Supplementary data

Table S1. Properties of DNA mimic proteins

DNA mimic protein	PDB number	No. of NCRs[a]	No. of NCRs on surface[b]	No. of residues on surface[b]	Percentage of NCRs on surface	Protein size (total residues)
DinI	1GHH	13	12	65	18%	81
HI1450	1NNV	31	30	78	38%	107
Ocr	1S7Z	29	28	75	37%	106
TAF Ⅱ 230	1TBA	15	15	58	26%	67
UGI	1UGI	18	17	61	28%	83
P53	2B3G	11	10	78	13%	117
MfpA	2BM4	25	22	111	20%	180
CarS	2KSS	16	15	65	23%	86
Gam	2UUZ	20	19	80	24%	89
ArdA	2W82	37	35	112	31%	163
ICP11	2ZUG	15	14	61	23%	80
DMP19	3VJZ	27	20	98	20%	164

[a] NCR: negatively charged residue

[b] A residue is regarded as an exposed residue on the protein surface if its relative solvent accessibility is greater than 16% (47).

Table S2. Properties of potential DNA mimic proteins

Query protein	Identified protein	Gene name	Description	Z-Score	RMSD (Å)	No. of aligned residues	Sequence identity (%)	No. of NCRs[a]	No. of NCRs on surface[b]	No. of residues on surface[b]	Percentage of NCRs on surface	Protein size (total residues)
DinI	1OTG	hpcD	5-Carboxymethyl-2-Hydroxymuconate Isomerase	5.8	2.9	72	13	15	15	93	16%	125
DinI	2OP8	ywhB	Probable Tautomerase ywhB	4.1	1.9	44	7	13	13	55	24%	61
DinI	3E6Q	PA1966	Putative 5-Carboxymethyl-2-Hydroxymuconate Isomerase	5.6	2.8	72	10	17	16	95	17%	127
DinI	3MF7	cis-caaD	Cis-3-Chloroacrylic Acid Dehalogenase	5.5	2.6	72	8	12	11	87	13%	118
DMP19	2J9W	vps28	vps28-Prov Protein	5.1	3.4	87	6	17	16	67	24%	99
Gam	1A7W	hmfB	Histone hmfB	5.1	1.7	46	9	12	10	60	17%	68
Gam	1II8	rad50	rad50 ABC-Atpase	4.9	2.4	57	14	33	31	132	23%	174
Gam	2PEO	rbcX	rbcX Protein	4.1	3.4	57	7	15	15	96	16%	113
Gam	3B0B	CENPX	Centromere Protein S	5.4	2	50	8	13	13	67	19%	75
MfpA	2F3L	RFR-Domain	RFR-Domain	16.2	1.9	114	16	21	19	81	23%	132
MfpA	2O6W	Rfr23	Repeat five residue (Rfr) protein or pentapeptide repeat protein	14.2	1.3	114	18	20	19	99	19%	147
MfpA	3N90	T13E15.7	Thylakoid lumenal 15 kda protein 1, chloroplastic	11.8	2	90	14	16	14	89	16%	145
UGI	1XS0	ivy	Inhibitor of vertebrate lysozyme	4.6	3	66	6	13	13	89	15%	129
UGI	2KCD	SSP0047	Uncharacterized protein SSP0047	5.8	2.9	71	13	17	17	88	19%	120

[a] NCR: negatively charged residue

[b] A residue is regarded as an exposed residue on the protein surface if its relative solvent accessibility is greater than 16% (47).

Table S3. Protein identification by LC-nanoESI-MS/MS.

Band	Protein name (species)	Acc. No.	MS/mps*	Sequence coverage
1	Preprotein translocase subunit SecA [*Staphylococcus aureus* subsp. *aureus* MRSA252]	YP_040234	4303/95	68%
	DNA topoisomerase IV A subunit (GyrA/ParC) [*Staphylococcus aureus*]	AAA73952	319/6	8%
2	Catabolite control protein A (CcpA) [*Staphylococcus aureus* subsp. *aureus* 55/2053]	ZP_05602273	2594/60	87%
3	Acetyl-CoA carboxylase subunit beta [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372225	1750/30	69%
	50S ribosomal protein L2 [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372771	1337/32	68%
	Phosphosugar-binding transcriptional regulator (RpiR family) [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372839	943/13	48%
4	50S ribosomal protein L3 [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372774	1514/39	75%
	Accessory gene regulator A (agrA) [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372563	1213/23	68%
	Two-component system response regulator (WalR/VicR) [*Staphylococcus aureus*]	CAB65399	918/19	60%
	GTP-sensing transcriptional pleiotropic repressor (CodY) [*Staphylococcus aureus* A9765]	ZP_06328220	764/12	53%
5	30S ribosomal protein S3 [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_372768	1708/78	76%
	DNA damage-inducible repressor (LexA) [*Staphylococcus aureus* subsp. *aureus* Mu50]	NP_371863	762/16	53%
6	Ribosomal protein S4 [*Staphylococcus aureus* subsp. *aureus* 55/2053]	YP_041184	1218/23	63%
	Uracil-DNA glycosylase [*Staphylococcus hominis* SK119]	ZP_04059821	168/3	25%

* MS/mps: Mowse Score/matched peptides.

