Directed Evolution of Recombinase Specificity by Split Gene Reassembly

Charles A. Gersbach, Thomas Gaj, Russell M. Gordley, and Carlos F. Barbas, III* The Skaggs Institute for Chemical Biology and the Departments of Chemistry and Molecular Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California, 92037

Supplementary Material

Table S1: Primer Sequences

5' XbaI pBla:	5 ' - CACCACTCTAGAACCCCTATTTGTTTATTTTTC-3 '
3' KpnI AmpR:	5 ' - CACCACGGTACCTTACCAATGCTTAATCAGTG-3 '
5' AmpR mid-Spe-Hind:	5 ' - ACTAGTCACCACAAGCTTACTTACTCTAGCTTCCCGGCAAC - 3 '
3' AmpR mid-Spe-Hind:	5 ' - TAAGCTTGTGGTGGTGACTAGTTAGTTCGCCAGTTAATAGTTTG - 3 '
5'-XbaI C4-20G-GFP:	5 ' - TTAATTAAGAGTCTAGAGCGGGAGGCGTGTCCAAAACCATGGTTTACAGCACGCCTCCCGCAGATCTAGGAGGAATTTAAAATGAG- 3 '
3'-HindIII C4-20G-GFP:	5 ' - ACTGACCTAGAGAAGCTTGCGGGAGGCGTGCTGTAAACCATGGTTTTGGACACGCCTCCCGCCTGCAGTTATTTGTACAGTTCATC-3 '
3'-HindIII C4-20T-GFP:	5 ' - ACTGACCTAGAGAAGCTTGCGGGAGGCGTGCGAAATATTATAAATTATCACACGCCTCCCGCCTGCAGTTATTTGTACAGTTCATC- 3 '
5'-XbaI C4-20E-GFP:	5 ' - TTAATTAAGAGTCTAGAGCGGGAGGCGTGGTGAGCACCATGGAGCTGGCCACGCCTCCCGCAGATCTAGGAGGAATTTAAAATGAG- 3 '
3'-HindIII C4-20E-GFP:	5 ' - ACTGACCTAGAGAAGCTTGCGGGGAGGCGTGGCCAGCTCCATGGTGCTCACCACGCCTCCCGCCTGCAGTTATTTGTACAGTTCATC- 3 '
ResGin-cat fo1 prim1:	5 ' - ACCACGATGACTGACCTAGAGCTCAGGAGGAATTTAAAATGCTGATTGGCTATGTAAGGG - 3 '
3'-ZF SS-AXEX prim2:	5 ' - CAGTATCACCTCGAGGAATTCTCTAGAGGCGCGCCTTATTGGCCGGCC

Supporting Information Figure S1: Sequences of Evolved Gin Variants on GE Substrate

