Systematic Evaluation of the Dependence of Deoxyribozyme Catalysis on Random Region Length

Tania E. Velez, Jaydeep Singh, Ying Xiao, Emily C. Allen, On Yi Wong, Madhavaiah Chandra, Sarah C. Kwon, and Scott K. Silverman*

Department of Chemistry, University of Illinois at Urbana-Champaign, 600 South Mathews Avenue, Urbana, IL 61801, USA

Deoxyribozyme sequences

Α		1 10	20	0
8\//	12	GAACGTGGTG	ССТССТААСА	20
8\//	14	AGATGTGGTG	CGTGCCAAAA	20
8\//	\ 4 \5		GCGATGGATC	20
8\//	10	TAGACGTAAA	CTGGAGTTG.	. 10
8\//	10	CGGCAATATG	CACAGTTGGG	2 20
۵۷/ ۹۱/۸	10	CGAATGTGGT	GCGTGCTAAA	20
81/4	11	GAACTGTGGT	GCGTGCCAAA	20
8\/Δ'	23	AGGGGGCGTGA	GGGGTTCTTC	20
8\/Δ'	25	TTAGGGAGGG	CCACCAGCTT	
770	22	CAGGATAATG	GCGGTTTGGT	
770	25	GGATCATGGC	GGTCTGGTTT	
071	45	GCCCGGGAAT	GAGTCTTAGC	20
521	10	1 10	20	0 30
		i i	Ĩ	
8VE	31	ATGGGGCACA	GTTCTCTCAT	ACCCCTGGAA 30
8VE	32	TGGTTCGCAC	TTTCCAGGAC	CAGGTAACCAC 30
8VE	34	CGGACCCGGG	CTCGACCTCG	G TGCTGAGCAT 30
8VE	35	TACAGCACAG	GAGTTACGTC	CGGGTAAGTG 30
8VE	37	CCTTGGTGAG	AACGCACCTC	CGGACGTGG 30
8VE	39	CGGGGGAATGG	AGGCGTCCCA	ATGCAAATCG 30
8VB	12	ACCGCGCGGA	AGGCCTTTCT	CGAAGGGCGA 30
8VB	16	CCACTCCGTG	CTCCTCTTGA	TGAGTAGGGC 30
8VB	18	TACACTCATG	GCGGTGTGAT	TCGATGCCGA 30
8VB2	20	TCGCGGCACA	TTAGTGTGAG	G TGGATCACGT 30
8VB2	21	ATCGGGTATT	ACGCGGACGG	G TTGCCCACCA 30
8VB2	22	CCGGCAGTGT	GCTTGGGACA	GCTTTGCTGG 30
8VB2	25	CAGGGGCGGT	AGGCGTTACA	CTCAAATTGA 30
8VB2	26	CCAGGATCAA	TGATAAGCCG	G AGTCAAAGGG 30
7Z.	J7	GGATCATGGC	GGTTTGGTTA	CGCTTTGCGT 30
7ZJ	12	GCCCACCAGG	CTAGGGAGTG	GTTACGAGGA 30
		1 10	20	0 30 40
971	1	GTGTGCTGGT		
521	- '	0101001001	0000110000	
B		1 10	20	0 30
D				
810	13	GCGCIGGGAG	GCACAIGCIG	GGTTGCACCG 30
810	18	• • • • • • • • • • •	IGAAAG.	. 16
81M	12		IAGG.	
		1 10 I I	20 I	
7TQ2	20	CAAGGAGAGC	TGTACAAGCT	CGGGTCGTGT TCAAAGG G ATCATAGTGA GTAC-AAAAC GG 59
7TC	22	. T		
710	23	. T		
7TQ	11	. T		
7TQ	12	. T		
7TQ	16	. T		G
7TQ	53	. T		
7104	16	ТΔ	CGTAG T A	

Figure S1. Sequences of the initially random $(N_{20}-N_{60})$ catalytic regions of the various new deoxyribozymes. The systematic alphabetic nomenclature for each selection experiment (and therefore the resulting deoxyribozymes) is described in the figure captions. (A) Deoxyribozymes for DNA cleavage. (B) Deoxyribozymes for tyrosine-RNA nucleopeptide linkage formation. In all cases, the binding arm sequence to the 5'-side of the catalytic region was 5'-CCGTCGCCATCTCTTC-3', and the binding arm sequence to the 3'-side of the catalytic region was 5'-ATAGTGAGTCGTATTA-3'. For 8VA23, a G \rightarrow A binding arm mutation (5'-ATAA...) was identified by sequencing of the clone and found to be required for activity; only a trace amount of DNA cleavage was observed when the original G was included. For 9ZH5, a T \rightarrow C binding arm mutation (5'-ATAGTGAGC...) was identified by sequencing of the clone and found to be required for activity; no DNA cleavage was observed when the original T was included.