Table S4. Hydrogen bonding interactions between SAUDG and SAUGI.

SAUDG	**SAUGI**	**Distance (Å) between the two residues**
Pro64 O	Lys22 NZ	6.0 (2.8, 3.2)
Lys79 N	Glu24 OE2	5.6 (2.9, 2.7)
His62 N	Cys25 O	5.4 (2.7, 2.7)
His62 NE2	Ser27 OG	2.8
Ser83 N	Glu26 OE1	3.1
Tyr61 OH	Glu26 OE1	5.1 (2.8, 2.3)
Ser83 OG	Glu26 OE2	2.7
Ser182 OG	Glu26 OE2	5.1 (2.6, 2.5)
Gln58 OE1	Glu29 N	2.7
Gln58 NE2	Glu29 O	2.8
Arg130 NH1	Glu29 OE1	3.0
Lys159 N	Glu30 OE1	2.8
His180 N	Glu30 OE2	3.0
Arg188 NH1	Asp33 O	2.7
Arg188 NH2	Leu36 O	2.8
Ser182 OG	Thr55 OG1	5.9 (2.6, 3.3)
Gln66 NE2	Tyr57 OH	2.6
Lys79 O	Tyr57 OH	5.4 (2.7, 2.7)
Lys79 NZ	Asp59 OD2	2.8
Asn86 OD1	Tyr67 OH	5.4 (2.9, 2.5)

A single distance indicates a direct hydrogen bond. Additional distances in parentheses indicate that the bond is formed by two hydrogen bonds and mediated by a water molecule.

Supplementary Figure Legends

Figure S1. Potential DNA mimic proteins similar to (A) UGI, (B) DMP19, (C) DinI, (D) Gam, and (E) MfpA. The query and selected candidate DNA mimic proteins are colored pale yellow and cyan, respectively. Carboxyl groups of aspartic acid and glutamic acid are represented by red spheres.

Figure S2. Comparisons of SSP0047 (2KCD) and its homologs with UGI. (A) Protein structures and charge distribution (B) Amino acid sequence alignment.

Figure S3. A comparison of the distances between the negative spots on SSP0047 (2KCD) and B-form DNA. The distances between the carboxyl groups on SSP0047 and the phosphate groups on B-form DNA were measured by the program PyMOL.

Figure S4. A His-pulldown assay to find the interacting partners of SAUGI. When total protein extracted from *Staphylococcus aureus* was used as prey, the N-terminal His_{10}-tagged SAUGI bait pulled down several proteins. LC-nano ESI-MS/MS identified eight DNA binding proteins from these six protein bands (see Table S3 for details).

Figure S5. Analytical ultracentrifugation and gel filtration analysis to determine the molecular weights of SAUGI, SAUDG and the SAUGI/SAUDG complex. (A) The molecular weights of SAUGI, C-terminal His_6-tagged SAUDG and SAUGI/C-terminal His_6-tagged SAUDG complex were estimated by Superdex 75 10/300 gel filtration. The standard plot of Kav against log MW (inset) was generated by the elution volumes of the proteins ovalbumin (O; 43 kDa), carbonic anhydrase (CA; 29 kDa), RNase A (RA; 13.7 kDa) and Aprotinin (A; 6.5 kDa). The Kav of each protein (SAUGI, C-terminal His_6-tagged SAUDG and SAUGI/C-terminal His_6-tagged SAUDG complex) was then used to find its molecular weight. (B) For sedimentation velocity (SV) analysis using analytical ultracentrifugation, all samples were diluted to a suitable concentration (OD 280 absorption between 0.1~0.8; SAUGI: 0.2 mg/ml, SAUDG: 0.3 mg/ml and SAUGI/SAUDG complex: 0.3 mg/ml) using 30 mM Tris-HCl pH 7.4, 5% glycerol, 100 mM NaCl, 1 mM EDTA and 5 mM DTT. All analytical ultracentrifugations were performed as described in our previous studies (24, 48). The observed MWs were calculated by the program SEDFIT (http://www.analyticalultracentrifugation.com) using the experimental S value from the SV data. The S value of each protein peak was SAUGI: 1.792, SAUDG: 2.676 and SAUGI/SAUDG complex: 3.313. A full comparison of the theoretical and observed MW values is given in Table 2.

