Supplementary data for

Alternative σ^I/anti-σ^I factors represent a unique form of bacterial σ/anti-σ complex

Zhen Wei[1,2,3,4], Chao Chen[1,2,3], Ya-Jun Liu[1,2,3], Sheng Dong[1,2,3], Jie Li[1,2,3,4], Kuan Qi[1,2,3,4], Shiyue Liu[1,2,3,4], Xiaoke Ding[1,2,3], Lizett Ortiz de Ora[5], Iván Muñoz-Gutiérrez[6], Yifei Li[1,2,3], Hongwei Yao[7], Raphael Lamed[5], Edward A. Bayer[6], Qiu Cui[1,2,3], Yingang Feng[1,2,3,*]

[1] CAS Key Laboratory of Biofuels, Qingdao Institute of BioEnergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China
[2] Shandong Provincial Key Laboratory of Energy Genetics, Qingdao Institute of BioEnergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China
[3] Qingdao Engineering Laboratory of Single Cell Oil, Qingdao Institute of BioEnergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, China
[4] University of Chinese Academy of Sciences, Beijing 100049, China
[5] Department of Molecular Microbiology and Biotechnology, Tel Aviv University, Tel Aviv, Israel
[6] Department of Biomolecular Sciences, The Weizmann Institute of Science, Rehovot, Israel
[7] High-Field Nuclear Magnetic Resonance Center, Xiamen University, 422 South Siming Road, Xiamen 361005, China

[*] To whom correspondence should be addressed. Mail address: 189 Songling Road, Laoshan District, Qingdao 266101, China. Tel.: +86-532-80662706. E-mail: fengyg@qibebt.ac.cn

Present address: Lizett Ortiz de Ora, Department of Chemistry, University of California, Irvine, California, USA; Iván Muñoz-Gutiérrez, Outreach Research Training and Minority Science Programs, School of Biological Sciences, University of California, Irvine, California, USA; Yifei Li, Vonsun Pharmatech (Suzhou) Co., Ltd., Room 213, Building A4, No 218 Xinghu Street, Suzhou industrial park, Suzhou, Jiangsu, 215000, China.

Table of Contents

Figure S1. Sequence alignments of the eight pairs of $RsgI_N$ (A) and $SigI_C$ (B) from *C. thermocellum*.
Figure S2. The ^1H-^{15}N HSQC spectrum of $RsgI1_N$ with peak assignments.
Figure S3. The NMR titrations of RsgI1N and SigI1 domains.
Figure S4. Mass spectrometry results for RNAP component identification of the purified *C. thermocellum* RNAP.
Figure S5. The ^1H-^{15}N HSQC spectrum of the $RsgI1_N$-$Sig1I_C$ complex with peak assignments.
Figure S6. Structure comparison of $RsgI1_N$ in free form and in the $RsgI1_N$-$SigI1_C$ complex.
Figure S7. The ^1H NMR spectra of $RsgI1_N$ mutants.
Figure S8. NMR titrations for mutated sites that contributed to specific recognition of $RsgI1_N$ for $SigI1_C$.

Table S1. Primer sequences and restriction enzymes used for protein expression constructs in this study.
Table S2. Specific primer sequences for plasmid mutagenesis used in this study.
Table S3. Sequence identities among different pairs of $RsgI_N$ or $SigI_C$.
Table S4. Structural statistics of $RsgI1_N$ and the $SigI1_C$-$RsgI1_N$ complex.
Table S5. Numbers of identical residues in the interfaces of the $RsgI_N$-$SigI_C$ complexes.
Table S6. Intermolecular interactions in the structure of the $SigI1_C$-$RsgI1_N$ complex and the structural model of the $SigI2_C$-$RsgI2_N$ complex.

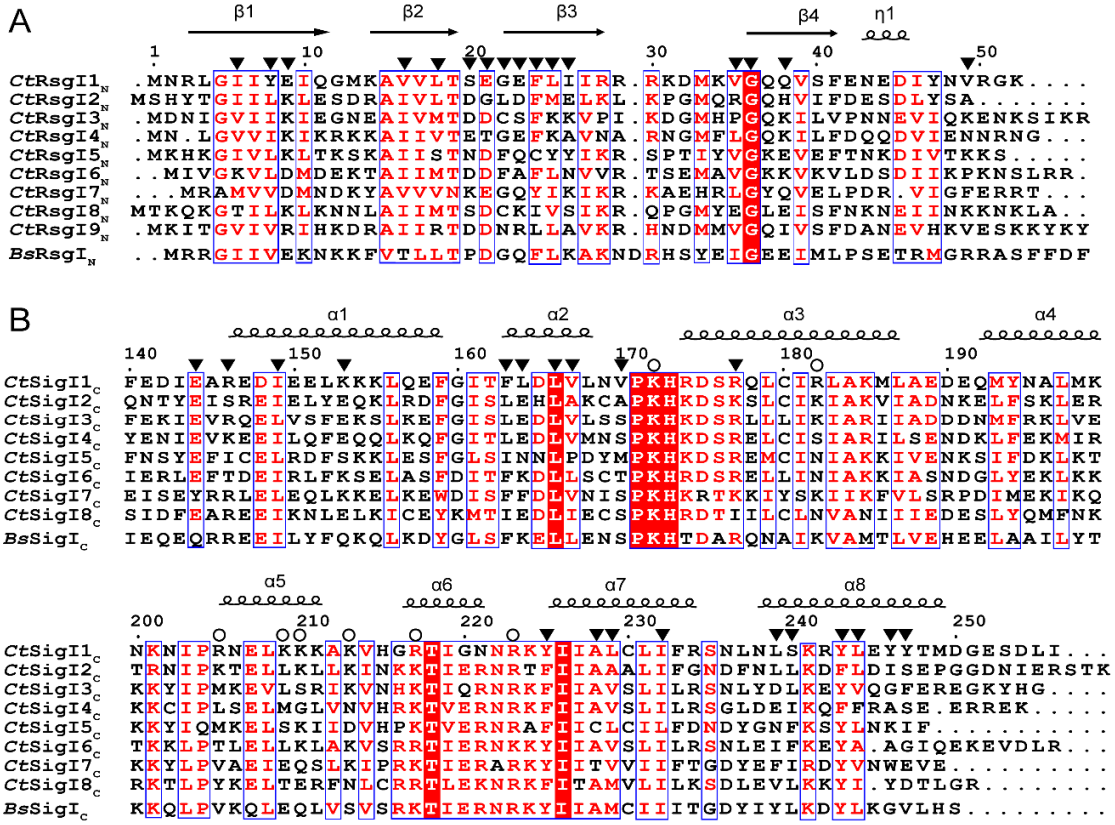


Figure S1. Sequence alignments of the eight pairs of RsgI$_N$ (A) and SigI$_C$ (B) from *C. thermocellum*. The sequences of RsgI$_N$ and SigI$_C$ from *Bacillus subtilis* (*Bs*RsgI$_N$ and *Bs*SigI$_C$, respectively) are also shown at the bottom of the alignments. Secondary structures of the SigI1$_C$-RsgI1$_N$ complex are indicated at the top of the sequences. Red-highlighted frames indicate residues that are identical, while blue frames represent conserved residues. Residues involved in the interactions between SigI1$_C$ and RsgI1$_N$ are indicated by filled triangles. Positively charged residues for potential promoter binding are indicated by open circles.

