Science Advances
AAAS

advances.sciencemag.org/cgi/content/full/6/10/eaaz2309/DC1

Supplementary Materials for

Computer simulations explain mutation-induced effects on the DNA editing by adenine base editors

Kartik L. Rallapalli*, Alexis C. Komor*, Francesco Paesani*

*Corresponding author. Email: krallapa@ucsd.edu (K.L.R.); akomor@ucsd.edu (A.C.K.); fpaesani@ucsd.edu (F.P.)

Published 4 March 2020, *Sci. Adv.* **6**, eaaz2309 (2020)
DOI: 10.1126/sciadv.aaz2309

This PDF file includes:

Supplementary Materials and Methods

System setup

The TadA0.1 was constructed using the previously published crystal structure of *E.coli* TadA (PDB ID: 1z3a) (*10*). All crystallographic water molecules within 3 Å distance of the protein surface were preserved during the modelling process. The wild-type TadA structure was transformed into the ABE monomers using the mutagenesis plug-in available in PyMOL (*31*). *S. aureus* TadA enzyme (PDB :2b3j) was identified as a structural homologue of *E. coli* TadA using Blastp algorithm (fig. S9). The *ec*TadA and *sa*TadA were combined to yield the *ec*TadA-RNA complex. Since the natural substrate of TadA is tRNA anti-codon loop (stem looped secondary structure) and the system of interest contains single stranded DNA (ssDNA), the *ec*TadA-RNA complex was remodeled using steered molecular dynamics (SMD) simulations, to better represent the ABE-ssDNA complex (based on the cryo-EM structure of CRISPR-Cas9 (PDB ID: 5y36) (*14*). SMD simulation with an external force of 100 kcal/mol $Å^2$ along the collective variable defined by the distance between the 5′ and 3′-terminal bases of RNA hairpin loop. Over a span of 100 ns SMD simulation, the terminal bases were moved 30 Å apart in order to capture the disordered ssDNA using the R-loop as the reference model (fig. S9). The TadA*0.1-ssDNA complex was mutated to higher TadA*-ssDNA complexes using the PyMOL mutagenesis tool (*31*). A similar approach was adopted to generate the modified TadA*1.2(N108D) and TadA*2.1(N108D) complex, where Asn108 was reversed back to aspartate. All titratable residues were assigned protonation states predicted using the H++ server at pH 7 (*34*, *35*).

Using the LEap tool from the AmberTools suite (*38*), the resulting structures of the enzyme and the enzyme-substrate complex were embedded into a pre-equilibrated truncated octahedron box of explicit water molecules, which were described by the TIP3P model (*33*), with a 10 Å buffer distance. Varying number of Na^+ ions were added to each system to maintain charge neutrality and the simulation cell was then replicated infinitely in three dimensions to impose periodic boundary conditions.

Molecular dynamics simulations

Molecular dynamics (MD) simulations were performed using the CUDA version of PMEMD of Amber18 in the canonical ensemble (constant number of atoms, temperature, and volume) (*39-41*). The protein and the DNA atoms were represented using the AMBER ff14SB force field and the *bsc1* parameters, respectively (*36-38*). The structures were subjected to energy minimization using a combination of steepest descent and conjugate gradient algorithms. A total of 50000 steps were used in the minimization, with the first 2000 steps carried out with the steepest descent algorithm and the remaining 48000 steps carried out with conjugate gradient algorithm. Initial velocities were assigned to all atoms in the system through gradual heating to 298.15 K using Langevin dynamics with a collision frequency of 1 ps^{-1}. The equilibration was performed in an isobaric-isothermal (NPT: constant number of atoms, pressure, and temperature) ensemble at 298.15 K and 1 bar, using Langevin dynamics for temperature regulation and a Monte Carlo barostat for pressure control. Both heating and equilibration were carried out with additional harmonic restraints (with a force constant, k, of 2 kcal/mol $Å^2$) on all protein and DNA residues. To capture the effects of the neighboring nuclear environment (i.e. Cas9 and the genome), harmonic restraints were imposed on the bases on 5′- and 3′-termini of the substrate DNA sequence during all subsequent simulations. The non-bonded interactions were cut off at 8 Å and

the long-range electrostatics were evaluated using the particle mesh Ewald method. The hydrogen atom bond length was constrained by implementing the SHAKE algorithm. The final snapshots of the 10 ns equilibrating were used to initiate the various biased and unbiased MD production trajectories. All MD simulations were propagated in time using the velocity Verlet with a time step of 2 fs.

Steered molecular dynamics and umbrella sampling simulations

The distance between the centers of mass of the protein and the ssDNA was chosen as the collective variable (ξ) to monitor the binding and unbinding of the protein-ssDNA complexes. The equilibrated structures were subjected to constant force pulling and pushing along the _ value using a 100 kcal/mol Å^2 moving harmonic restraint. These steered MD trajectories were used to uniformly change the ξ value at the rate of 0.1 Å/ns from 19 Å to the target values of 19±10 Å over 100 ns (fig. S7).

The initial configurations for seeding the 41 windows ($\xi \in [10, 30]$; step 0.5 Å) of the umbrella sampling (US) simulations were extracted from the SMD trajectories. Each US window was subjected to short equilibration, which was followed by production for 5 ns under umbrella restraints with a force constant, k, of 20 kcal/mol Å^2. Four independent US simulations were carried out for each of the 41 windows for all the protein-ssDNA complexes, which were then used to determine the potentials of mean force (PMFs) representing the free energy profiles associated with the protein-ssDNA binding process along ξ. Hence, a total of 820 ns (20 ns ×41 windows) was sampled for each complex, which was found enough to converge the PMFs. Table S2 lists all the systems that were modelled along with the corresponding type of simulation that was performed.

Calculations of binding free energy and error bars

The US simulations led to the accumulation of 100000 values of the instantaneous reaction coordinate in each of the 41 windows along the collective variable ξ describing the protein-DNA binding process. This raw data (fig. S10) represents the biased probability distribution along ξ. The WHAM algorithm was then used with a convergence threshold of 10^{-8} to obtain the unbiased probability distribution and subsequently derive the actual PMF along ξ (*42*). The standard deviation associated with 4 independent PMF profiles was calculated to estimate the error in convergence due to sampling. Additional error analysis was performed using the block averaging method to evaluate the uncertainty associated with the normalization procedure implemented within WHAM algorithm (*43*).