deoxyribozyme ^a	mass	mass	L error, %	mass	mass	R error, %	hydrolysis
	L calcd.	L found	(found - calcd.)	R calcd.	R found	(found - calcd.)	site ^b
8VA5 (N ₂₀)	4823.2	4823.9	+0.01	5701.7	5702.3	+0.01	TAT^ <u>C</u> GAA, 5′-p
8VA6 (N ₂₀)	5192.4	5195.6	+0.06	5332.5	5335.5	+0.06	TAT <u>C</u> ^GAA, 3′-p
8VB4 (N ₃₀)	4285.8	4288.2	+0.06	6239.1	6242.3	+0.05	T^AT <u>C</u> GAA, 3′-p
8VB20 (N ₃₀)	4519.0	4516.4	-0.06	6005.9	6002.0	-0.06	TA^T <u>C</u> GAA, 5′-p
8VB21 (N ₃₀)	4903.2	4910.3	+0.14	5621.7	5629.8	+0.14	TAT^ <u>C</u> GAA, 3′-p
7ZG2 (N ₂₀)	6084.0	6083.3	-0.01	5996.9	5995.9	-0.02	TAT <u>C^G</u> GAA, 5′-p
7ZG5 (N ₂₀)	6084.0	6082.1	-0.03	5996.9	5995.0	-0.03	TAT <u>C^G</u> GAA, 5′-p
9ZH5 (N ₂₀)	6084.0	6087.4	+0.06	5996.9	6000.1	+0.05	TAT <u>C^G</u> GAA, 5′-p
7ZJ7 (N ₃₀)	6084.0	6086.1	+0.03	5996.9	5999.1	+0.04	TAT <u>C^G</u> GAA, 5′-p
7ZJ12 (N ₃₀)	6084.0	6085.5	+0.02	5996.9	5998.0	+0.02	TAT <u>C^G</u> GAA, 5′-p
9ZL1 (N ₄₀)	6084.0	6085.3	+0.02	5996.9	5997.6	+0.01	TAT <u>C</u> ^ <u>G</u> GAA, 5′-p

MALDI mass spectrometry data

Table S1. MALDI mass spectrometry data for products of the new deoxyribozymes that catalyze DNA hydrolysis. L = left-hand cleavage product; R = right-hand cleavage product.

^{*a*} The number of originally random nucleotides is shown in parentheses by the deoxyribozyme name.

^b The DNA hydrolysis site is marked with \uparrow within the illustrated portion of the DNA substrate sequence, where the underlined <u>C</u> or <u>CG</u> is unpaired (all other substrate nucleotides are base-paired with the deoxyribozyme binding arms; see full substrate sequences in Experimental Procedures). The location of the phosphate that remains after hydrolysis, either 5' or 3', is indicated.

deoxyribozyme ^a	mass	mass	L error, %	mass	mass	R error, %	deglycosylation
	L calcd.	L found	(found - calcd.)	R calcd.	R found	(found - calcd.)	site ^b
8VA2 (N ₂₀)	4599.0	4603.8	+0.10	5701.7	5707.7	+0.11	TA T<u>C</u>GAA
8VA4 (N ₂₀)	4599.0	4602.0	+0.07	5701.7	not obs.	_	TA T<u>C</u>GAA
	4903.2	4903.7	+0.01	5412.5	5412.5	0	TAT <u>C</u> GAA
8VA8 (N ₂₀)	5192.4	5192.8	+0.01	5083.3	5083.9	+0.01	TAT <u>C</u> AA
	5521.6	5521.8	+0.004	4770.1	4770.5	+0.01	TAT <u>C</u> G A A
8VA10 (N ₂₀)	4599.0	4601.6	+0.06	5701.7	5704.6	+0.05	TA T<u>C</u>GAA
8VA11 (N ₂₀)	4599.0	4599.0	0	5701.7	5701.5	-0.004	TA T<u>C</u>GAA
8VA23 (N ₂₀)	4903.2	4907.3	+0.08	5412.5	5416.5	+0.07	TAT <u>C</u> GAA
8VA25 (N ₂₀)	5521.6	5522.0	+0.01	4770.1	4770.7	+0.01	TAT <u>C</u> G A A
8VB1 (N ₃₀)	4285.8	4288.2	+0.06	6005.9	6009.0	+0.05	T A T <u>C</u> GAA
8VB2 (N ₃₀)	5192.4	5190.6	-0.03	5083.3	5081.4	-0.04	TAT <u>C</u> AA
	5521.6	5519.7	-0.03	4770.1	4768.2	-0.04	TAT <u>C</u> G A A
8VB5 (N ₃₀)	5192.4	5193.6	+0.02	5083.3	5084.2	+0.02	TAT <u>C</u> GAA
8VB7 (N ₃₀)	4903.2	4906.0	+0.06	5412.5	not obs.	-	TAT <u>C</u> GAA
	5192.4	5196.0	+0.07	5083.3	5084.8	+0.03	TAT <u>C</u> AA
8VB9 (N ₃₀)	5521.6	5528.8	+0.13	4770.1	4776.5	+0.13	TAT <u>C</u> G A A
8VB12 (N ₃₀)	5192.4	5191.8	-0.01	5083.3	5082.5	-0.02	TAT <u>C</u> AA
8VB16 (N ₃₀)	5192.4	5192.5	+0.002	5083.3	5083.7	+0.01	TAT <u>C</u> AA
	5521.6	5521.5	-0.002	4770.1	4770.7	+0.01	TAT <u>C</u> G A A
8VB18 (N ₃₀)	5192.4	5196.0	+0.07	5083.3	5086.7	+0.07	TAT <u>C</u> AA
8VB22 (N ₃₀)	5521.6	5532.7	+0.20	4770.1	4780.3	+0.21	TAT <u>C</u> G A A
8VB25 (N ₃₀)	5521.6	5528.6	+0.13	4770.1	4776.0	+0.12	TAT <u>C</u> G A A
8VB26 (N ₃₀)	5192.4	5197.3	+0.09	5083.3	5087.7	+0.09	TAT <u>C</u> AA