Figure S6. (A) Comparison of the crystal structures of unbound UDGs from four different species and (B) the proposed DNA binding region of SAUDG as suggested by alignment with DNA-bound human UDG. The RMSD fit is 0.726 Å for SAUDG and unbound human UDG (PDB ID: 1AKZ), 0.791 Å for SAUDG and *E.coli* UDG (PDB ID: 1LQJ), 0.484 Å for SAUDG and *B. subtilis* UDG (PDB ID: 3ZOR) and 1.493 Å for SAUDG and DNA-bound human UDG (PDB ID: 1SSP). In (A), the first circled region highlights the protruding residues Leu182 of SAUDG, Leu191 of *E. coli* UDG, Phe191 of *B. subtilis* UDG and Leu272 of human UDG. The circled regions in the right figure show the three loops (4-Pro, Gly-Ser and Leu272) that accomplish compression of the phosphates in the human UDG/DNA complex. These loops were conserved in these UDG structures.

Figure S7. The secondary structural elements and the amino acid sequences of SAUGI (left) and SAUDG (right). The α-helices and β-sheets are respectively represented by cylinders and arrows, and colored red and yellow in SAUDG and cyan and magenta in SAUGI.

Figure S8. 2FoFc maps of the three SAUGI-interacting regions of SAUDG shown

in Figure 3B to D. The left stereograms (sigma level = 1.5 σ) show SAUDG bound in the SAUGI/SAUDG complex; the stereograms on the right (sigma level = 1.0 σ) show the same regions in unbound SAUDG.

Figure S9. A comparison of the SAUDG-bound SAUGI and unbound SSP0047 from *Staphylococcus saprophyticus* (2KCD). The two circles in the merged figure (center) highlight the two loop regions that show structural differences.

Figure S10. Binding affinities between SAUGI and UGI to SAUDG and human UDG were determined by surface plasmon resonance. Increasing concentrations of both uracil-DNA glycosylase inhibitors were injected to immobilized SAUDG or human UDG, and the sensograms were recorded on a BIAcore T200. The colored lines show experimentally recorded values at different concentrations, and the black lines are a fit of the data to a 1:1 Langmuir model.

Figure S11. Conformational changes in UDGs upon binding to their respective uracil-DNA glycosylase inhibitors. The RMSD fit is 0.345 Å for unbound SAUDG and SAUGI-bound SAUDG, 0.479 Å for unbound *E.coli* UDG (PDB ID: 1LQJ) and UGI-bound *E.coli* UDG (PDB ID: 1LQM), 0.217 Å for unbound *B. subtilis* UDG

(PDB ID: 3ZOR) and p56-bound *B. subtilis* UDG (PDB ID: 1ZOQ). The most significant changes between the bound and unbound UDGs are shown in detail.

Figure S12. A BS3 cross-linking assay suggests an interaction between SAUGI and C-terminal His-tagged *S. aureus* LEX transcription factor (SALEX). In lane 3, the presence of a new band (asterisk) after the addition of 1 mM BS3 suggests cross-linking occurs between SAUGI and SALEX.