Analysis protocol

The structural parameters of the protein and protein-DNA complexes during the unbiased MD simulations were analyzed using the cpptraj tool (*17, 18*). To measure the spatial extent of motion of each residue of the apo-proteins and to determine the effects of individual mutations on the overall structure of the protein, an RMS mass-weighted calculation was carried out on the entire 500 ns trajectory, using the initial frame as a reference. We performed hierarchical agglomerative clustering of the 250000 trajectory snapshots of the unbiased MD simulations of each of the ABE models using the position of Cα atoms. Clusters representing highly similar poses were merged using the group average method for clustering implemented in cpptraj. This yielded a total of 10 representative conformations for each TadA* models (Fig. 2, table S1). We defined the primary interaction shell as a shell of 4 Å radius around the target A and its 5′- and

3′-flanking bases. Protein residues in this primary interaction shell were identified by analyzing the unbiased MD trajectories using cpptraj with distance-based mask of 4 Å around the three bases in the active site. The atom-list per frame data obtained using the cpptraj module was re-normalized to give the percentage residue contacts, using the following formula (Fig. 3, fig. S3)

$$Percentage\ contact = \frac{Total\ atomic\ contact\ during\ all\ frames\ \times 100}{Number\ of\ atoms\ in\ the\ amino\ acid\ \times Total\ number\ of\ frames}$$

The average number of hydrogen-bonding interactions between the residues in the primary interaction shell and the substrate DNA was computed using the `hbond` feature in cpptraj with the default hydrogen-bond definition (3 Å distance between donor and acceptor atoms, and 135º angle between the donor, hydrogen, and acceptor atoms). The second shell residues were identified within a 4 Å distance of all the residues of the first shell, and the associated percentage contacts were calculated in a manner similar to that described for residues in the primary interaction shell (Fig. 4, fig. S1A and S4). The PDB2PQR webserver in conjunction with the APBS server was used to calculate the electrostatic maps for the TadA*0.1 and TadA*2.1 models (*44*) (fig. S4 and S4C). The visualization of the biomolecular trajectories was rendered using Chimera (*32*) and the data was plotted using Matplotlib (*45*).

Supplementary Figures

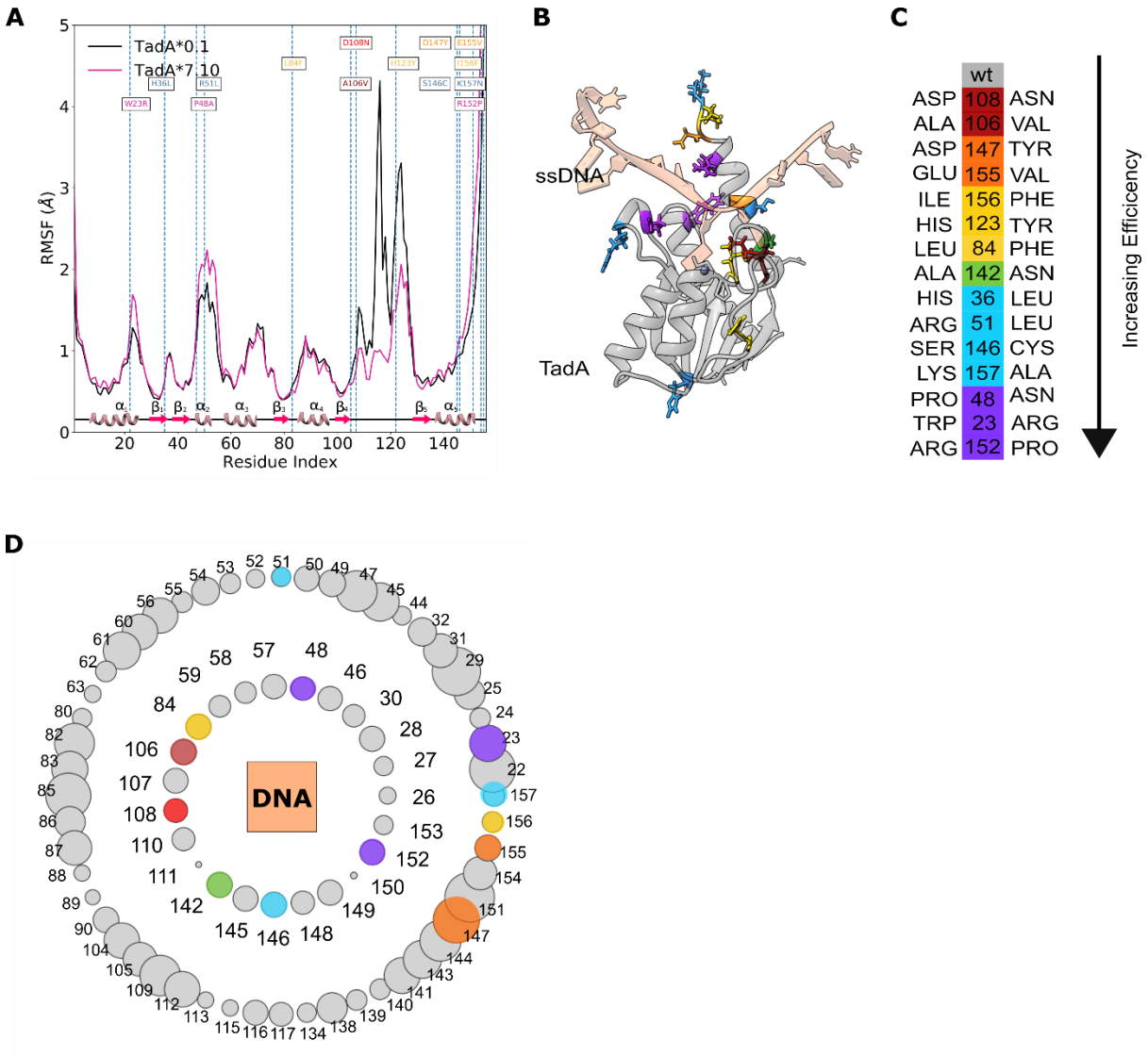


Fig. S1. Asteroid plot for TadA*7.10-ssDNA. (**A**) The root mean squared fluctuations (RMSFs) as a function of the residues of each of the TadA*7.10 compared to the TadA*0.1. (**B**) The residues that were mutated during the 7 rounds of directed evolution highlighted as sticks on the structure of TadA* protein and color-coded as per (**C**). (**D**) The first and second interaction shell of around the 3 nucleotides in the active site of the TadA*7.10-ssDNA complex. The size of the nodes corresponds to the time for which the amino acid resides in the first/second shell for the TadA*0.1-ssDNA simulations.