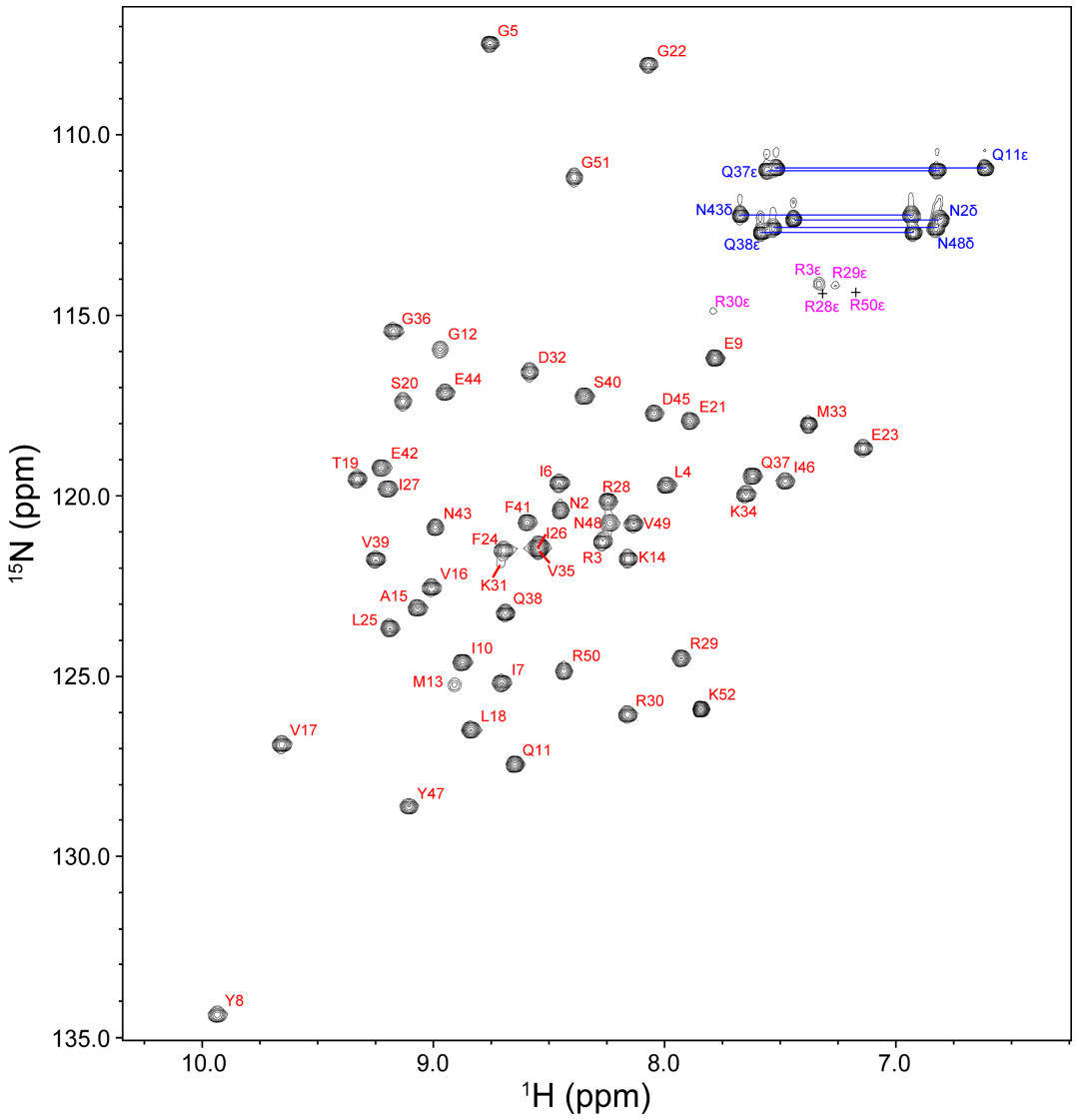


Figure S2. The ^{1}H-^{15}N HSQC NMR spectrum of RsgI1$_N$ with peak assignments. Resonance peaks for backbone and side chain amide groups were labeled with one-letter code for amino acid residue type followed by the sequence position. Side chain NH_2 resonance peaks of Asn and Gln were labeled in blue with an accompanying δ and ε, respectively, and were connected by a horizontal line. Resonance peaks corresponding to the side chain $NH_ε$ of Arg were labeled with an accompanying ε in magenta. The resonance peaks of Arg side chains were folded into the spectrum from their original ^{15}N chemical shifts by adding a spectrum width of 30 ppm.