Supporting figures

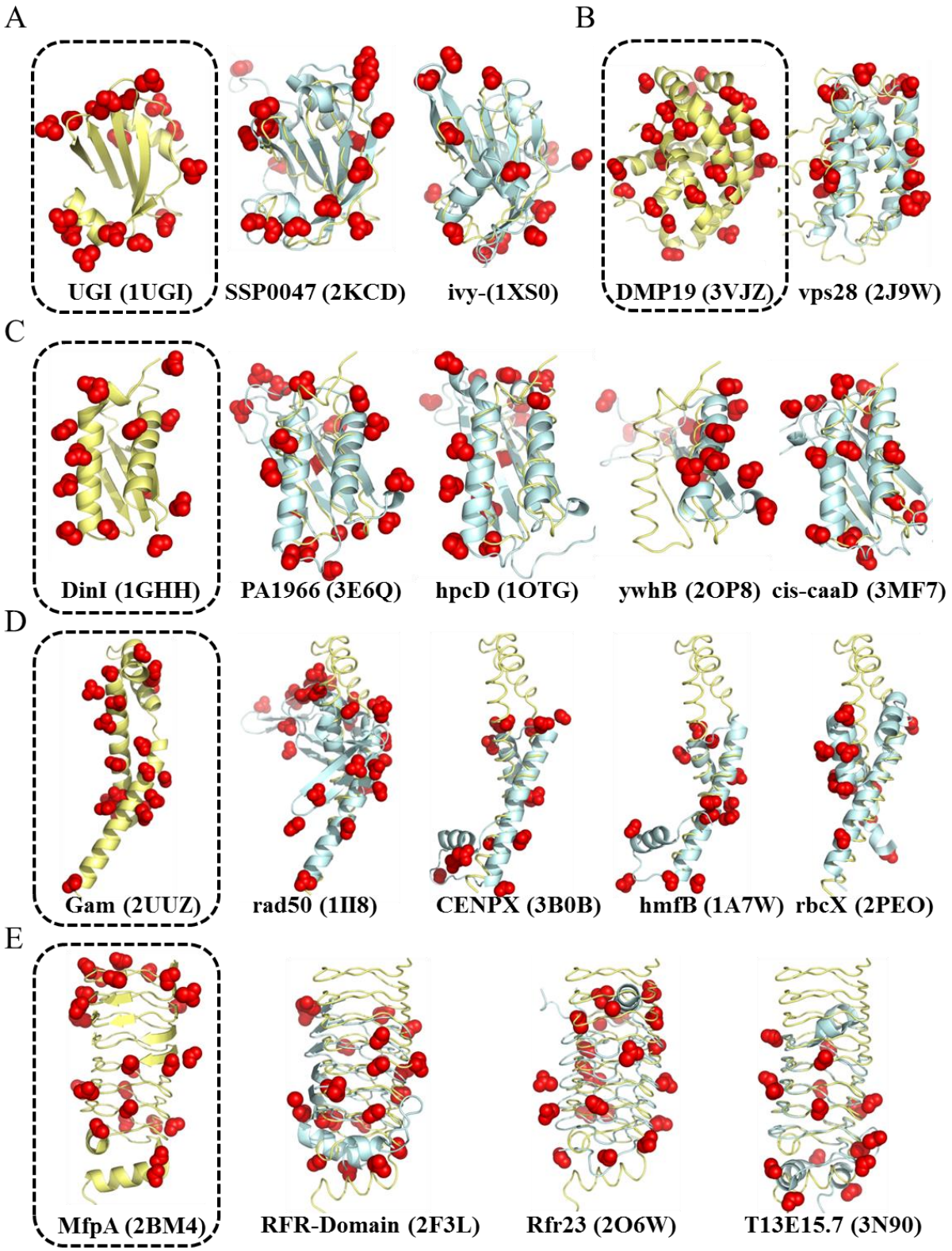

Figure S1

Figure S2 (A)