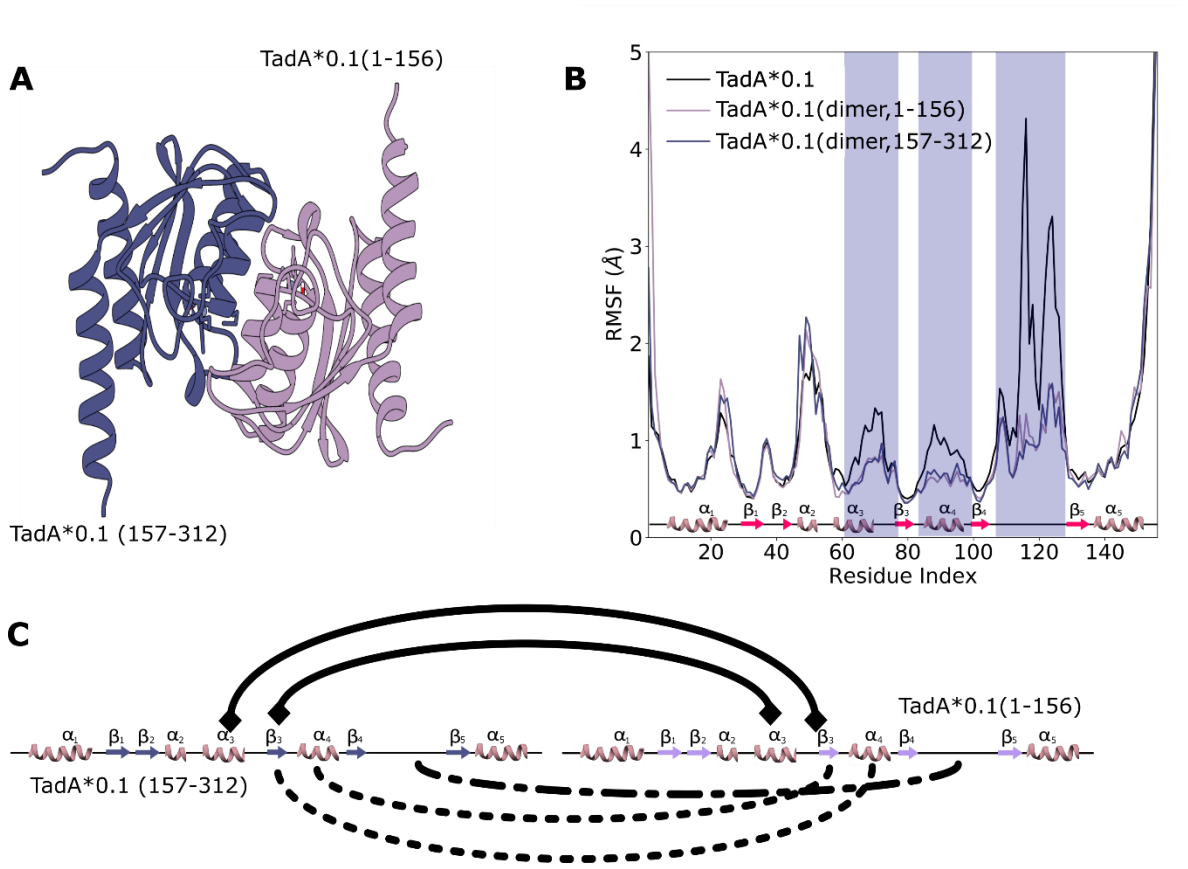


Fig. S2. Comparison of the structural flexibility of the TadA monomer with TadA dimer. (**A**) Structure of TadA dimer highlighting the monomeric subunits. (**B**) The root mean squared fluctuations (RMSFs) as a function of the residues of each of the TadA monomer compared to the monomeric subunits of TadA dimer. (**C**) The schematic diagram of the TadA dimer highlighting the regions that serve as the dimerization interface and undergo the most change in structural dynamics upon dimerization.

A

	Percentage Contact	H-bond acceptor	H-bond donor
26	45	0.2	0.1
27	59	0.0	0.0
28	100	0.0	0.0
30	79	0.0	0.0
46	94	0.3	0.3
48	87	0.0	0.0
57	95	0.1	0.1
58	73	0.0	0.0
59	74	0.2	0.0
84	99	0.0	0.0
106	97	0.0	0.0
107	97	0.8	0.1
108	90	0.0	0.0
110	83	0.0	0.7
111	6	0.0	0.0
142	100	0.0	0.0
145	100	0.0	0.0
146	99	0.0	0.8
148	86	0.0	0.0
149	100	0.0	0.0
150	7	0.0	0.0
152	97	0.0	0.7
153	58	0.0	0.6

B

	Percentage Contact	H-bond acceptor	H-bond donor
26	66	0.2	0.0
27	72	0.1	0.0
28	100	0.0	0.0
30	46	0.0	0.0
46	48	0.1	0.2
48	21	0.0	0.0
57	31	0.0	0.0
58	31	0.0	0.0
59	38	0.3	0.0
84	100	0.0	0.0
106	98	0.0	0.0
107	99	0.7	0.0
108	100	0.1	1.0
110	96	0.0	0.7
111	56	0.0	0.0
142	99	0.0	0.0
145	99	0.0	0.0
146	94	0.0	0.0
148	86	0.0	0.0
149	94	0.0	0.0
150	3	0.0	0.0
152	94	0.0	0.7
153	57	0.0	0.6

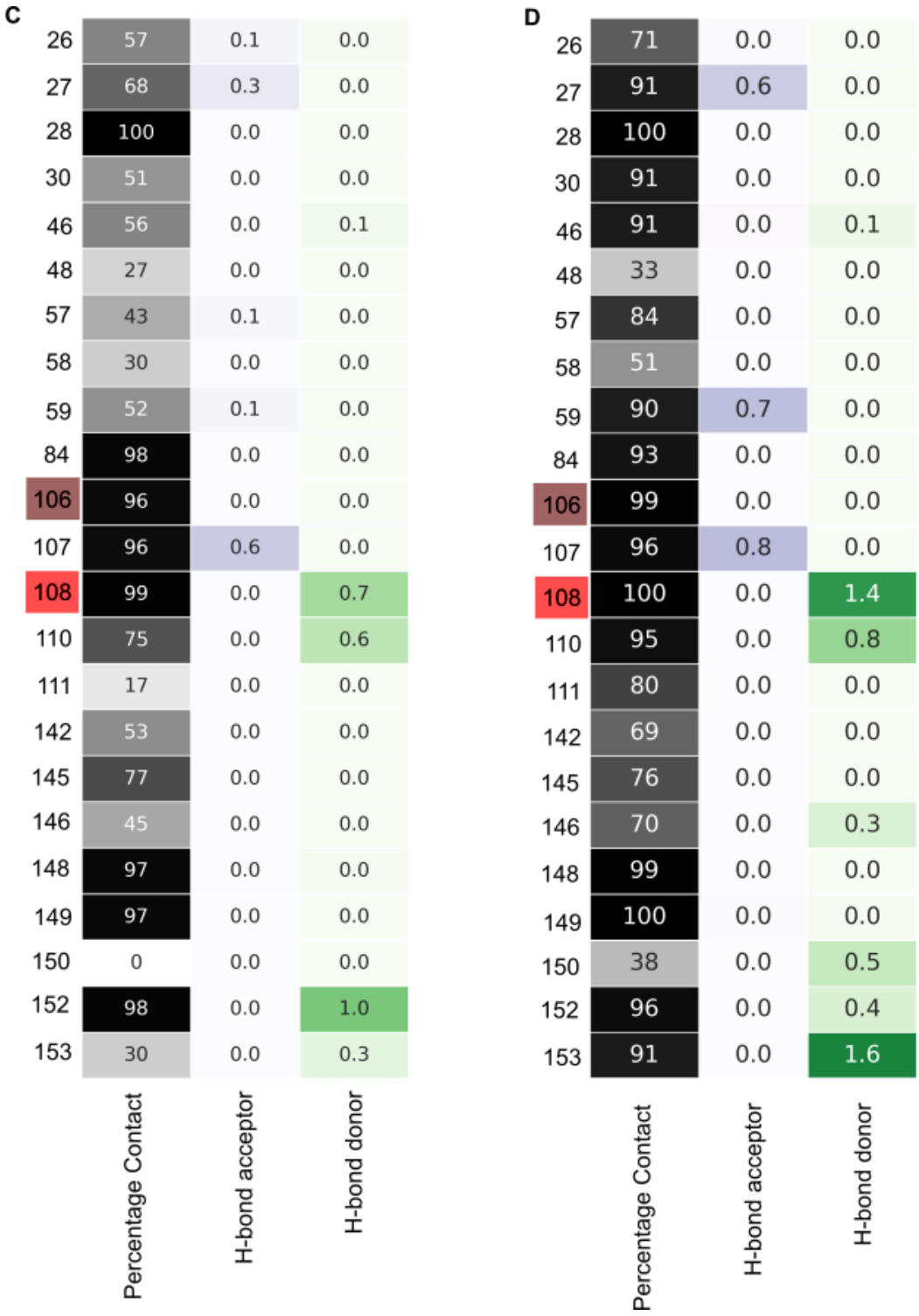


Fig. S3. Percentage contact and the fractional H-bonding between the three nucleotides and the first interaction shell amino acids. For (**A**) TadA*0.1, (**B**) TadA*1.1, (**C**) TadA*1.2 and (**D**) TadA*2.1. The cells are colour coded as per the generation in which the mutations were identified and the intensity of the colours in the columns signifies the magnitude of the percentage contact and the H-bonding strength.