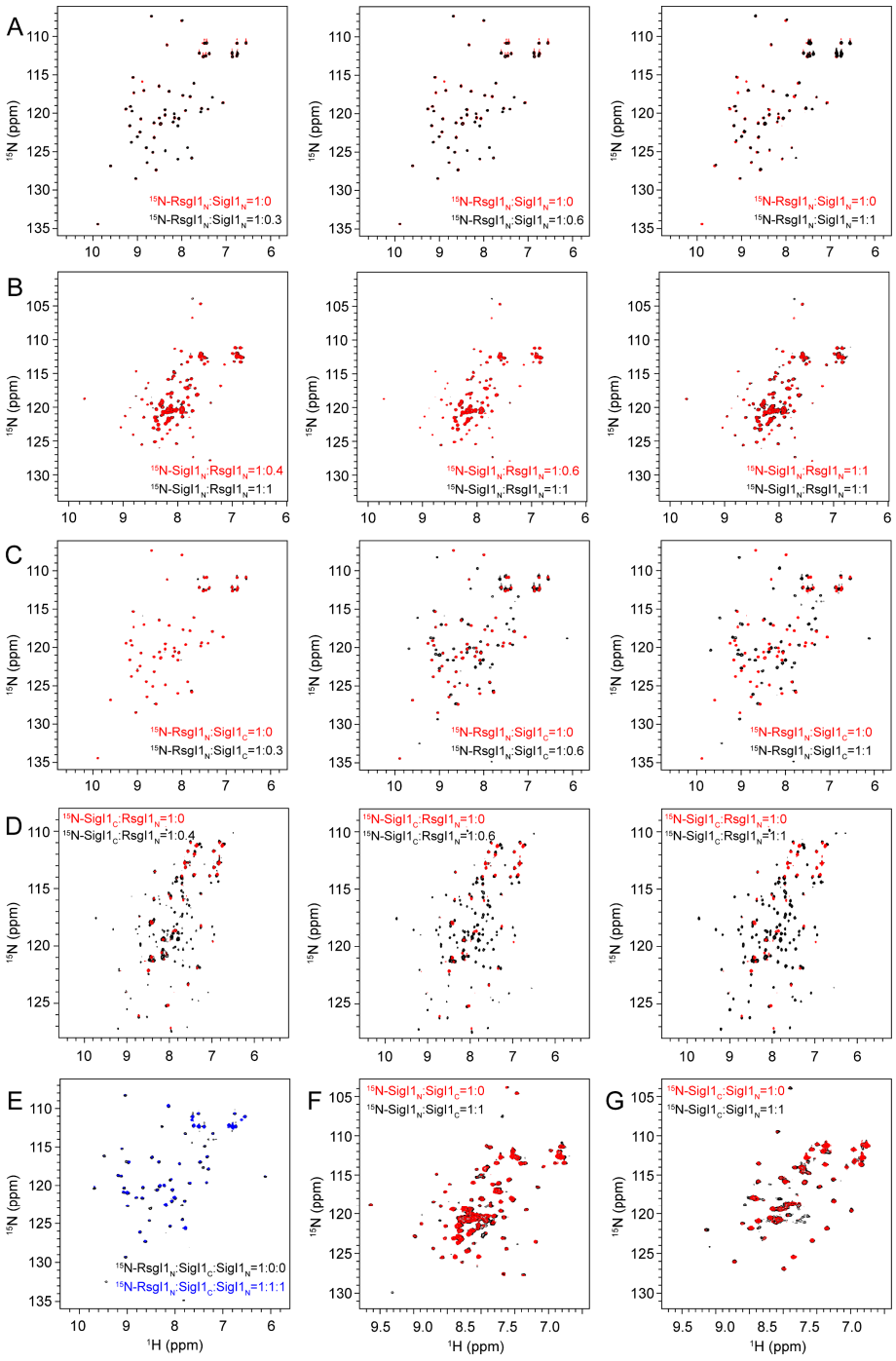


Figure S3. NMR titrations of RsgI1$_N$ and SigI1 domains. (A) ^{1}H-^{15}N HSQC spectra of RsgI1$_N$ titrated with different ratios of SigI1$_N$. (B) ^{1}H-^{15}N HSQC spectra of SigI1$_N$ titrated with different ratios of RsgI1$_N$. (C) ^{1}H-^{15}N HSQC spectra of RsgI1$_N$ titrated with different ratios of SigI1$_C$. (D) ^{1}H-^{15}N HSQC spectra of SigI1$_C$ titrated with different ratios of RsgI1$_N$. (E) ^{1}H-^{15}N HSQC spectra of RsgI1$_N$ in complex with SigI1$_C$ titrated with SigI1$_N$. (F) ^{1}H-^{15}N HSQC spectra of SigI1$_N$ titrated with SigI1$_C$. (G) ^{1}H-^{15}N HSQC spectra of SigI1$_C$ titrated with SigI1$_N$.

RNA polymerase, α subunit

MIEIEKPKIE CVVCSEDNRY GKFVVEPLER GYGITLGNSL RRILLSSLPG VAVTSIKIDG VLHEFSTIPG VIEDVTEIIL NIKELSLNFH GEGPKVIYID
AEGEGEVKAK DIKADADVEI LNPEHKIATL SGDHRLYMEM TIDKGRGYVS AEKNKHPGQP IGVIPVDSIF TPVHKVNYTV ENTRVGQVTD YDKLTLEVWT
NGSIKPDEAI SLGAKILSEH LNLFIDLSDN AKNAEIMVEK EETKKEKVLE MTIEELDLSV RSYNCLKRAG INTVEDLISR TEEDMMKVRN LGRKSLEEVV
NKLKALGLSL APSED

RNA polymerase, β subunit

MVHPVKLGRN VRMSYSKIDE VIDMPNLIEI QKNSYEQFLK EGFKEVFKDV NPITDYTGNL ILEFVDYSLD EPPKYSVDEC KERDATYAAP LKVKVRLINK
ETGEVKEQEI FMGDFPLMTE TGTFIINGAE RVIVSQLVRS PGIYYAMKID KAGKQLFSNT VIPNRGAWLE YETDSNDVLS VRIDRTRKLP LTVLVRALGY
GTDLEITELF GEDERILATI QKDSTKTEEE GLLETYKRLR PGEPPTVESA KALLHGLFFD PKRYDLAKPG RFKFNKKLSI AARIHGFIAG ENIKDPDTGE
LIVAEGETIS REKAETIQNA GVNTVIIRVD GKNVKVIGND MVDIKRYVDE DPKEIGINEK VKRDVLMEIL EEYKGKGDDA IKKALQERID DLIPKHITKE
DIISSISYII GLSYGIGSTD DIDHLGNRRL RSVGELLQNQ FRIGLSRMER VVRERMTIQD LDVVTPQALI NIRPVAAAIK EFFGSSQLSQ FMDQTNPLAE
LTHKRRLSAL GPGGLSRERA GFEVRDVHHS HYGRMCPIET PEGPNIGLIG SLSTYARVNE YGFIETPYRK VSKEEPGKVT NEIVYLTADE EDEYIIAQAN
EPLDEEGRFI SNKVVCRYKE EFIEVDPSKI DFMDVSPKQI VSVATSMIPF LENDDANRAL MGANMQRQAV PLIKTESPIV GTGIEYRAAR DSGVVILAKN
PGVVEKVTAN EIIIRTKDGK RDTYKLLKYM RSNQGTCINQ RPIVKKGEEV EAGDVIADGP STDNGEIALG KNVLVGFMTW EGYNYEDAIL ISERLVKDDV
FTSIHIEEYE AEARDTKLGP EDITREIPNV SEDALKDLNS EGIIRIGAEV RAGDILVGKV TPKGETELTA EERLLRAIFG EKAREVRDTS LRVPHGESGI
VVDVKIFTRE NGDELAPGVN KLVRVYVAQK RKISVGDKMA GRHGNKGVIS RILPVEDMPF LPDGTPLDIV LNPLGVPSRM NIGQVLEVHL GYAAKALGWK
VATPVFDGAT EEDIVQTLRK AGLAEDGKSI LYDGRTGEPF ENRVTVGYMY MLKLAHLVDD KIHARSTGPY SLVTQQPLGG KAQFGGQRFG EMEVWALEAY
GAAYTLQEIL TVKSDDVVGR VKTYEAIVKG ENVPEPGIPE SFKVLIKELQ SLCLDVKVYS EEQEEIAIKE SVEDDLEELN VNIEGREDEV NFNEFNDIGE
EITDEDLEVE DFDLQDLNDD DINPDDTIDA ELDDNLFDDD FDDTFDDDDL

RNA polymerase, β' subunit

MFELNNFDSI RIGLASPEKI REWSRGEVKK PETINYRTLK PERDGLFCER IFGPQKDWEC HCGKYKRIRY KGIVCDRCGV EVTRSKVRRE RMGHIELAAP
VSHIWYFKGI PSRMGLLLDM SPRALEKILY FAAYVVIDPG QTPLSKKQIL SEKEYRDSLE KFGPKFRAGM GAEAVRELLQ EINLDELSAE LREEIKQSTG
QKRVRAIKRL EVVEAFRQSQ NKPEWMILDV IPVIPPELRP MVQLDGGRFA TSDLNDLYRR VINRNNRLKR LLDLGAPDII VRNEKRMLQE AVDALIDNGR
RGRPVTGPGN RPLKSLSDML KGKQGRFRQN LLGKRVDYSG RSVIVVGPEL KIYQCGLPKE MALELFKPFV MKKLVNDGLA HNIKSAKRMV ERVRNEVWDV
LEEVIKEHPV LLNRAPTLHR LGIQAFEPVL VEGRALKLHP LVCTAYNADF DGDQMAIHVP LSAEAQAEAR FLMLSANNLL KPQDGKPVAV PTQDMVLGSY
YLTILKEGAK GEGRVFTSMD EAVMAYDNGE IELHSKIKVR MKRVVDGVEK SKIIETTLGR LIFNEAIPQD LGFVDRSDPD KIFDLEVDFL VGKNELLKII
DKSIKVHGTT KTAILLDKIK ELGFKYSTKG AITISISDMV IPEVKAKYIK ETEEKIEKIT KQYKRGLISD EERYNSVIAA WTEASENITR ALINNLDRFN
PVYMMSQSGA RGNINQIKQL AGMRGLMADT SGKTIEFPIK ANFREGLTVM EFFISTHGAR KGLADTALRT ADSGYLTRRL VDVSQDVIVR ETDCGTRKGI
EVTDIKDGNE VIEELSERII GRYPVGNIVH PETGEIIVEA GRMITDQDAE KIVKAGIKKV RIRSVLTCHS EYGVCAKCYG ANLATGEECN VGEAVGIIAA
QSIGEPGTQL TMRTFHTGGV AGEDITQGLP RVEELFEARK PKGLAIISEI KGTVKISETK KKREIVVTSE DGETRSYLIP YGSRIKVSDG DQVEAGDELT
EGSVNPHDIL KIKGVEAVQT YLVHEVQKVY RMQGVDINDK HIEVIVRQML RKVKVEDPGD TSLLPGGLVD VFDFEEENAK AIAEGKKPAV AKRALLGITK
AALATDSFLS AASFQETTRV LTEAAIKGKV DPLVGLKENV IIGKLIPAGT GMSRYKDITI STVTE

RNA polymerase, ω subunit

MKEKKERVSS MIEPSINSLL EKVDSRYTLV VATAKRARQL TDGANKLTNC ESDKPVTVAI NEINENKITY IFTKSGIK

Figure S4. Mass spectrometry results for RNAP component identification of the purified *C. thermocellum* RNAP. The identified peptides in mass spectrometry are highlighted in green.