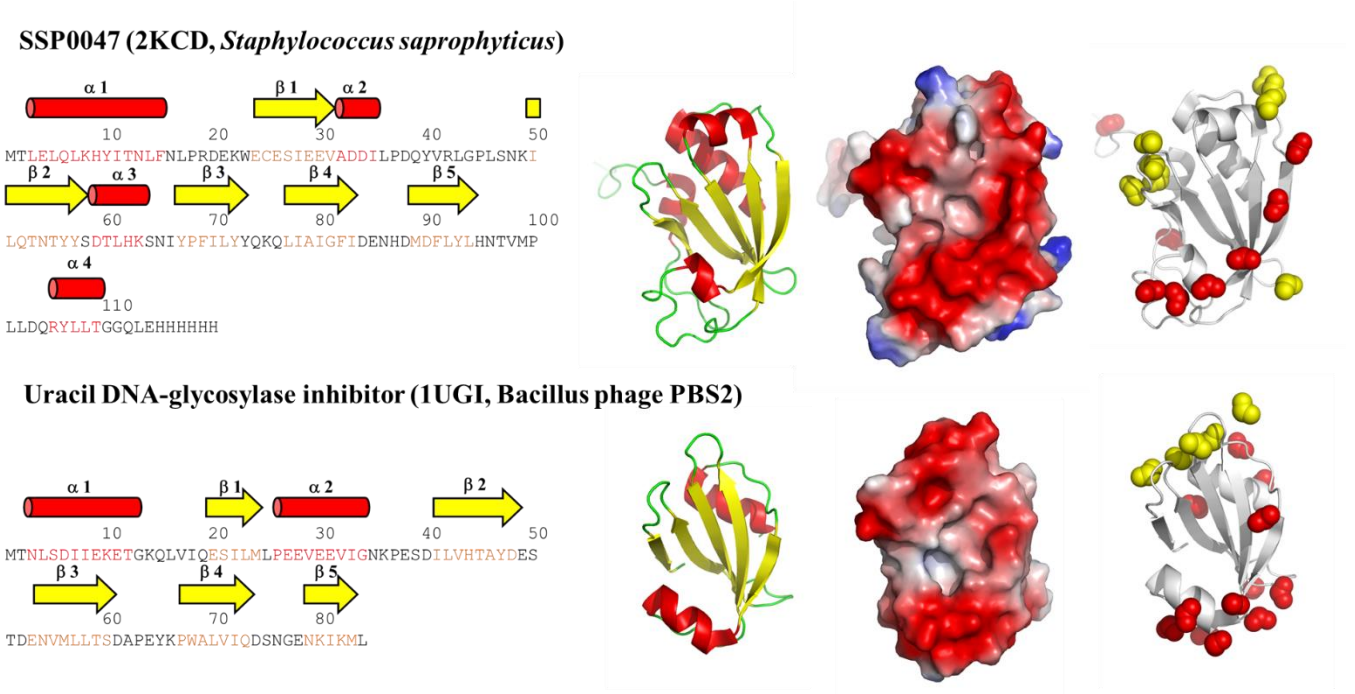

Figure S2 (B)

```
                                                  10        20        30        40        50
                                                  |         |         |         |         |
SSP0047 homologs [Staphylococcus equorum]         1 MTLEQQLKHYITNLFNLPKDEKWECESIEEVADDILPEQYVRLGPLSSKI  50
SSP0047 homologs [Staphylococcus epidermidis]     1 MTLEQQLKHYITNLFNLPKDEKWECESIEEAADDILPEQHVRLGPLSNKI  50
SSP0047 [Staphylococcus saprophyticus ]           1 MTLELQLKHYITNLFNLPRDEKWECESIEEVADDILPDQYVRLGPLSNKI  50
SSP0047 homologs [Staphylococcus aureus]          1 MTLELQLKHYITNLFNLPKDEKWECESIEEIADDILPDQYVRLGALS-NK  49
Uracil-DNA glycosylase inhibitor [Bacillus phage PBS2] 1 ----------MTNLSDIIEKETGKQLVIQESILMLPEEVEEVIGNKPESD  40
Consensus                                                  :*** :: ..*. :   *:*    :  :     :*  . .

                                                  60        70        80        90        100
                                                  |         |         |         |         |
SSP0047 homologs [Staphylococcus equorum]        51 LQTNTFYSDTLHKSNIYPFILYYQKQLIAIGYIDENNDMDFLYLHNTIAP 100
SSP0047 homologs [Staphylococcus epidermidis]    51 LQTNTYYSYTLHKSNIYPFILYYQKQLIAIGYIDENHDKDFLYLHNTIMP 100
SSP0047 [Staphylococcus saprophyticus ]          51 LQTNTYYSDTLHKSNIYPFILYYQKQLIAIGFIDENHDMDFLYLHNTVMP 100
SSP0047 homologs [Staphylococcus aureus]         50 LQTYTYYSDTLHESNIYPFILYYQKQLIAIGYIDENHDMDFLYLHNTIM- 98
Uracil-DNA glycosylase inhibitor [Bacillus phage PBS2] 41 ILVHTAYDESTDENVMLLTSDAPEYKPWALVIQDSNGENKIKML------ 84
Consensus                                         16 :  .  *  *.  :  .:.  :        :  :   *:    *.*  :  .:   *        30

                                                  110
                                                  |
SSP0047 homologs [Staphylococcus equorum]       101 LLDQRYLLTGGQ 112
SSP0047 homologs [Staphylococcus epidermidis]   101 LLDQRYLLTGGQ 112
SSP0047 [Staphylococcus saprophyticus ]         101 LLDQRYLLTGGQ 112
SSP0047 homologs [Staphylococcus aureus]         99 LLDQRYLLTGGQ 110
Uracil-DNA glycosylase inhibitor [Bacillus phage PBS2] 84 ------------ 84
Consensus
```

Figure S2

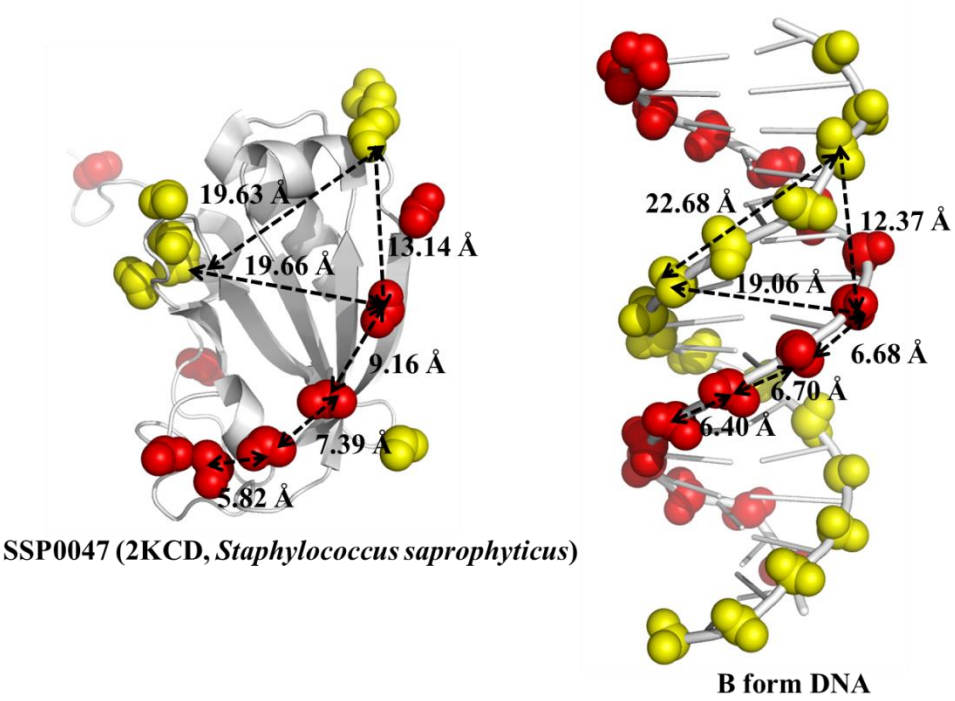
19.63 Å
13.14 Å
19.66 Å
9.16 Å
7.39 Å
5.82 Å
SSP0047 (2KCD, Staphylococcus saprophyticus)
22.68 Å
12.37 Å
19.06 Å
6.68 Å
6.70 Å
6.40 Å
B form DNA