Table S2. MALDI mass spectrometry data for products of the new deoxyribozymes that catalyze DNA deglycosylation and strand scission by two β -elimination reactions. L = left-hand cleavage product; R = right-hand cleavage product.

^{*a*} The number of originally random nucleotides is shown in parentheses by the deoxyribozyme name.

^b The deglycosylation site is marked in bold italics within the illustrated portion of the DNA substrate sequence, where the underlined C is unpaired (all other substrate nucleotides are base-paired with the deoxyribozyme binding arms; see full substrate sequence in the Experimental Procedures). For deoxyribozymes that deglycosylate at either of two adjacent nucleotides, data for each of the two sites is given on consecutive lines.

deoxyribozyme ^a	mass	mass	error, %
	calcd.	found	(found - calcd.)
8TM3 (N ₃₀)	12357.2	12366.5	+0.08
7TQ20 (N ₆₀)	12357.2	12364.1	+0.06
7TQ46 (N ₆₀)	12357.2	12349.8	-0.06

 Table S3. MALDI mass spectrometry data for products of the new deoxyribozymes that catalyze tyrosine-RNA nucleopeptide linkage formation.

^a The number of originally random nucleotides is shown in parentheses by the deoxyribozyme name.

Kinetic data for deoxyribozymes from the initial selections for DNA cleavage



Figure S2. Kinetic plots for additional individual deoxyribozymes from the initial selections for DNA cleavage (see Figure 3 for other deoxyribozymes).

<u>*k*_{obs} values for kinetic plots</u>

deoxyribozyme ^a	$k_{\rm obs}, {\rm h}^{-1}$
8VA2 (N ₂₀)	0.12
8VA4 (N ₂₀)	0.16
8VA5 (N ₂₀)	0.011*
8VA6 (N ₂₀)	0.19
8VA8 (N ₂₀)	0.009*
8VA10 (N ₂₀)	0.22
8VA11 (N ₂₀)	0.19
8VA23 (N ₂₀)	0.051
8VA25 (N ₂₀)	0.15
8VB1 (N ₃₀)	0.052
8VB2 (N ₃₀)	0.41
8VB4 (N ₃₀)	0.18
8VB5 (N ₃₀)	0.52
8VB7 (N ₃₀)	0.14
8VB9 (N ₃₀)	0.17
8VB12 (N ₃₀)	0.17
8VB16 (N ₃₀)	0.22
8VB18 (N ₃₀)	0.23
8VB20 (N ₃₀)	0.006*
8VB21 (N ₃₀)	0.094
8VB22 (N ₃₀)	0.080
8VB25 (N ₃₀)	0.086
8VB26 (N ₃₀)	0.041
7ZG2 (N ₂₀)	0.053
7ZG5 (N ₂₀)	0.30
9ZH5 (N ₂₀)	0.063
7ZJ7 (N ₃₀)	0.19
7ZJ12 (N ₃₀)	0.29
9ZL1 (N ₄₀)	0.52
8TM3 (N ₃₀)	0.15
11MN5 (N ₄₀) ³³	0.95
7TQ20 (N ₆₀)	0.073

Table S4. k_{obs} values for the kinetic plots shown in Figures 3, 5 and 6 and Figure S2. Unmarked values were obtained from direct fits to first-order kinetics, $Y = Y_{max} \cdot (1 - e^{-kt})$. Values marked with an asterisk were obtained from linear fits due to the low rate constants. ^{*a*} The number of originally random nucleotides is shown in parentheses by the deoxyribozyme name.