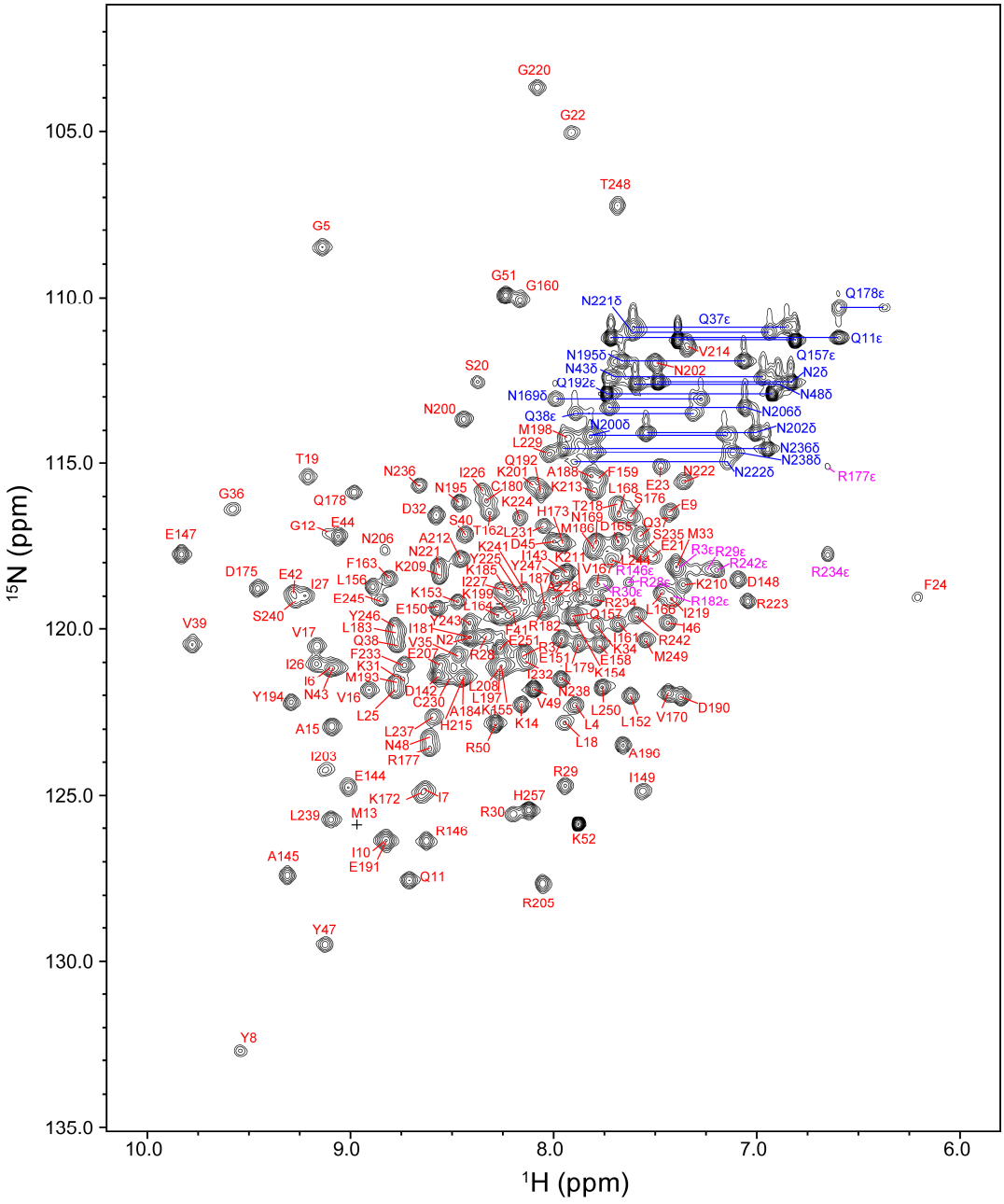


Figure S5. The ^{1}H-^{15}N HSQC NMR spectrum of the RsgI1$_{N}$-Sig1I$_{C}$ complex with peak assignments. Resonance peaks for backbone and side chain amide groups were labeled with one-letter code for amino acid residue type followed by the sequence position. Side chain NH_2 resonance peaks of Asn and Gln were labeled in blue with an accompanying δ and ε, respectively, and were connected by a horizontal line. Resonance peaks corresponding to the side chain NH_ε of Arg were labeled with an accompanying ε in magenta. The resonance peaks of Arg side chains were folded into the spectrum from their original ^{15}N chemical shifts by adding a spectrum width of 34 ppm.