Figure S3

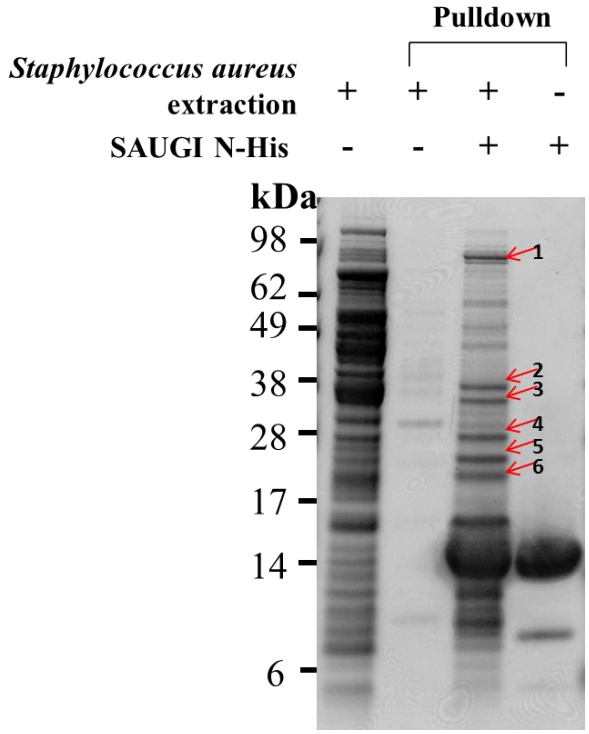
Pulldown
Staphylococcus aureus extraction
+ + + -
SAUGI N-His
- - + +
kDa
98
62
49
38
28
17
14
6
1
2
3
4
5
6

Figure S4

Figure S5 (A)

