure S3 Amino acid alignments

SUHW ZnF domains

```
STEPHE          EEFHKYCNPNPDNPCYCCGEDVSTAHVGHIRCKFCPKSFKAYEYMERHL-SSIHSESD-- 195
GAMBIAE         EEFHKYCNPNPDNPCYCCGEDVSTAHVGHIRCKFCPKSFKAYEYMERHL-SSVHSESD-- 192
AEDES           EEFFKYCNPNPENPCYCCGEDVSTAHVGHMKCKLCPKSFKAYEYLERHLQA-VHA--DTT 171
CULEX           EEFFKYCDPNPENPCYCCGEDISTAHTGHIRCKLCPKSFKAYEYMERHL-ASVHAEDDT- 173
                ***.***:***:*********:****.**::**:**********:**** : :*:  *

STEPHE          AFPCCYCNAKC-STQEILDEHLKTHDEGKPYACMACGKDFTRRYHLDRHEKHTKCGIVPK 254
GAMBIAE         AFPCGYCNAKC-STQSVLDEHVKTHDEGKPFACMACGKDFTRRYHLDRHEKHTKCGTVTK 251
AEDES           EFGCTQCNAKCPS-VKVLGQHMMTHSAGKPFSCLKCGKDFTRKYHLERHLNHSSCGEIPK 230
CULEX           -FACPQCNSKCSS-AKVLNEHLLTHSEGKPFSCLRCSKDFTRKYHLERHLNHSNCGDNPK 231
                 * *  **:** *  .:*.:*: **. ***::*: *.*****:***:** :*:.**  .*

STEPHE          EVEVLPCEVCGKEFTRIDNLREHLRSHMGQGA-RKRDYQCPYCSKSFYGSSLLNIHIRTH 313
GAMBIAE         EIAVLPCEVCGKEFTRIDNLREHLRSHMGQGA-RKRDYQCPYCAKSFYGSSLLNIHIRTH 310
AEDES           --YLLPCEVCGKEFTRLDNLREHLRYHMGE-AKRKRDYQCPHCEKAFYGSSLLNIHIRTH 287
CULEX           --HMQPCEVCGKEFARLDNLREHLRFHMGQ-ANRKRDYQCPYCPKSFYGSSLLNIHIRTH 288
                   : *********:*:******** ***: * ********:* *:**************

STEPHE          TGEKPFPCDLCPMSFPSTGALRKHRRKHTGERPYRCDECSATFAARETLNRHRKTHTGEK 373
GAMBIAE         TGEKPFPCDLCPMSFPSTGALRKHRRKHTGERPYRCDECSATFAARETLNRHRKTHTGEK 370
AEDES           TGEKPFPCDLCTKTFPSTGALRKHRRSHTGERPYRCAECSATFAARETLNRHRKTHTGER 347
CULEX           TGEKPFPCDLCTKSFPSTGALRKHRRSHTGERPYRCDQCSATFAARETLNRHRKTHTGEK 348
                ***********. :************.********* :*********************:

STEPHE          PHVCQECGKRFIQATQLRSHMLNHSGDSAFQCSE-CDATFNRKARLTEHIKFTHKG-EQP 431
GAMBIAE         PHECDQCGKRFIQATQLRAHALTHSGENAFQCSE-CDASFNRKARLNDHIKFTHRG-EHP 428
AEDES           PHECTICGKKFIQATQLRAHMFNHTGENGFNC-DQCEAVFNRKARLDEHIRFVH-QKEAP 405
CULEX           PHECTICHKGFIQASQLRSHMFNHTGEGGFDC-EQCGATFSRRTRLTEHVKFVHKQ-EQM 406
                ** *  * * ****:***:* :.*:*:..*:* : * * *.*::** :*::*.*   *

STEPHE          FECEKCSKTFLRKDDLARHELIHS-DVKPYECLVCRKFFTTKNAILL--HQRTHL-EENP 487
GAMBIAE         HECDQCPKTFLRKEDLARHLLIHS-ESKPFECLVCNKFFTTKHAIQL--HQRTHD-EDNP 484
AEDES           LTCEVCSKSFIRKEDLNRHLDTHSDE-KNFSCNHCGKLFASKAA--LKIHERTHIIEE-P 461
CULEX           LPCPTCGKEFVRKEDLQRHQDTH-AEIKNFECSECSKRFVSRTA--LNTHLKVHIVEE-P 462
                  *  * * *:**:** **   *  : * :.*  * * *.:: *  *  * :.*  *: *

STEPHE          -VQCKVC--SGTFKRSDCL-----MRH------MRTKH--RDFLDKIIDDTE---KEKME 528
GAMBIAE         -VQCKVC--SGTFKREDCL-----LRH------MRSKH--RDYLDQIIEDTE---KEKMK 525
AEDES           SV-CSICNHS--FIRRDCLVRHIRTRHAGCPEAM-SKIPPR----KIQSS-EPPR-E--E 509
CULEX           AK-CTICNKE--FLRKDCL-----TRH------MRSKH--R--------D-EYYMKE--E 495
                   *.:*  .  * * ***      **      * :*   *        . *    *  :
```