1	,10	20	30	40		50	60	70
Gin <mark>M</mark>	LIGYVRVSTN	ID <mark>Q</mark> NTDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	VVWKLDRLGF	RS <mark>MK</mark> HLISI
GinL7C7 <mark>M</mark>	LIGYVRVST <mark>N</mark>	IG <mark>Q</mark> S <mark>TDLQR</mark> DA	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	E <mark>RLQKGDTI</mark>	. <mark>VVWKLDRLGF</mark>	RS <mark>V</mark> KHLISI
Gin-GE1 (9) <mark>M</mark>	LIGYVRVST <mark>S</mark>	D <mark>Q</mark> STDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	. <mark>VVWKLDRLGF</mark>	RS <mark>LK</mark> HLISI
Gin-GE8 (5) <mark>M</mark>	LIGYVRVST	ID <mark>Q</mark> NTDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	. <mark>VVWKLDRLGF</mark>	RS <mark>M</mark> EHLISI
Gin-GE9 (4) <mark>M</mark>	LIGYVRVST <mark>S</mark>	D <mark>Q</mark> STDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	. <mark>VVWKLDRLGF</mark>	RS <mark>L</mark> KHLISI
Gin-GE13(4)M	LIGYVRVST <mark>N</mark>	ID <mark>Q</mark> NTDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	JVVWKLDRLGF	RS <mark>M</mark> EHLISI
Gin-GE17 <mark>M</mark>	LIGYVRVST <mark>N</mark>	ID <mark>Q</mark> NTDLQR <mark>N</mark> A	LV <mark>YAGCEQIFE</mark>	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	JVVWKLDRLGF	RS <mark>MK</mark> HLISI
Gin-GE20(2)M	LIGYVRVST <mark>N</mark>	ID <mark>Q</mark> NTDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	JVVWKLDRLGF	RS <mark>V</mark> KHLISI
Gin-GE24(2)M	LIGYVRVST <mark>N</mark>	ID <mark>QN</mark> TDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	JVVWKLDRLGF	RS <mark>MK</mark> HLISI
Gin-GE27(2)M	LIGYVRVST <mark>N</mark>	IG <mark>QN</mark> TDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	JVVWKLDRLGF	RS <mark>LK</mark> HLISI
Gin-GE31 <mark>M</mark>	LIGYVRVST <mark>N</mark>	IG <mark>QN</mark> TDLQR <mark>N</mark> A	LV <mark>C</mark> AGCEQIFE	DKLSGTRTE	RPGLKRAL	<mark>K</mark> RLQKGDTI	LVVWKLDRLGF	RS <mark>M</mark> EHLISI
ConsensusM	LIGYVRVSTN	IDQNTDLQRNA	LVCAGCEQIFE	DKLSGTRTE	RPGLKRAL	KRLQKGDTI	JVVWKLDRLGF	RSMKHLISI
_	80	90	100	110	120	130	140	15
Gin <mark>L</mark>	VGELRERGI <mark>N</mark>	IFRSLTDS <mark>I</mark> DT	SPMGRFFF <mark>Y</mark> V	/ <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY
GinL7C7	VGELRERGI <mark>N</mark>	J <mark>FRSLTDS</mark> ▼DT	SPMGRFFF <mark>H</mark> V	/ <mark>MG</mark> ALAE <mark>V</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY ·
Gin-GE1 (9)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> IDT	SPMGRFFF <mark>Y</mark> V	/ <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY ·
Gin-GE8 (5)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS<mark>L</mark>DM</mark>	SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY ·
Gin-GE9 (4)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> ID <mark>T</mark>	SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	5GSGEKPY <mark>1</mark>
Gin-GE13(4)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS<mark>L</mark>DM</mark>	<mark>S</mark> SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY <mark>.</mark>
Gin-GE17	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> IDT	P <mark>SPMGRFFF</mark> Y\	7 <mark>M</mark> CALAE <mark>M</mark> EF	ELI <mark>V</mark> ERTM	<mark>AGLAAAR</mark> S <mark>F</mark>	GRIGGRPPKS	SGSGEKPY <mark>.</mark>
Gin-GE20(2)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> I <mark>D</mark> T	SPMGRFFF <mark>Y</mark> V	/T <mark>G</mark> ALAE <mark>V</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	5GSGEKPY <mark>1</mark>
Gin-GE24(2)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> IDT	<mark>S</mark> SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	G <mark>LI I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY <mark>.</mark>
Gin-GE27(2)	VGELRERGI <mark>N</mark>	I <mark>FRSLTDS</mark> IDT	<mark>S</mark> SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY <mark>.</mark>
Gin-GE31	VGELRERGI	I <mark>FRSLTDS</mark> ID	<mark>S</mark> SPMGRFFF <mark>Y</mark> V	7 <mark>MG</mark> ALAE <mark>M</mark> EF	ELI <mark>I</mark> ERTM	AGLAAAR <mark>N</mark> K	GRIGGRPPKS	SGSGEKPY <mark>.</mark>
Consensus	VGELRERGIN	IFRSLTDSIDT	SSPMGRFFFYV	/MGALAEMEF	ELIIERTM	AGLAAARNK	GRIGGRPPK	SGSGEKPYI
								$\gamma \rightarrow \gamma$
								Finder

Finger Protein

Figure S1. Individual sequences of isolated variants from round 3 of evolutions of the Gin catalytic domain on the GE substrate, as shown in Figure 4. The number in parentheses indicates the number of sequenced colonies in which that particular sequence was repeated. The activities of Gin-GE1 and Gin-GE8 were tested independently, as shown in Figure 5A. Several mutations were repeated and/or similar in chemistry and are highlighted in Figure 6. Additionally, many selected residues showed chemical homology to mutations selected in previous evolutions, such as those in the GinL7C7 variant (D12G, N14S, M70V/L, I94V/L, M114V) (10).

Supporting Information Figure S2: Sequences of Evolved GinL7C7 Variants on EE Substrate



Figure S2. Individual sequences of isolated variants from round 3 of evolutions of the GinL7C7 catalytic domains on the EE substrate, as shown in Figure 4. All sequences were unique among the colonies that were sequenced. The activities of GinL7C7-EE2 and GinL7C7-EE3 were tested independently, as shown in Figure 5B. Several mutations were repeated and/or similar in chemistry and are highlighted in Figure 6.

	Gin-GE	Gin-GE	GinL7C7	GinL7C7-EE
	total sequences	unique sequences		unique sequences
V6A				3/8
N11S	13/30	2/9		
N14S	13/30	2/9	Х	NA
N20D			Х	NA
K50E			Х	NA
M70L/V	17/30	4/9	Х	NA
N86H	1/30	1/9		1/8 (N86D)
D95N				1/8
M114V	2/30	1/9	Х	NA
I120V	1/30	1/9		2/8

Table S2: Summary of Mutations Conserved Among the Serine Recombinase Family