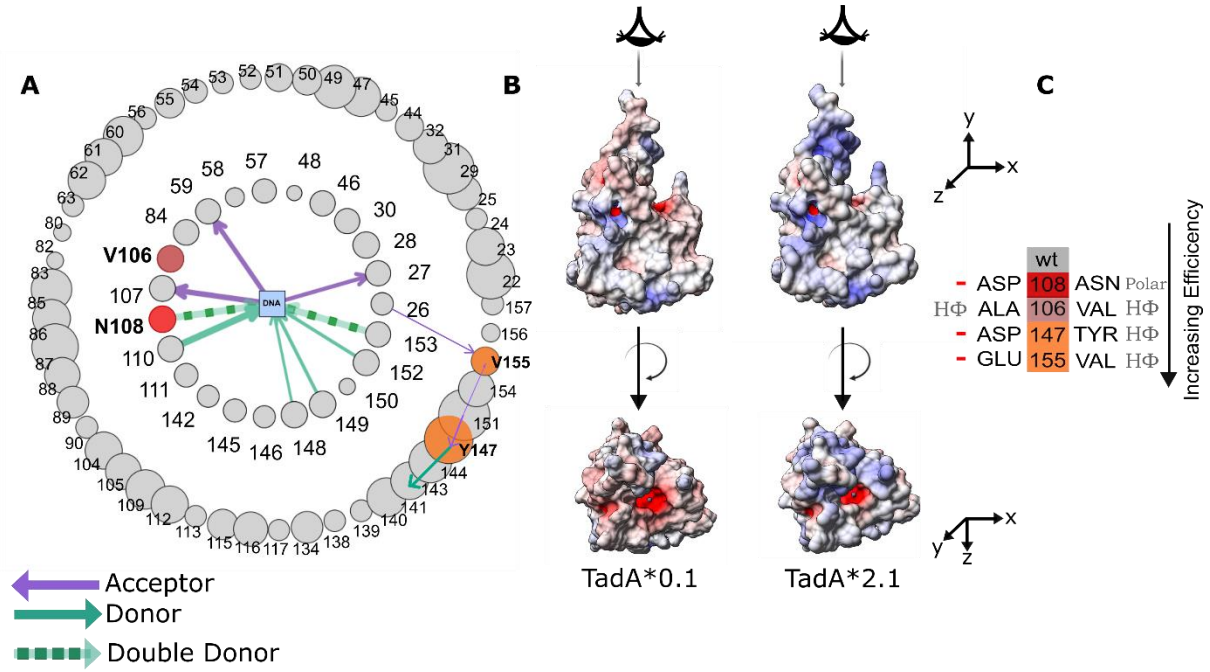


Fig. S4. Asteroid plot for TadA*0.1-ssDNA complex. (**A**) The first and second interaction shell of around the 3 nucleotides in the active site of the TadA*0.1-ssDNA complex. The size of the nodes corresponds to the time for which the amino acid resides in the first/second shell. (**B**) Electrostatic map of TadA*0.1 and TadA*2.1. (**C**) The chemical nature of the mutations that were introduced in the TadA* during the first two rounds of directed evolution.

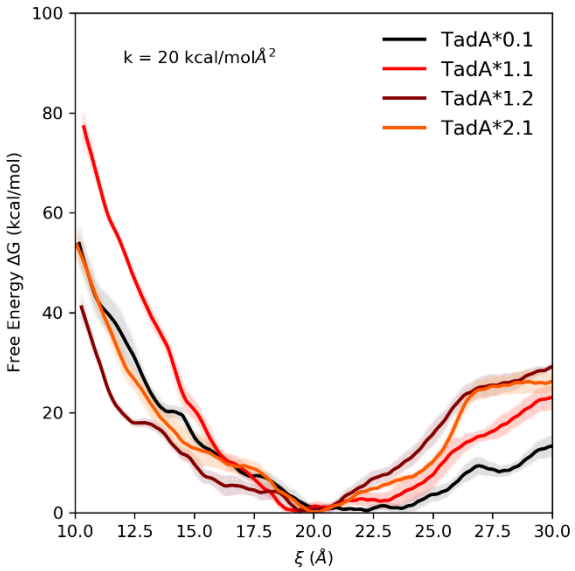


Fig. S5. Mutations lead to an increase in TadA* binding to the ssDNA (AAG). The free energy profile of binding of the ssDNA (5′-GTCA**A**GAAAC-3′) to various TadA*s. For each TadA*-ssDNA complex, the average PMF is shown as a function of the continuously changing ξ values. The shaded regions around the individual curves depict the standard deviation for 4 independent replicates of the umbrella sampling simulations.

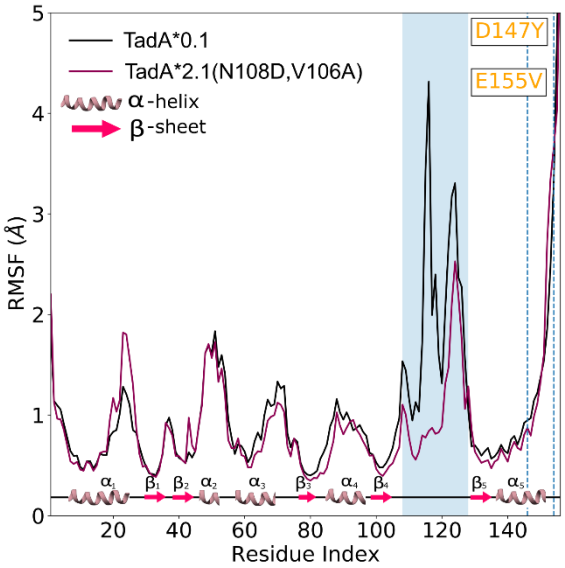


Fig. S6. Structural importance of Ala^{106}Val mutation. Residue level flexibility of TadA* shown in terms of the root-mean-squared-fluctuations (RMSFs) of the C_ atoms of the peptide backbone of TadA**2.1(N108D, V106A). The _4-beta5 loop region is highlighted. This plot, along with the RMSF analysis of TadA*1.2(N108D) and TadA*2.1(N108D) variants (Fig. 6B in the main text)) shows that Ala106Val mutation imparts structural stability to the β4-β5 loop of the TadA*, in a manner analogous to the Asp108Asn mutation.