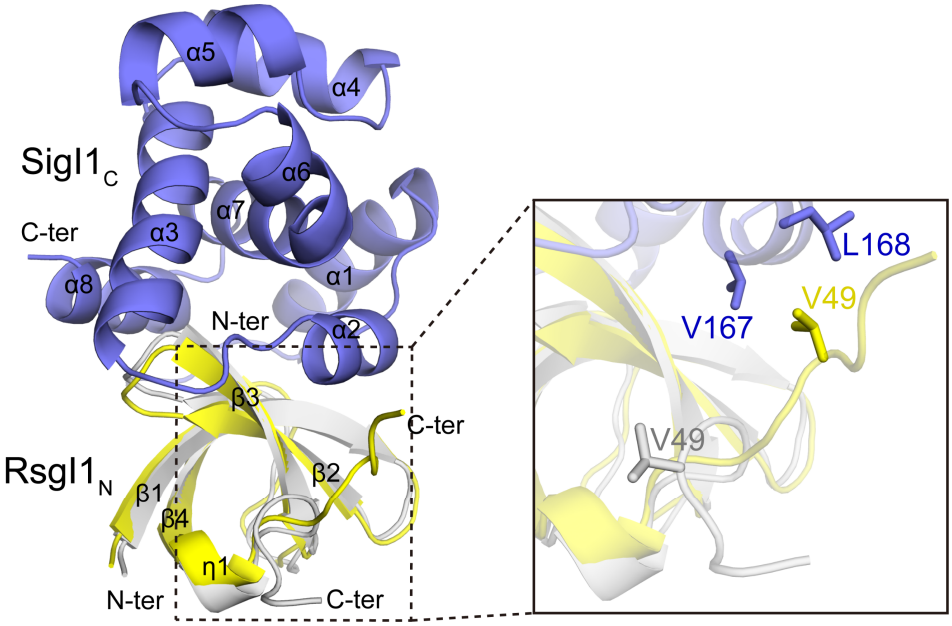
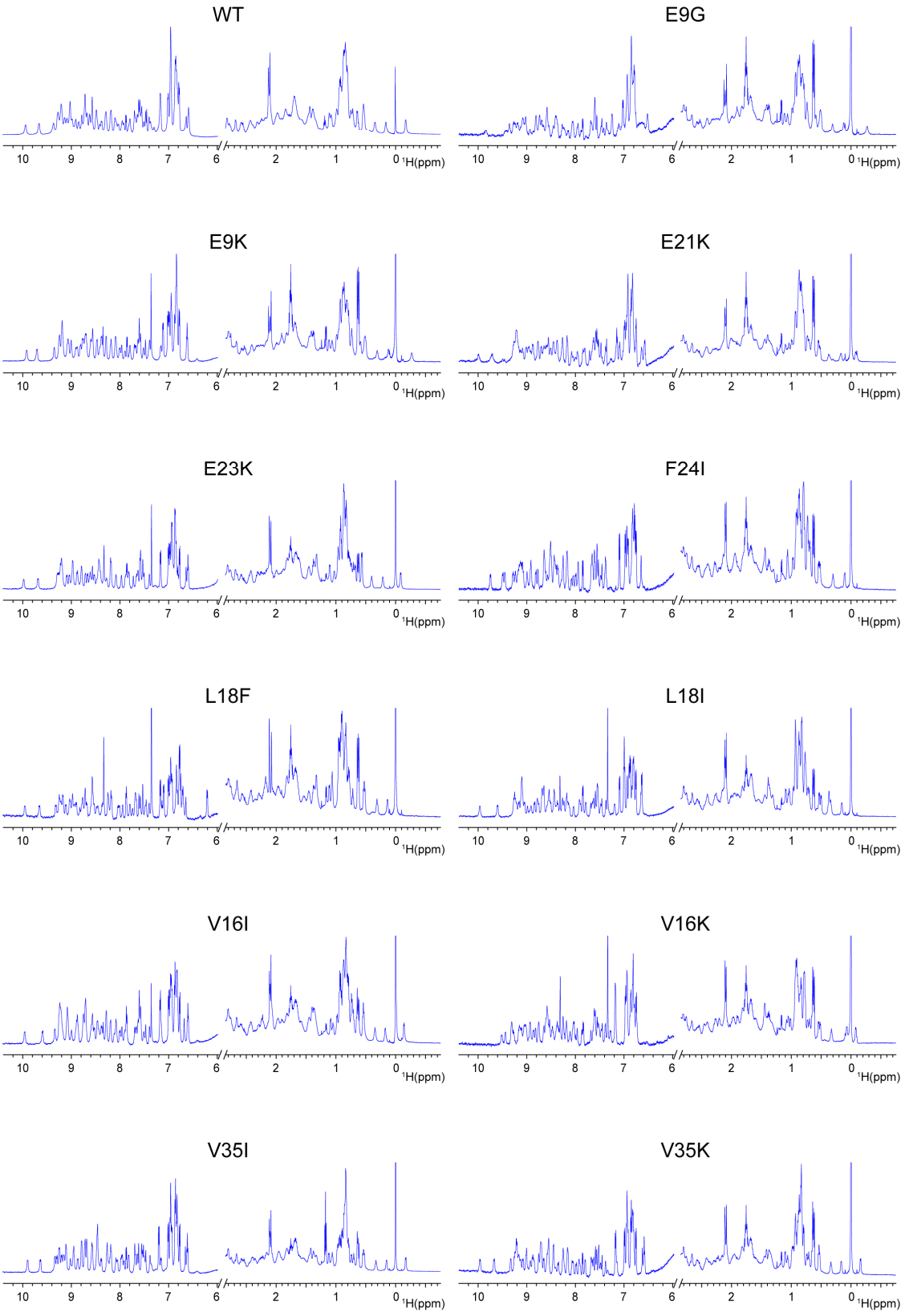


Figure S6. Structure comparison of $RsgI1_N$ in free form (light grey) and in the $RsgI1_N$-$SigI1_C$ complex (blue and yellow). The C-terminal difference caused by the interaction between V49 of $RsgI1_N$ and V167/L168 of $SigI1_C$ is shown in detail.


WT
E9G
E9K
E21K
E23K
F24I
L18F
L18I
V16I
V16K
V35I
V35K
10 9 8 7 6 2 1 0
¹H(ppm)

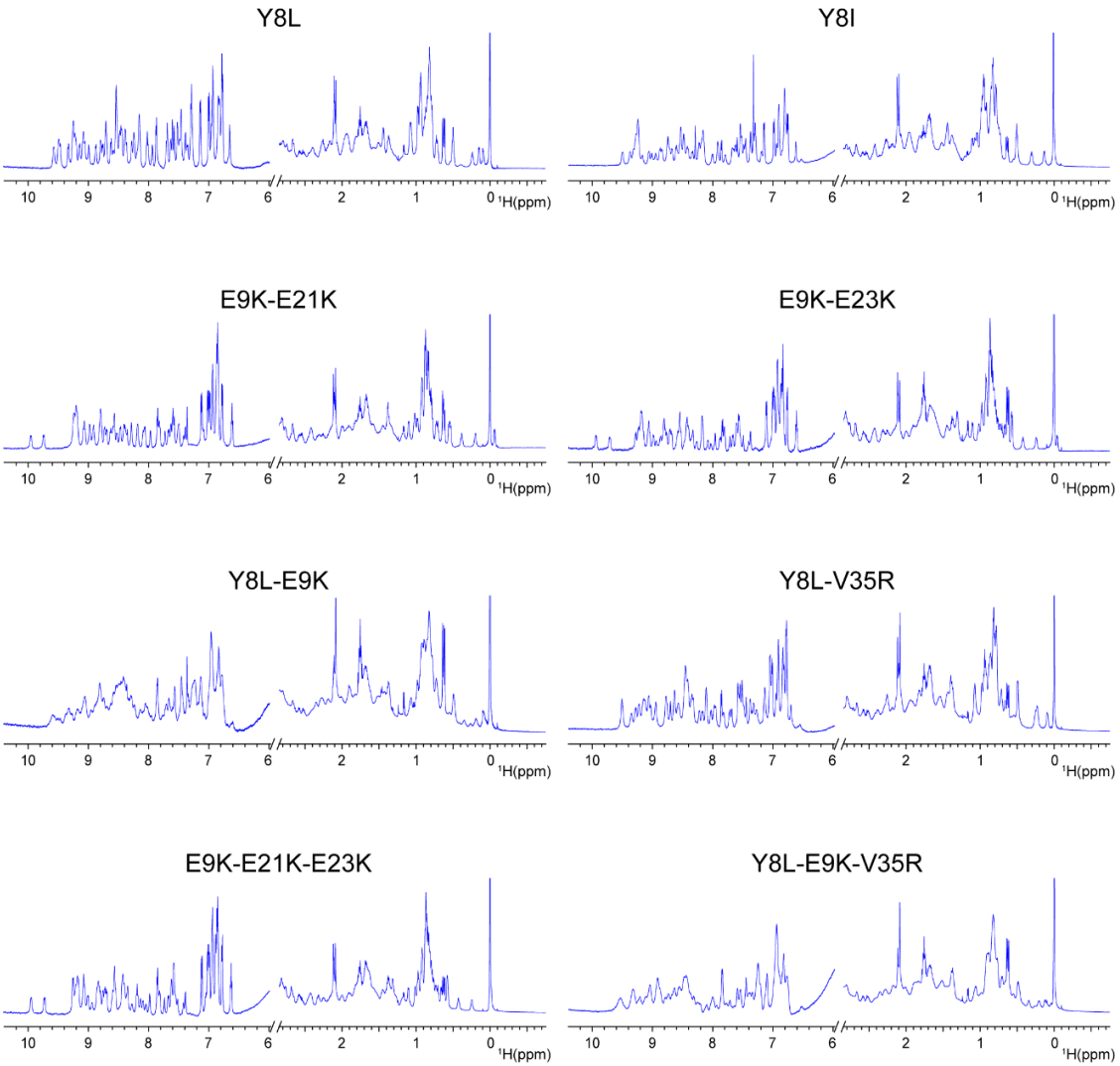


Figure S7. The ^{1}H NMR spectra of RsgI1$_{N}$ mutants. The proper folding of each mutant is verified by the well-dispersed peaks in both the low field (7.0-10.2 ppm) and the high field (-0.5-1.0 ppm) regions.