Mod(mdg4) BTB-POZ domain

```
STEPHE          MADDEQFSLCWNNFNSNLSAGFHESLQRGDLVDVTLAAEGQLVQAHRPVL 50
GAMBIAE         MADDEQFSLCWNNFNSNLSAGFHESLQRGDLVDVTLAAEGHLVKAHRLIL 50
AEDES           MADDEQFSLCWNNFNTNLSAGFHESLVRGDLVDVTLAAEGQLVKAHRLIL 50
CULEX           MADDEQFSLCWNNFNTNLSAGFHESLIRGDLVDVTLAAEGQLVKAHRLIL 50
                ***************:********** *************:**:***:::

STEPHE          SVCSPYFRKMFNQMPKNQHAFVFLKDVTHSALQDLIQFMYCGEVNVKQDA 100
GAMBIAE         SVCSPYFRKMFTQVPVNQHAFIFLKDVSHSALQDLIQFMYCGEVNVKQDA 100
AEDES           SVCSPYFRKMFTQMPANQHAFIFLKDVSHSALKDLIQFMYCGEVNVKQDA 100
CULEX           SVCSPYFRKMFTQMPANQHAFIFLKDVSHSALKDLIQFMYCGEVNVKQDA 100
                ***********.*:* *****:*****:****:*****************

STEPHE          LPAFISTAEALQIKGLTETGDSAPAQQSPVKEDPPA---VPVATASIST- 146
GAMBIAE         LPAFISTAEALQIKGLTETGDSAPTHQSPAKEEPAAAAAVPVTTATISTA 150
AEDES           LPAFISTAEALQIKGLTETGESAPTHPSPAKEAS------HIPSASSTTI 144
CULEX           LPAFISTAEALQIKGLTETGESAPAQQSPAKDSTS-----HHHLPSSGAA 145
                ********************:***:: **.*: .          .:  :
```

CP190 BTB-POZ domain

```
STEPH           MTRYAEAKSVKVDNWGVFFLQKLQNFFNKTDHCDLTLQFNDNSQLKVHRLVLSACTDYFN 60
GAMBIAE         MTRYAEAKSVKVDNWGVFFLQKLQNFFNKTDHCDLTLQFNDNSQLKVHRLVLSACTDYFN 60
AEDES           MTRITETKSVKVDNWGVFFLQKLQNFFNKTDHCDLTLQFNDNSQLKVHRLVLSACTDYFN 60
CULEX           MTRLTETKSVKVDNWGVFFLQKLQNFFNKTDHCDLTLQFNDNSQLKVHRLVLSACTDYFN 60
                *** :*:*****************************************************

STEPH           HLENVCEMYDDVLIMPMDLQADVIVPIVNFMYTGTLEFQYNMYDRLLKTATDMNMTVLLK 120
GAMBIAE         HLENVCEMYDDVLIMPMDLQADVIVPIVNFMYTGTLEFQYNMYDRLLKTATDMNMTVLLK 120
AEDES           MLENTCEMYDDVLIMPLDLQADVIVPIVNFMYTGTLEFQYNMYEKLLKTARDMNMTVLLK 120
CULEX           MLESTCEMYDDVLIMPLDLQADVIVPIVNFMYTGTLEFQYGIYEKLLKTAKDMSMTVLLK 120
                 **..***********:***********************.:*::***** **.******

STEPH           LLEAHRQTSSRIV---KQPVLLNKQAPRP-RGAPVYQG-------AGPRVMPAQRG---- 165
GAMBIAE         LLEAHRQTSSRIV---KQPVLLNKQAPRPQRGGPVYQG----NHGQQRGGLPVQRGPPSG 173
AEDES           LLEAHKQTTRMVQ-GGRQPIVLNKQAPNARPRSTAQSGVTYQMPASVARSYGSVKQTTQT 179
CULEX           LLEAHKQTTRINQNVGRQPVLLNKQAPSARGRNPQSS-----FHSAPSRSYGAIRQPAPV 175
                *****:**:       :**::****** .    .  .                 :
```

CP190 ZnF domain

```
STEPH           TAK---------PSAGSATSTGKPTTGT-TKTIDAKTMHALIAKGAENTTGPWLCLRCGI 498
GAMBIAE         MAKGGPSTTSQSSSSSTAATTGKPSGGGTTKTIDAKTMHALIAKGAENTTGPWLCLRCGI 516
AEDES           ------------RAPAPVAT-STP--G--PKKIDAKTMHALIAKGAENMTGPWLCLRCGI 506
CULEX           ------------QPAAPAPTPSTPSSG--PKKIDSKTMHALIAKGAENMTGPWLCLRCGI 499
                             ......: ..*  *  .*.**:************* ***********

STEPH           NGRPISIPSYKAFYNHLIHKHKERIDVRICEHCGFKAQQRNPHLIYHWFTEHKVKPTIKF 558
GAMBIAE         NGRPISIPSYKAFYNHLIHKHKERIDVRICEHCGFKAQQRNPHLLYHWHTVHNIKPALRF 576
AEDES           DGRPISIPSYKAFRNHLIVKHKEKIDARICEHCGWKAEQRNPNLFYHWLTEHNIKSTMKF 566
CULEX           DGRPISIPSYKAFRNHLIVKHKEKIDARICEHCGWRAEQRNPNLFYHWLTEHNIKSTIKF 559
                :************ **** ****:**.*******::*:****:*:*** * *::*.:::*

STEPH           PRCDECDHVAMTADDLQKHEEEAHSR---GDLQQCIYCNKVFAKEMELYDHMKEQHKARA 615
GAMBIAE         PRCGECDHVAMTADDLAKHEEEVHASSGGGDLQQCIYCNKVFAKEMELYDHMKEQHKQRA 636
AEDES           PTCAECGHVAMNTAALQKHEEENHQR---GDLQQCIYCNKVFAKEMDLYEHMKSLHKQRA 623
CULEX           PTCAECGHVAMNATLLAKHEEENHQR---GDLQQCIYCN--------------------- 595
                * * **.****.:  * ***** *     **********
```