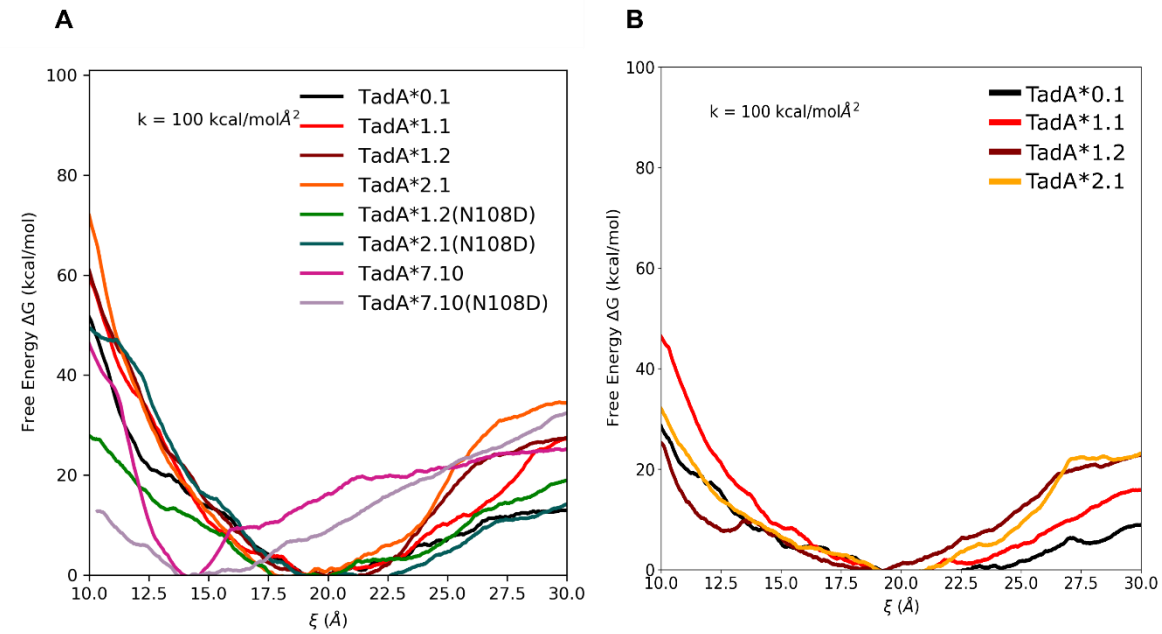


Fig. S7. PMF of the TadA*-ssDNA complexes calculated using steered MD simulations. The configurations extracted from the SMD trajectory were used to initiate umbrella simulations for (**A**) TadA*-ssDNA (T**A**C) and (**B**) TadA*-ssDNA (A**A**G). The incorporation of 14 mutations in the TadA* structure leads to a shift in the equilibrium position of the ssDNA towards the active site of the TadA*. The reversal of Asn108 to Asp for the initial mutants is accompanied by decrease in the binding affinity for the initial mutants, TadA*1.2 and TadA*2.1. For the highly mutated TadA*7.10, the Asn108Asp reversal mutation leads to a flattening of the free energy profile, which is indicates that the ssDNA can easily traverse out of the active site and hence the TadA*7.10(N108D) has lacks ssDNA editing efficiency.

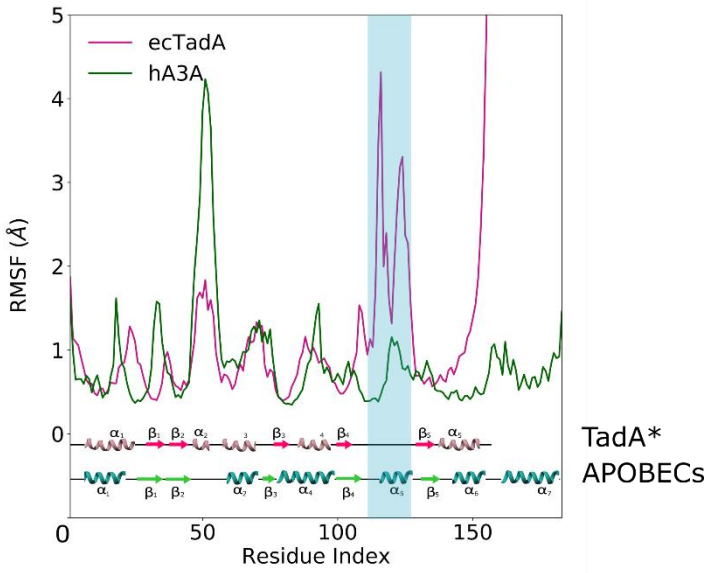


Fig. S8. Comparison of the structural flexibility of the ecTadA with hAPOBEC3A. Shown are the root mean squared fluctuations (RMSFs) as a function of the residues of each of the proteins.