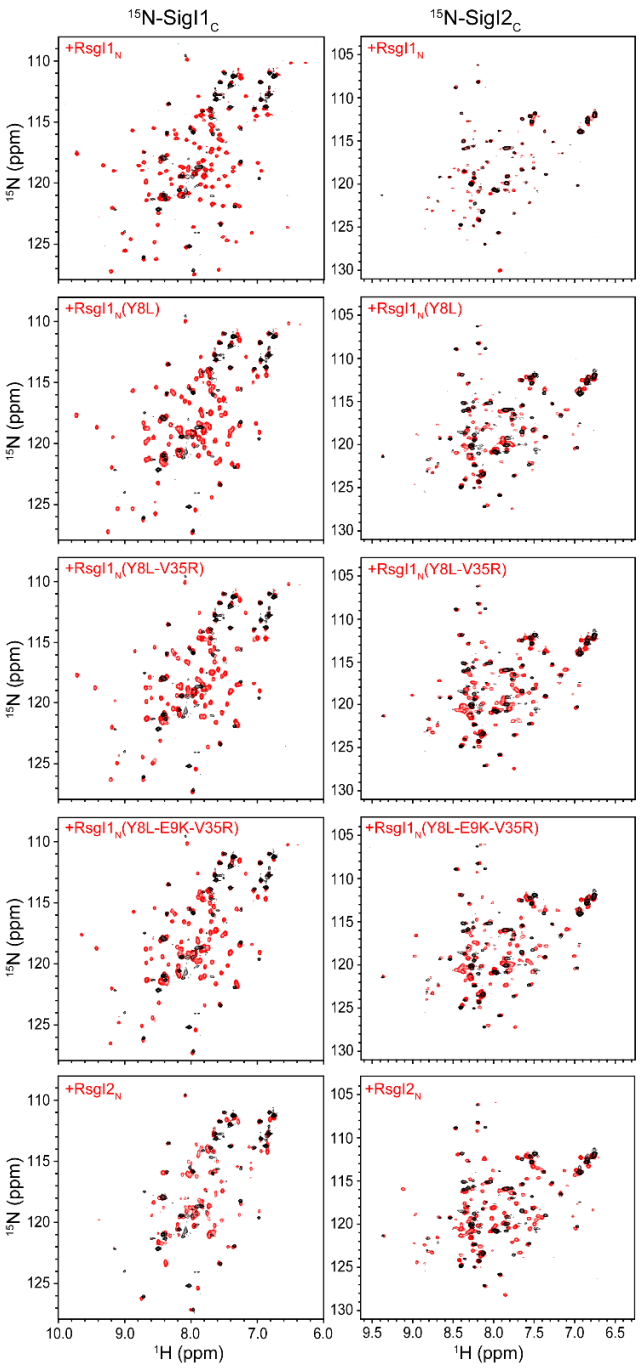


Figure S8. NMR titrations for mutated sites that contribute to specific recognition of RsgI1$_N$ for SigI1$_C$. The ^{1}H-^{15}N HSQC spectra of SigI1$_C$ (left) and SigI2$_C$ (right) without and with wild-type RsgI1$_N$, its mutants, or RsgI2$_N$ (molar ratio 1:1) are shown in black and red, respectively. The free SigI2$_C$ and SigI1$_C$ have similar spectra, and only a few crowded peaks can be observed in the central region of the spectra. The addition of wild-type RsgI2$_N$ into SigI1$_C$ resulted in a well-dispersed spectrum, but the addition of wild-type RsgI1$_N$ into SigI2$_C$ resulted in only a slight change in the spectrum, thus indicating the specificity of the two RsgI factors. By adding the single, double, and triple mutants of RsgI1$_N$, the number of dispersed peaks successively increased in the spectrum

of $SigI2_C$, which indicates that the $SigI2_C$-$RsgI1_N$ interaction is enhanced by these mutations. In contrast, upon the addition of these mutants, the spectrum of $SigI1_C$ displayed successively less dispersed peaks but more crowded peaks, compared with the spectra of $SigI1_C$ and the wild-type $RsgI1_N$. Therefore, the evidence supports the contention that these residues play important roles for specific σ^I/anti-σ^I factor recognition between the two complementary pairs of σ^I-anti-σ^I factors. The triple mutants of $RsgI1_N$, however, still failed to abolish the interaction with $SigI1_C$ completely, and their interactions with $SigI2_C$ were not as strong as that of the wild-type $RsgI2_N$,

Table S1. Primer sequences and restriction enzymes used for protein expression constructs in this study. The restriction enzyme sites in the primer sequences are underlined.

Recombinant plasmid	Primer sequences	Restriction enzymes
pET28a-SMT3-SigI1	F, 5'- ATTGGATCCGTGGAAGTCCGGAAAAT R, 5'-ATTCTCGAGTTATATAAGGTCACTTTCCCC	*Bam*HI, *Xho*I
pET30a-SigI1$_N$	F, 5'-TTACATATGGAAGTCCGGAAAATTAATACC R, 5'-TTTCTCGAGGTTATAAAAATATTCGTCATCA	*Nde*I, *Xho*I
pET30a-SigI1$_C$	F, 5'-TTACATATGGAAGATATTGAGGCAAGG R, 5'-TTTCTCGAGCATAGTATAATACTCCAAATA	*Nde*I, *Xho*I
pET28a-SMT3-RsgI1$_N$	F, 5'-ATTGGATCCATGAACAGATTGGGAATAAT R, 5'-ATTCTCGAGTTACTTTCCCCTGACA	*Bam*HI, *Xho*I
pET28a-SigI2	F, 5'- ATACCATGGGCATTGATTTGTTTTCCCCTAAGG R, 5'-TTACTCGAGTTTTGTGCTCCGTTCAATGTTG	*Nco*I, *Xho*I
pET28a-SigI2$_C$	F, 5'-TTACCATGGGCAGTGAAATTGATTTACAAA R, 5'-TTACTCGAGTTTTGTGCTCCGTTCAAT	*Nco*I, *Xho*I
pET28a-SMT3-RsgI2$_N$	F, 5'-ATTGGATCCATGTCACATTACACGGGAAT R, 5'-ATTCTCGAGTTAGGCTGAATACAAGTCGGA	*Bam*HI, *Xho*I

Table S2. Specific primer sequences for plasmid mutagenesis used in this study. The mutation sites in primer sequences are shown in red.