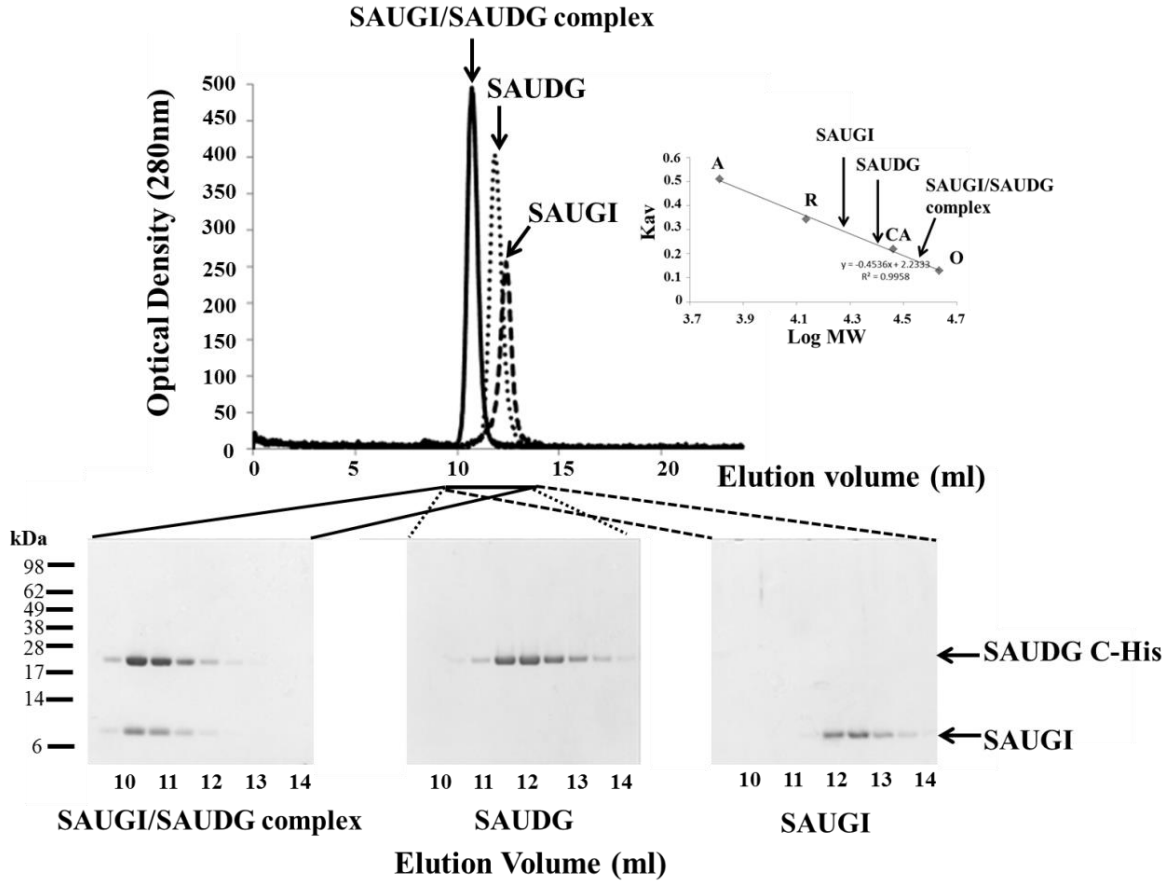

Figure S5 (B)

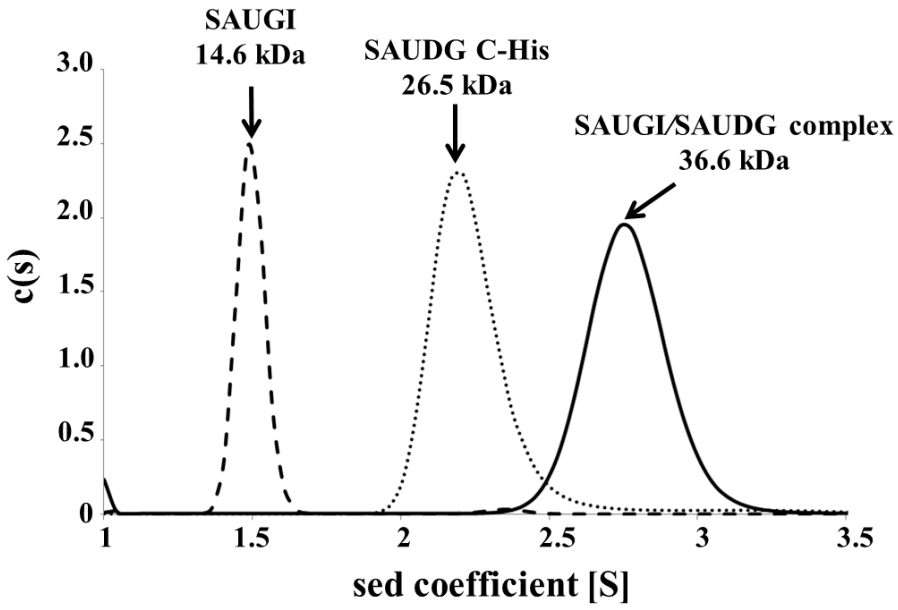

Figure S5

Figure S6 (A)