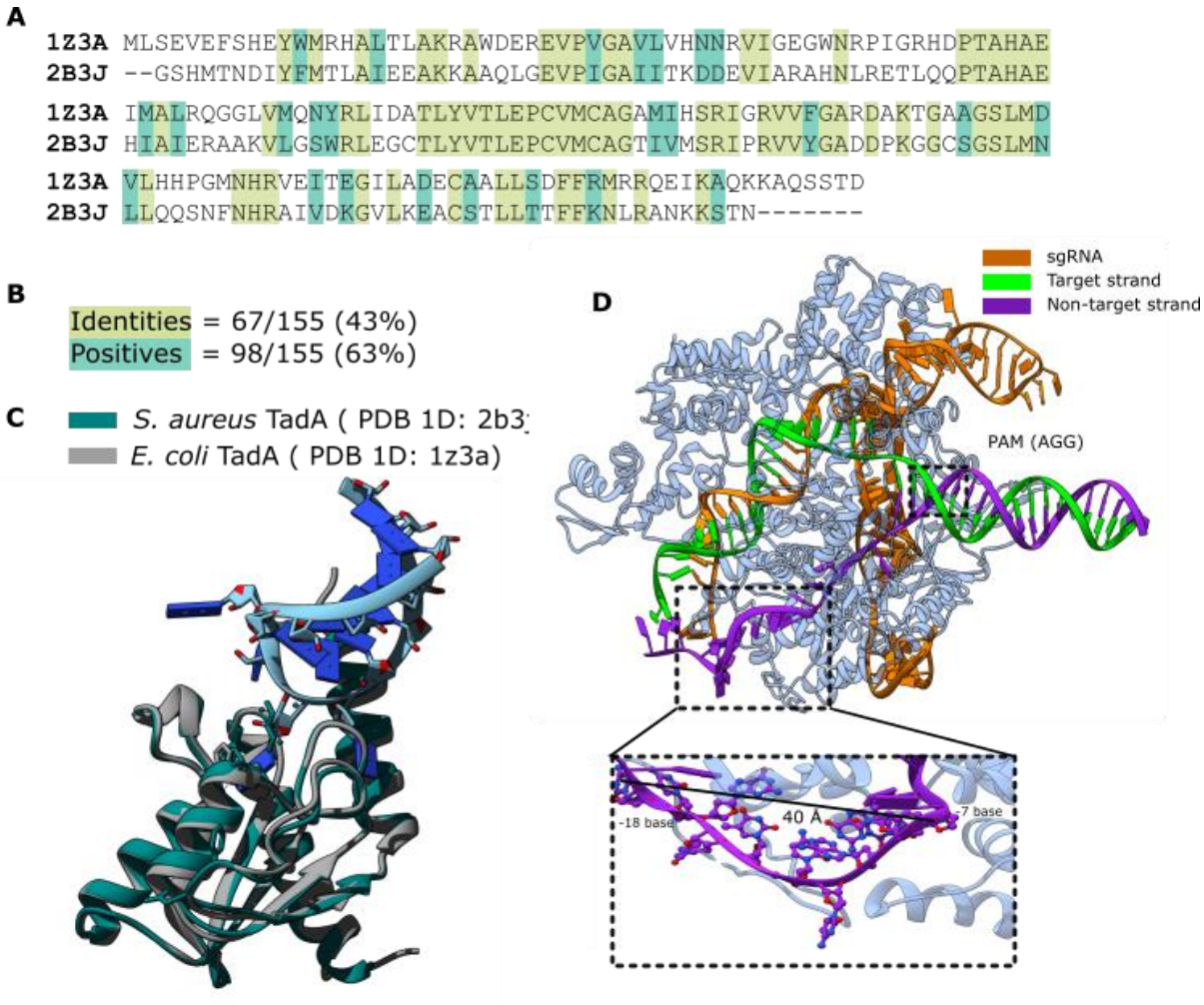
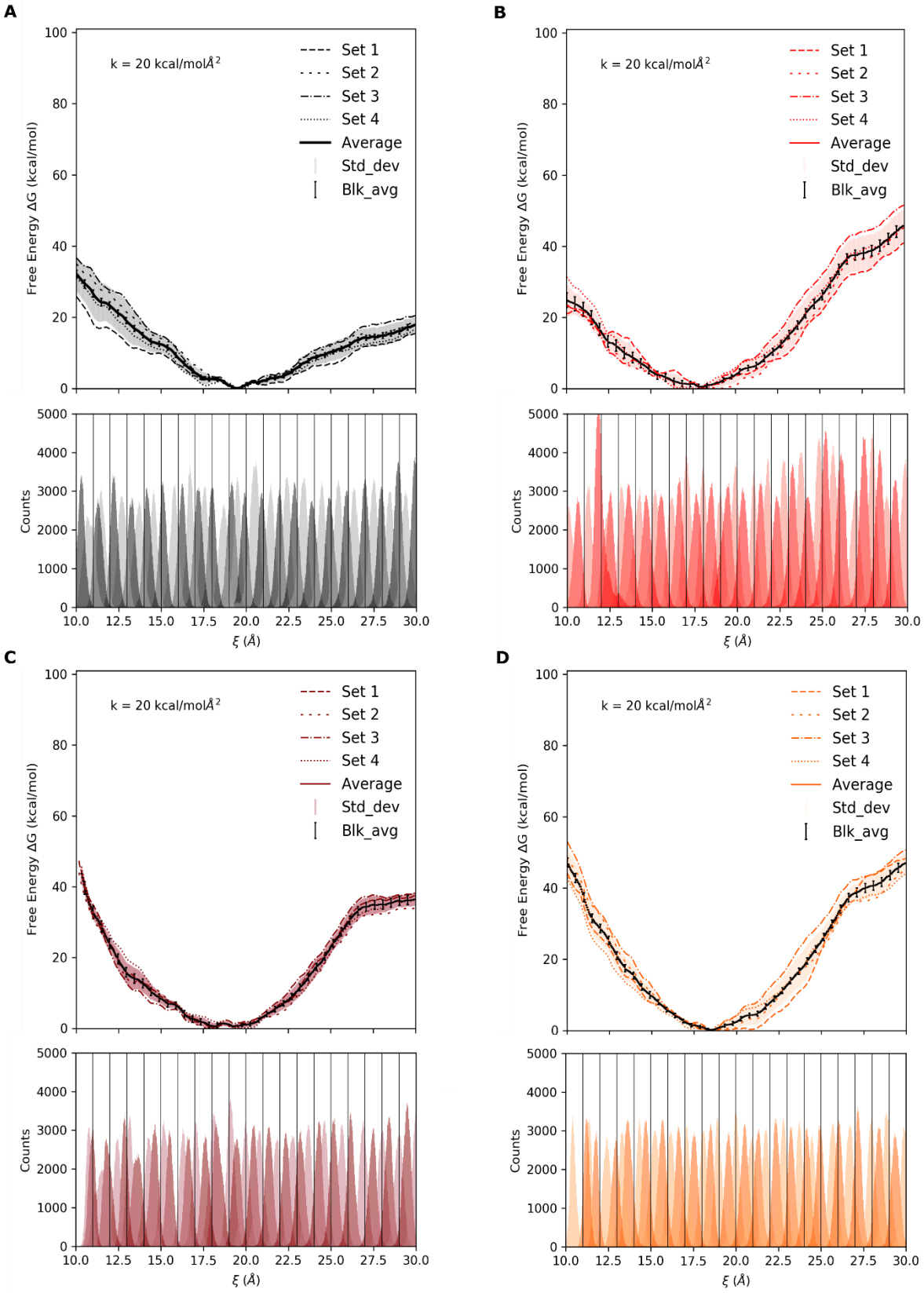


Fig. S9. Modeling of TadA*-ssDNA. (**A**) Sequence alignment of *E. coli* TadA with *S. aureus* TadA. The aligned sequences are highlighted to depict identical and similar residues between the two sequences. (b) Percentage identity and similarity scores of the TadA sequences, calculated using Blastp. (**C**) Structural alignment of *S. aureus* TadA-RNA and *E. coli* TadA. (**D**) Structure of *S. pyogenes* Cas9 in complex with the sgRNA and the target DNA. The 11 nucleotides of the distal end of the non-target strand are highlighted as the purple sticks in the figure inset, bases -7 to -18 from the PAM site were used are reference to model the RNA from fig. S7 (C) into the exposed ssDNA.


A
B
C
D
k = 20 kcal/molÅ²
Set 1
Set 2
Set 3
Set 4
Average
Std_dev
Blk_avg
Free Energy ΔG (kcal/mol)
Counts
ξ (Å)