Mutant	Primer sequences	Plasmid template
$RsgI1_N$(Y8L)	F,5'-GAACAGATTGGGAATAATATTAGAAATTCAGGGCATGAAAG	pET28a-SMT3-$RsgI1_N$
	R,5'-CTTTCATGCCCTGAATTTCTAATATTATTCCCAATCTGTTC	
$RsgI1_N$(Y8I)	F,5'-GAACAGATTGGGAATAATAATTGAAATTCAGGGCATGAAAGC	pET28a-SMT3-$RsgI1_N$
	R,5'-GCTTTCATGCCCTGAATTTCAATTATTATTCCCAATCTGTTC	
$RsgI1_N$(E9K)	F,5'-ACAGATTGGGAATAATATATAAAATTCAGGGCATGAAAGCT	pET28a-SMT3-$RsgI1_N$
	R,5'-AGCTTTCATGCCCTGAATTTTATATATTATTCCCAATCTGT	
$RsgI1_N$(E9G)	F,5'-ACAGATTGGGAATAATATATGGAATTCAGGGCATGAAAGCT	pET28a-SMT3-$RsgI1_N$
	R,5'-AGCTTTCATGCCCTGAATTCCATATATTATTCCCAATCTGT	
$RsgI1_N$(V16K)	F,5'-GAAATTCAGGGCATGAAAGCTAAAGTTCTGACAAGCGAAGGCG	pET28a-SMT3-$RsgI1_N$
	R,5'-CGCCTTCGCTTGTCAGAACTTTAGCTTTCATGCCCTGAATTTC	
$RsgI1_N$(V16I)	F,5'-ATTCAGGGCATGAAAGCTATTGTTCTGACAAGCGAAGGCGAAT R,5'-	pET28a-SMT3-$RsgI1_N$
	ATTCGCCTTCGCTTGTCAGAACAATAGCTTTCATGCCCTGAAT	
$RsgI1_N$(L18I)	F,5'-GGGCATGAAAGCTGTAGTTATTACAAGCGAAGGCGAATTT	pET28a-SMT3-$RsgI1_N$
	R,5'-AAATTCGCCTTCGCTTGTAATAACTACAGCTTTCATGCCC	
$RsgI1_N$(L18F)	F,5'-CAGGGCATGAAAGCTGTAGTTTTTACAAGCGAAGGCGAAT	pET28a-SMT3-$RsgI1_N$
	R,5'-ATTCGCCTTCGCTTGTAAAAACTACAGCTTTCATGCCCTG	
$RsgI1_N$(E21K)	F,5'-GAAAGCTGTAGTTCTGACAAGCAAAGGCGAATTTTTGATTATTCGC	pET28a-SMT3-$RsgI1_N$
	R,5'-GCGAATAATCAAAAATTCGCCTTTGCTTGTCAGAACTACAGCTTTC	
$RsgI1_N$(E23K)	F,5'-TAGTTCTGACAAGCGAAGGCAAATTTTTGATTATTCGCAG	pET28a-SMT3-$RsgI1_N$
	R,5'-CTGCGAATAATCAAAAATTTGCCTTCGCTTGTCAGAACTA	
$RsgI1_N$(F24I)	F,5'-GTTCTGACAAGCGAAGGCGAAATTTTGATTATTCGCAGGCGCAAAG	pET28a-SMT3-$RsgI1_N$
	R,5'-CTTTGCGCCTGCGAATAATCAAAATTCGCCTTCGCTTGTCAGAAC	
$RsgI1_N$(I26E)	F,5'-GACAAGCGAAGGCGAATTTTTGGAAATTCGCAGGCGCAAAGATATG	pET28a-SMT3-$RsgI1_N$
	R,5'-CATATCTTTGCGCCTGCGAATTTCCAAAAATTCGCCTTCGCTTGTC	
$RsgI1_N$(V35K)	F,5'-GCAGGCGCAAAGATATGAAGAAAGGACAGCAGGTGAGTTTTG	pET28a-SMT3-$RsgI1_N$
	R,5'-CAAAACTCACCTGCTGTCCTTTCTTCATATCTTTGCGCCTGC	
$RsgI1_N$(V35I)	F,5'-ATTCGCAGGCGCAAAGATATGAAGATTGGACAGCAGGTGAGT	pET28a-SMT3-$RsgI1_N$
	R,5'-ACTCACCTGCTGTCCAATCTTCATATCTTTGCGCCTGCGAAT	
$RsgI1_N$(E9K-E21K)	F,5'-GAAAGCTGTAGTTCTGACAAGCAAAGGCGAATTTTTGATTATTCG	pET28a-SMT3-$RsgI1_N$(E9K)
	R,5'-CGAATAATCAAAAATTCGCCTTTGCTTGTCAGAACTACAGCTTTC	
$RsgI1_N$(E9K-E23K)	F,5'-TAGTTCTGACAAGCGAAGGCAAATTTTTGATTATTCGCAG	pET28a-SMT3-$RsgI1_N$(E9K)
	R,5'-CTGCGAATAATCAAAAATTTGCCTTCGCTTGTCAGAACTA	
$RsgI1_N$(Y8L-E9K)	F,5'-GAACAGATTGGGAATAATATTAAAAATTCAGGGCATGAAAGC	pET28a-SMT3-$RsgI1_N$
	R,5'-GCTTTCATGCCCTGAATTTTTAATATTATTCCCAATCTGTTC	
$RsgI1_N$(Y8L-V35R)	F,5'-GCAGGCGCAAAGATATGAAGCGTGGACAGCAGGTGAGTTTTG	pET28a-SMT3-$RsgI1_N$(Y8L)
	R,5'-CAAAACTCACCTGCTGTCCACGCTTCATATCTTTGCGCCTGC	
$RsgI1_N$(E9K-E21K-E23K)	F,5'-CTGTAGTTCTGACAAGCAAAGGCAAATTTTTGATTATTCGCAGG	pET28a-SMT3-$RsgI1_N$(E9K)
	R,5'-CCTGCGAATAATCAAAAATTTGCCTTTGCTTGTCAGAACTACAG	
$RsgI1_N$(Y8L-E9K-V35R)	F,5'-GATTGGGAATAATATTAAAAATTCAGGGCATGAAAGCTG	pET28a-SMT3-$RsgI1_N$(Y8L-
	R,5'-CAGCTTTCATGCCCTGAATTTTTAATATTATTCCCAATC	V35R)
$SigI1_C$(K209E)	F,5'-CATACCCAGAAATGAATTAGAAAAGAAAGCCAAGGTTCACGG	pET30a-$SigI1_C$
	R,5'-CCGTGAACCTTGGCTTTCTTTTCTAATTCATTTCTGGGTATG	
$SigI1_C$(R217E)	F,5'-GAAAGCCAAGGTTCACGGGGAGACCATAGGCAATAATAG	pET30a-$SigI1_C$
	R,5'-CTATTATTGCCTATGGTCTCCCCGTGAACCTTGGCTTTC	
$SigI1_C$(K209E-R217E)	F,5'-GAAAGCCAAGGTTCACGGGGAGACCATAGGCAATAATAG	pET30a-$SigI1_C$(K209E)
	R,5'-CTATTATTGCCTATGGTCTCCCCGTGAACCTTGGCTTTC	

Table S3. Sequence identities among different pairs of $RsgI_N$ or $SigI_C$.

$RsgI_N$ (%)	*Ct* RsgI1	*Ct* RsgI2	*Ct* RsgI3	*Ct* RsgI4	*Ct* RsgI5	*Ct* RsgI6	*Ct* RsgI7	*Ct* RsgI8	*Ct* RsgI9	*Bs* RsgI
*Ct*RsgI1	100	28.9	26.9	32.7	28.9	23.1	19.2	26.9	34.6	26.9
*Ct*RsgI2		100	30.0	26.0	28.0	24.0	12.0	30.0	32.0	15.4
*Ct*RsgI3			100	39.6	19.2	27.3	17.0	35.7	35.1	16.1
*Ct*RsgI4				100	23.5	26.9	19.2	25.0	32.1	15.4
*Ct*RsgI5					100	33.3	17.7	43.1	31.4	17.6
*Ct*RsgI6						100	21.2	30.8	29.1	16.7
*Ct*RsgI7							100	11.5	15.4	21.2
*Ct*RsgI8								100	30.9	14.5
*Ct*RsgI9									100	15.5
*Bs*RsgI										100
$SigI_C$ (%)	*Ct* SigI1	*Ct* SigI2	*Ct* SigI3	*Ct* SigI4	*Ct* SigI5	*Ct* SigI6	*Ct* SigI7	*Ct* SigI8		*Bs* SigI
*Ct*SigI1	100	36.0	42.7	39.3	33.6	45.3	27.3	40.9		36.0
*Ct*SigI2		100	43.1	40.2	39.1	38.5	28.2	30.0		32.4
*Ct*SigI3			100	52.1	40.4	44.4	33.6	34.6		35.1
*Ct*SigI4				100	39.1	45.3	28.2	36.4		38.7
*Ct*SigI5					100	37.3	30.0	25.5		35.5
*Ct*SigI6						100	32.7	38.2		36.0
*Ct*SigI7							100	24.6		36.9
*Ct*SigI8								100		31.5
*Bs*SigI										100

Table S4. Structural statistics of $RsgI1_N$ and the $SigI1_C$-$RsgI1_N$ complex.

Parameters	$RsgI1_N$	$SigI1_C$-$RsgI1_N$ complex
NOE restraints		
Intra-residue	591	1470
Sequential	346	864
Medium-range	184	570
Long-range	514	1012
Ambiguous	688	2356
Total	2323	6272
Number per residue	43.6	36.7
Hydrogen bond restraints	21	160
Torsion angle restraints		
Phi (Φ) angle restraints	45	148
Psi (Ψ) angle restraints	45	148
Chi (χ) angle restraints	23	73
Violations		
Max. NOE violation (Å)	0.140	0.179
Max. torsion angle violation (°)	2.01	4.57
R.M.S.D from mean structure (Å)		
Residues in regular secondary structure [a]		
Backbone heavy atoms	0.18±0.03	0.30±0.06
All heavy atoms	0.63±0.08	0.75±0.05
All residues [b]		
Backbone heavy atoms	1.40±0.26	0.70±0.16
All heavy atoms	1.83±0.25	1.10±0.11
Ramachandran statistics		
Most favored region (%)	89.9	91.9
Additionally allowed (%)	10.0	7.4
Generously allowed (%)	0.1	0.3
Disallowed (%)	0.0	0.5 [c]
WHAT_CHECK Z-scores [d]		
1st generation packing quality	-0.477	-0.275
2nd generation packing quality	2.279	-0.292
Ramachandran plot appearance	-2.622	-2.461
chi-1/chi-2 rotamer normality	-2.839	-2.448
Backbone conformation	-0.039	-0.240
Inside/Outside distribution	1.013	0.978

[a] The residues in regular secondary structure include 2-11, 14-19, 24-28, 38-42, and 43-45 of $RsgI1_N$ and 146-159, 163-167, 174-188, 191-200, 205-211, 216-221, 223-234, and 238-248 of $SigI1_C$.

[b] The artificial residues at the N- and C-termini introduced by cloning are excluded.

[c] All residues in the disallowed region are in the terminal disordered loops.

[d] For Z-scores, a more positive value is better.

Table S5. Number of identical residues in the interface of the $RsgI_N$-$SigI_C$ complexes.

$RsgI_N$	*Ct* RsgI1	*Ct* RsgI2	*Ct* RsgI3	*Ct* RsgI4	*Ct* RsgI5	*Ct* RsgI6	*Ct* RsgI7	*Ct* RsgI8	*Ct* RsgI9	*Bs* RsgI
*Ct*RsgI1	16	4	2	4	3	4	5	2	4	7
*Ct*RsgI2		16	5	4	5	5	1	4	3	4
*Ct*RsgI3			16	8	5	7	2	7	5	4
*Ct*RsgI4				16	3	4	4	3	4	3
*Ct*RsgI5					16	6	2	7	4	5
*Ct*RsgI6						16	2	5	6	4
*Ct*RsgI7							16	1	2	5
*Ct*RsgI8								16	3	3
*Ct*RsgI9									16	4
*Bs*RsgI										16
$SigI_C$	*Ct* SigI1	*Ct* SigI2	*Ct* SigI3	*Ct* SigI4	*Ct* SigI5	*Ct* SigI6	*Ct* SigI7	*Ct* SigI8		*Bs* SigI
*Ct*SigI1	20	7	8	7	6	10	8	8		10
*Ct*SigI2		20	9	11	4	5	3	9		6
*Ct*SigI3			20	12	6	7	9	11		9
*Ct*SigI4				20	4	8	6	9		6
*Ct*SigI5					20	5	3	5		4
*Ct*SigI6						20	7	6		10
*Ct*SigI7							20	5		7
*Ct*SigI8								20		9
*Bs*SigI										20

Table S6. Intermolecular interactions in the structure of the SigI1$_C$-RsgI1$_N$ complex and the structural model of the SigI2$_C$-RsgI2$_N$ complex.

	RsgI1$_N$-SigI1$_C$	RsgI2$_N$-SigI2$_C$
Hydrophobic interactions [a]	(I6)...(L239)	(I7)...(L245, L246, F249)
		(L9)...(I155, L169, F243, L246)
	(V16)...(F163, L164, V167)	(I17)...(L169)
	(L18)...(L229, I232, L244)	(L19)...(L246, L250)
		(L23)...(I187, F249, L250)
	(F24)...(F163, L166, V167, V170, Y225, A228, L229, I232)	(F25)...(L172, F231, A234)
	(L25)...(V167, V170)	
	(I26)...(L164, V167)	
	(V35)...(R146, I149)	
	(V49)...(V167,V170)	
Hydrogen bonds [b]	(Y8 $O^{\eta}H^{\eta}$)...(S240 O^{γ})	
	(S20 O)...(Y246 $O^{\eta}H^{\eta}$)	
	(E21 O^{ε})...(Y247 $O^{\eta}H^{\eta}$)	
	(G22 O)...(Y225 $O^{\eta}H^{\eta}$)	
		(R36 O)...(S152 $O^{\gamma}H^{\gamma}$)
	(G36 O)...(E144 NH)	(G37 O)...(E150 NH)
	(Q38 $O^{\varepsilon 1}$)...(Y243 $O^{\eta}H^{\eta}$)	
Ionic bonds [c]	(E9 O^{ε})...(K153 N^{ζ})	(K10 N^{ζ})...(E159 O^{ε})
		(R15 $N^{\eta 1,2}H^{\eta 1,2}$)...(E170 O^{ε})
	(E21 O^{ε})...(R177 $N^{\eta 1,2}H^{\eta 1,2}$)	
	(E23 O^{ε})...(R177 $N^{\eta 1,2}H^{\eta 1,2}$)	(D24 O^{δ})...(K183 N^{ζ})
		(R36 $N^{\eta 1,2}H^{\eta 1,2}$)...(E156 O^{ε})

[a] The listed hydrophobic interactions were determined by a 5-Å distance cut-off between the hydrophobic-group carbon atoms of two residues in more than half of the total number of structures.
[b] The listed hydrogen bond interactions were determined by a 2.5-Å distance cut-off between the hydrogen donor and receptor atoms in more than half of the total number of structures.
[c] The listed ionic bond interactions were determined by a 5-Å distance cut-off between the charged group nitrogen/oxygen atoms of two residues in more than half of the total number of structures.