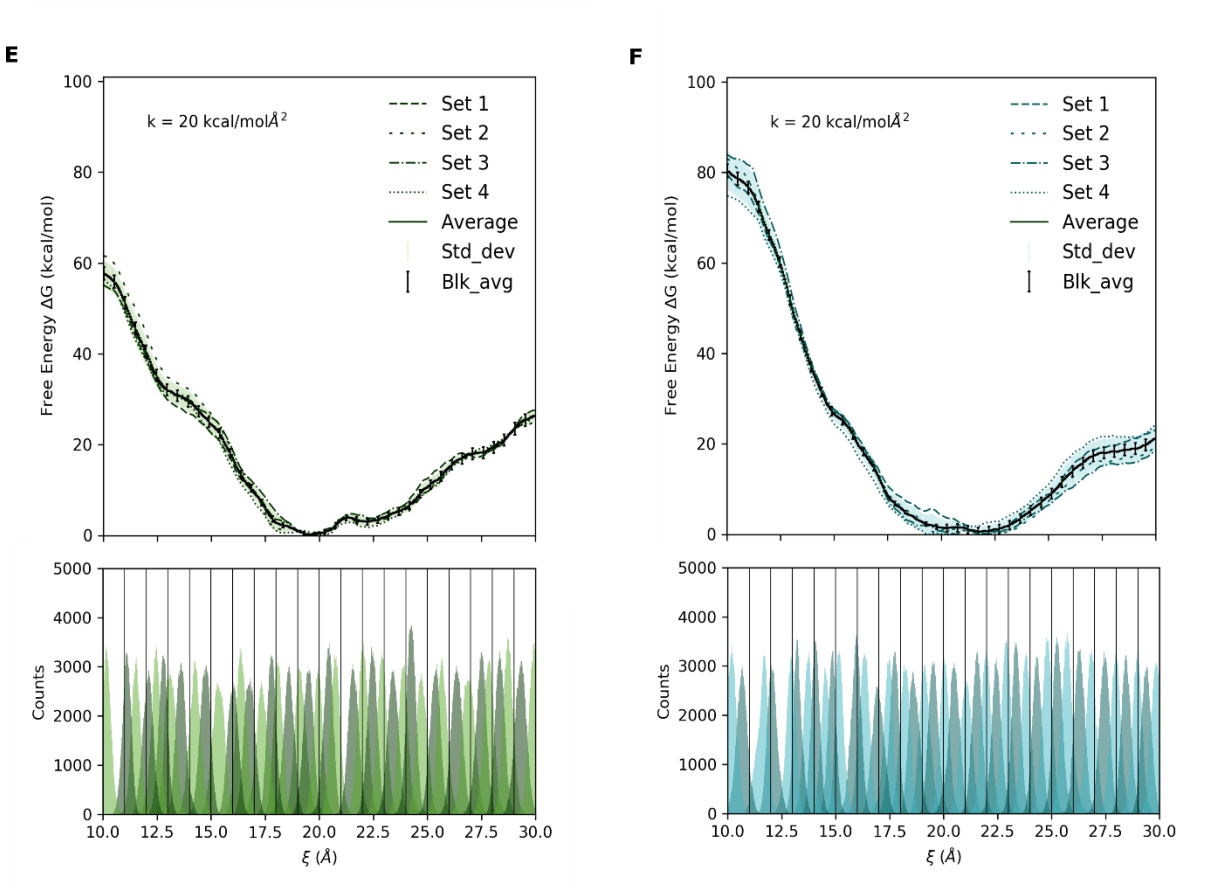


Fig. S10. Umbrella sampling data and biased statistics. Individual PMFs associated to 4 independently conducted umbrella sampling simulations and the biased probability distributions (histograms) obtained from individual windows that stratify the ξ space, for TadA*0.1-ssDNA (**A**), TadA*1.1-ssDNA (**B**), TadA*1.2-ssDNA (**C**), TadA*2.1-ssDNA (**D**), TadA*1.2(N108D)-ssDNA (**E**), and TadA*2.1(N108D)-ssDNA (**F**) complex. In each PMF profile the shaded region indicated the standard deviation of the individual PMFs and the error bars are the error calculated using block averaging method for individual windows.

Supplementary Tables

Table S1. Comparison of RMSD of the representative clusters. The alignment of all the clusters is done with respect to the Cluster 1, using the Cα distance of the residues composing the β4-β5 loop.

Cluster Number	RMSD TadA*0.1	RMSD TadA*1.1	RMSD TadA*1.2	RMSD TadA*2.1
1	0.000	0.000	0.000	0.000
2	3.017	0.725	0.437	0.559
3	0.348	0.796	0.445	0.693
4	1.333	1.484	0.824	0.861
5	0.435	0.777	1.293	0.670
6	2.626	0.571	1.141	0.660
7	3.216	0.798	1.101	0.585
8	0.864	0.781	1.130	0.891
9	2.377	0.578	1.434	0.916
10	3.331	0.930	1.083	0.405
Average	1.754	0.744	0.888	0.624

Table S2. Summary of the systems modeled and the types of simulations conducted in this study.

System	Number of Atoms	Simulation Type	Simulation Length(ns)
wtTadA-RNA	33906	Steered MD	100
TadA*0.1	35727	Unbiased MD	500
TadA*1.1	35729		
TadA*1.2	35741		
TadA*2.1	35751		
TadA*7.10	35648		
TadA*1.2(N108D)	35537		
TadA*2.1(N108D)	35752		
TadA**2.1(N108D,V106A)	35728		
APOBEC	79144		
TadA*0.1-ssDNA(TAC)	36125	Unbiased MD	500
TadA*1.1-ssDNA(TAC)	36504		
TadA*1.2-ssDNA(TAC)	36507		
TadA*2.1-ssDNA(TAC)	36557		
TadA*0.1-ssDNA(TAC)	36125	Steered MD	200
TadA*1.1-ssDNA(TAC)	36504		
TadA*1.2-ssDNA(TAC)	36507		
TadA*2.1-ssDNA(TAC)	36557		
TadA*7.10-ssDNA(TAC)	35257		
TadA*1.2(N108D)-ssDNA(TAC)	36510		
TadA*2.1(N108D)-ssDNA(TAC)	36554		
TadA*0.1-ssDNA(AAG)	30287		
TadA*1.1-ssDNA(AAG)	28764		
TadA*1.2-ssDNA(AAG)	28764		
TadA*2.1-ssDNA(AAG)	28733		
TadA*0.1-ssDNA(TAC)	36125	Umbrella Sampling	41 windows × 4 sets × 20ns
TadA*1.1-ssDNA(TAC)	36504		
TadA*1.2-ssDNA(TAC)	36507		
TadA*2.1-ssDNA(TAC)	36557		
TadA*1.2(N108D)-ssDNA(TAC)	36510		
TadA*2.1(N108D)-ssDNA(TAC)	36554		
TadA*0.1-ssDNA(AAG)	30287		
TadA*1.1-ssDNA(AAG)	28764		
TadA*1.2-ssDNA(AAG)	28764		
TadA*2.1-ssDNA(AAG)	28733		

Table S3. DNA sequences used for simulations and in mammalian tissue culture experiments. The sequence context (target motif) of the edited A is shown for each site. PAM sequences are shown in blue

Target Motif	Protospacer and PAM Sequence
TAC	GACTACAGACT (Sequence used in simulations)
AAG	GTCAAGAAAC (Sequence used in simulations)
CAC	GAACA$_5$CAAAGCATAGACTGCGGG
TAT	GAGTA$_5$TGAGGCATAGACTGCAGG
CAA	GAGCA$_5$AAGAGAATAGACTGTAGG
AAA	GAGCAA$_6$AGAGAATAGACTGTAGG
GAC	GGATTGA$_7$CCCAGGCCAGGGCTGG
CAC	CAGAGA$_7$CTGGAATTCGTCAGGG

Table S4. First round genomic DNA PCR sequences.

Primer Name	Primer Sequence
CAC (pos 5)-Fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATTGTCCAGCCCCATCTGTCAA
CAC (pos 5)-Rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTTTCAAGTTACTGCAGCCCAAGC
TAT-Fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNAGAGACTGATTGCGTGGAGT
TAT-Rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCACTCCAGCCTAGGCAACAA
CAA/AAA-Fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNCTGCACCTAGCCTCCATGTC
CAA/AAA-Rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCTTGCACTGAGACCGTGAA
GAC-Fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNATGTGGGCTGCCTAGAAAGG
GAC-Rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTCCCAGCCAAACTTGTCAACC
CAC (pos 7)-Fwd	ACACTCTTTCCCTACACGACGCTCTTCCGATCTNNNNGCCCTGCTTCTTTTTCTCTGGT
CAC (pos 7)-Rev	TGGAGTTCAGACGTGTGCTCTTCCGATCTACCATTAACGCAGCCAACTTCA

Supplementary sequences

red: TadA; purple: mutations of interest; blue: Cas9; green: EGFP

ABE0.1

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE1.1

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE1.2

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGV
RNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE1.2(N108D)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGV
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE2.1

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGV
RNAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE2.1(N108D)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGV
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSYFFRMRRQVIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE7.10(monomer)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEG
WNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGV
RNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE7.10(monomer, N108D)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRARDEREVPVGAVLVLNNRVIGEG
WNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAMIHSRIGRVVFGV
RDAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVFNAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAVITDEYKVPSKKF
KVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYLQEIFSNEMAKVD
DSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVDSTDKADLRLIYL
ALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGVDAKAILSARLSK
SRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSKDTYDDDLDNLLA
QIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQDLTLLKALVRQQ
LPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVKLNREDLLRKQRT
FDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGPLARGNSRFAWMT
RKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYEYFTVYNELTKVK
YVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSVEISGVEDRFNAS
LGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYAHLFDDKVMKQLK
RRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSLTFKEDIQKAQVS
GQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMARENQTTQKGQKN
SRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQELDINRLSDYDV
DHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLNAKLITQRKFDNL
TKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLIREVKVITLKSKL
VSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGDYKVYDVRKMIAK
SEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVWDKGRDFATVRKV
LSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDSPTVAYSVLVVAK
VEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKLPKYSLFELENGR
KRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVEQHKHYLDEIIEQ
ISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPAAFKYFDTTIDRK
RYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPKKKRKVGSGATNF
SLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGEGDATYGKLTLKF
ICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQERTIFFKDDGNYKT
RAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQKNGIKVNFKIRHN
IEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVLLEFVTAAGITLG
MDELYKSGGSPKKKRKV

ABE7.10(dimer)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAM
IHSRIGRVVFGVRNAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAV
ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV
DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK
DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ
DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE
YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV
EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL
TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGD
YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW
DKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS
PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPK
KKRKVGSGATNFSLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGE
GDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQER
TIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQK
NGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVL
LEFVTAAGITLGMDELYKSGGSPKKKRKV

ABE7.10(dimer, N108D)

MKRTADGSEFESPKKKRKVSEVEFSHEYWMRHALTLAKRAWDEREVPVGAVLVHNNRVIGEG
WNRPIGRHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTLEPCVMCAGAMIHSRIGRVVFGA
RDAKTGAAGSLMDVLHHPGMNHRVEITEGILADECAALLSDFFRMRRQEIKAQKKAQSSTDS
GGSSGGSSGSETPGTSESATPESSGGSSGGSSEVEFSHEYWMRHALTLAKRARDEREVPVGA
VLVLNNRVIGEGWNRAIGLHDPTAHAEIMALRQGGLVMQNYRLIDATLYVTFEPCVMCAGAM
IHSRIGRVVFGVRDAKTGAAGSLMDVLHYPGMNHRVEITEGILADECAALLCYFFRMPRQVF
NAQKKAQSSTDSGGSSGGSSGSETPGTSESATPESSGGSSGGSDKKYSIGLAIGTNSVGWAV
ITDEYKVPSKKFKVLGNTDRHSIKKNLIGALLFDSGETAEATRLKRTARRRYTRRKNRICYL
QEIFSNEMAKVDDSFFHRLEESFLVEEDKKHERHPIFGNIVDEVAYHEKYPTIYHLRKKLVD
STDKADLRLIYLALAHMIKFRGHFLIEGDLNPDNSDVDKLFIQLVQTYNQLFEENPINASGV
DAKAILSARLSKSRRLENLIAQLPGEKKNGLFGNLIALSLGLTPNFKSNFDLAEDAKLQLSK
DTYDDDLDNLLAQIGDQYADLFLAAKNLSDAILLSDILRVNTEITKAPLSASMIKRYDEHHQ
DLTLLKALVRQQLPEKYKEIFFDQSKNGYAGYIDGGASQEEFYKFIKPILEKMDGTEELLVK
LNREDLLRKQRTFDNGSIPHQIHLGELHAILRRQEDFYPFLKDNREKIEKILTFRIPYYVGP
LARGNSRFAWMTRKSEETITPWNFEEVVDKGASAQSFIERMTNFDKNLPNEKVLPKHSLLYE
YFTVYNELTKVKYVTEGMRKPAFLSGEQKKAIVDLLFKTNRKVTVKQLKEDYFKKIECFDSV
EISGVEDRFNASLGTYHDLLKIIKDKDFLDNEENEDILEDIVLTLTLFEDREMIEERLKTYA
HLFDDKVMKQLKRRRYTGWGRLSRKLINGIRDKQSGKTILDFLKSDGFANRNFMQLIHDDSL
TFKEDIQKAQVSGQGDSLHEHIANLAGSPAIKKGILQTVKVVDELVKVMGRHKPENIVIEMA
RENQTTQKGQKNSRERMKRIEEGIKELGSQILKEHPVENTQLQNEKLYLYYLQNGRDMYVDQ
ELDINRLSDYDVDHIVPQSFLKDDSIDNKVLTRSDKNRGKSDNVPSEEVVKKMKNYWRQLLN
AKLITQRKFDNLTKAERGGLSELDKAGFIKRQLVETRQITKHVAQILDSRMNTKYDENDKLI
REVKVITLKSKLVSDFRKDFQFYKVREINNYHHAHDAYLNAVVGTALIKKYPKLESEFVYGD
YKVYDVRKMIAKSEQEIGKATAKYFFYSNIMNFFKTEITLANGEIRKRPLIETNGETGEIVW
DKGRDFATVRKVLSMPQVNIVKKTEVQTGGFSKESILPKRNSDKLIARKKDWDPKKYGGFDS
PTVAYSVLVVAKVEKGKSKKLKSVKELLGITIMERSSFEKNPIDFLEAKGYKEVKKDLIIKL
PKYSLFELENGRKRMLASAGELQKGNELALPSKYVNFLYLASHYEKLKGSPEDNEQKQLFVE
QHKHYLDEIIEQISEFSKRVILADANLDKVLSAYNKHRDKPIREQAENIIHLFTLTNLGAPA
AFKYFDTTIDRKRYTSTKEVLDATLIHQSITGLYETRIDLSQLGGDSGGSKRTADGSEFEPK
KKRKVGSGATNFSLLKQAGDVEENPGPMVSKGEELFTGVVPILVELDGDVNGHKFSVSGEGE
GDATYGKLTLKFICTTGKLPVPWPTLVTTLTYGVQCFSRYPDHMKQHDFFKSAMPEGYVQER
TIFFKDDGNYKTRAEVKFEGDTLVNRIELKGIDFKEDGNILGHKLEYNYNSHNVYIMADKQK
NGIKVNFKIRHNIEDGSVQLADHYQQNTPIGDGPVLLPDNHYLSTQSALSKDPNEKRDHMVL
LEFVTAAGITLGMDELYKSGGSPKKKRKV