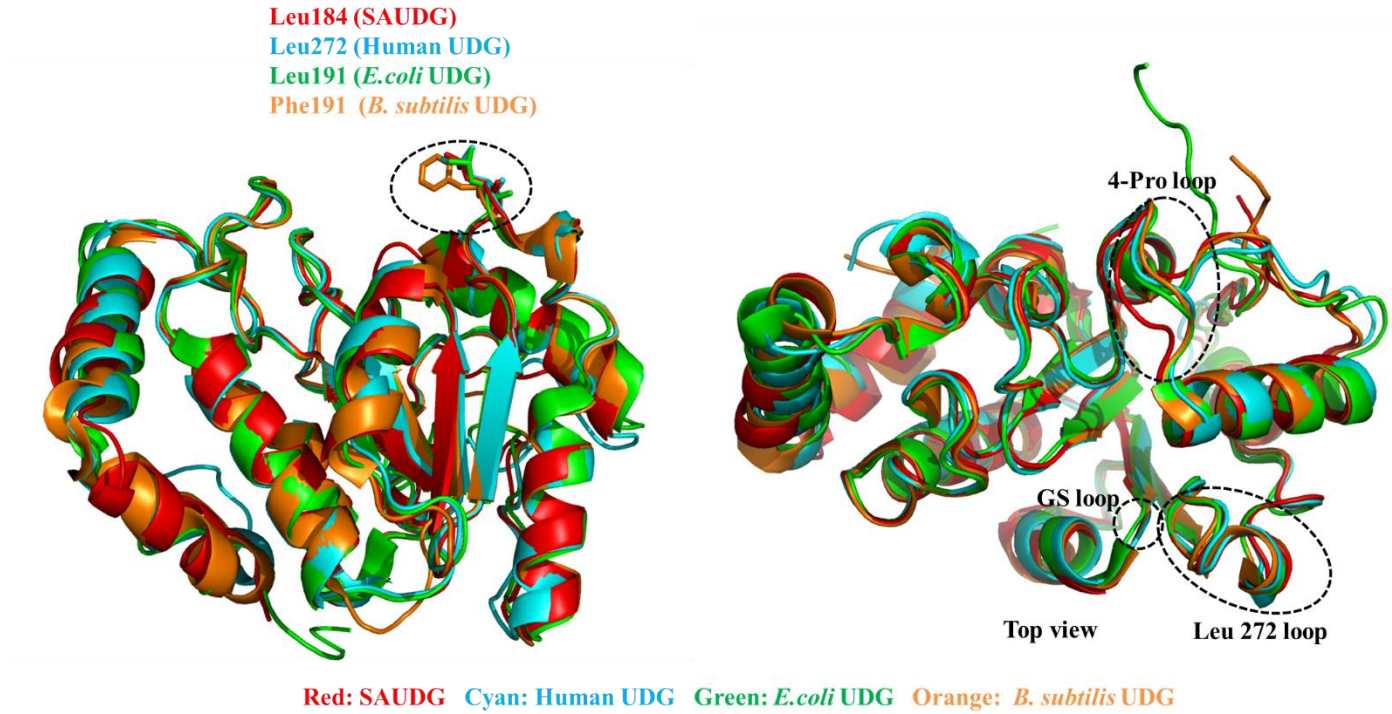

Figure S6 (B)

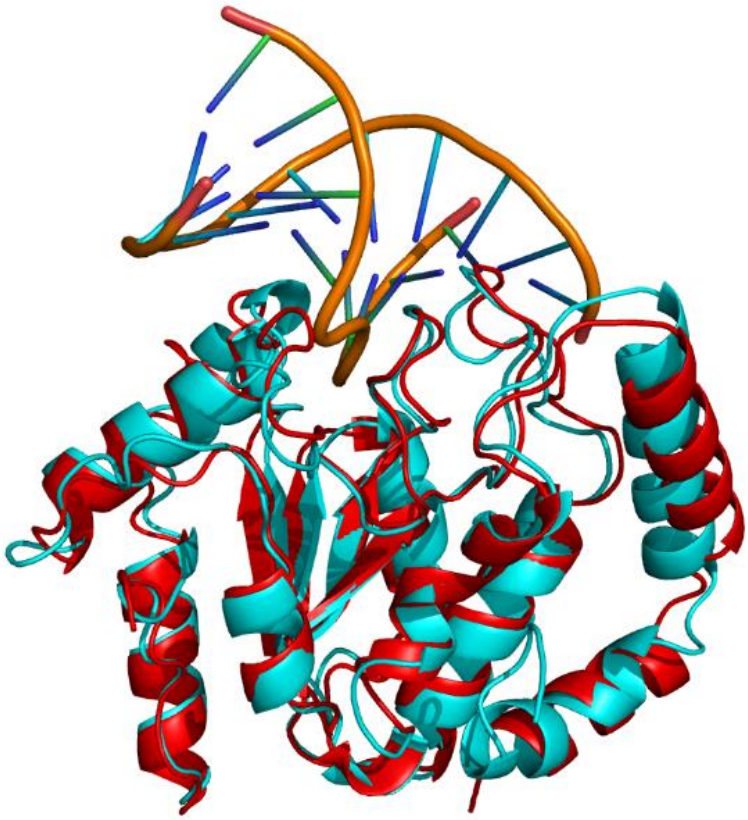

Figure S6

SAUGI
α 1
β 1
α 2
α 3
β 2
α 4
β 3
β 4
β 5
β 6
α 5
10
20
30
40
50
60
70
80
90
100
110
MTLELQLKHYITNLFNLPKDEKWECESIEEIADDILPDQYVRLGALSNKI
LQTYTYYSDTLHESNIYPFILYYQKQLIAIGYIDENHDMDFLYLHNTIMP
LLDQRYLLTGGQ
SAUDG
α 1
α 2
α 3
α 4
β 1
α 5
α 6
β 2
α 7
β 3
α 8
β 4
α 9
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
MEWSQIFHDITTKHDFKAMHDFLEKEYSTAIVYPDRENIYQAFDLTPFEN
IKWVILGQDPYHGPNQAHGLAFSVQPNAKFPPSLRNMYKELADDIGCVRQ
TPHLQDWAREGVLLLNTVLTVRQGEANSHRDIGWETFTDEIIKAVSDYKE
HVVFILWGKPAQQKIKLIDTSKHCLIKSVHPSPLSAYRGFFGSKPYSKAN
TYLESVGKSPINWCESEALEHHHHHH

Figure S7

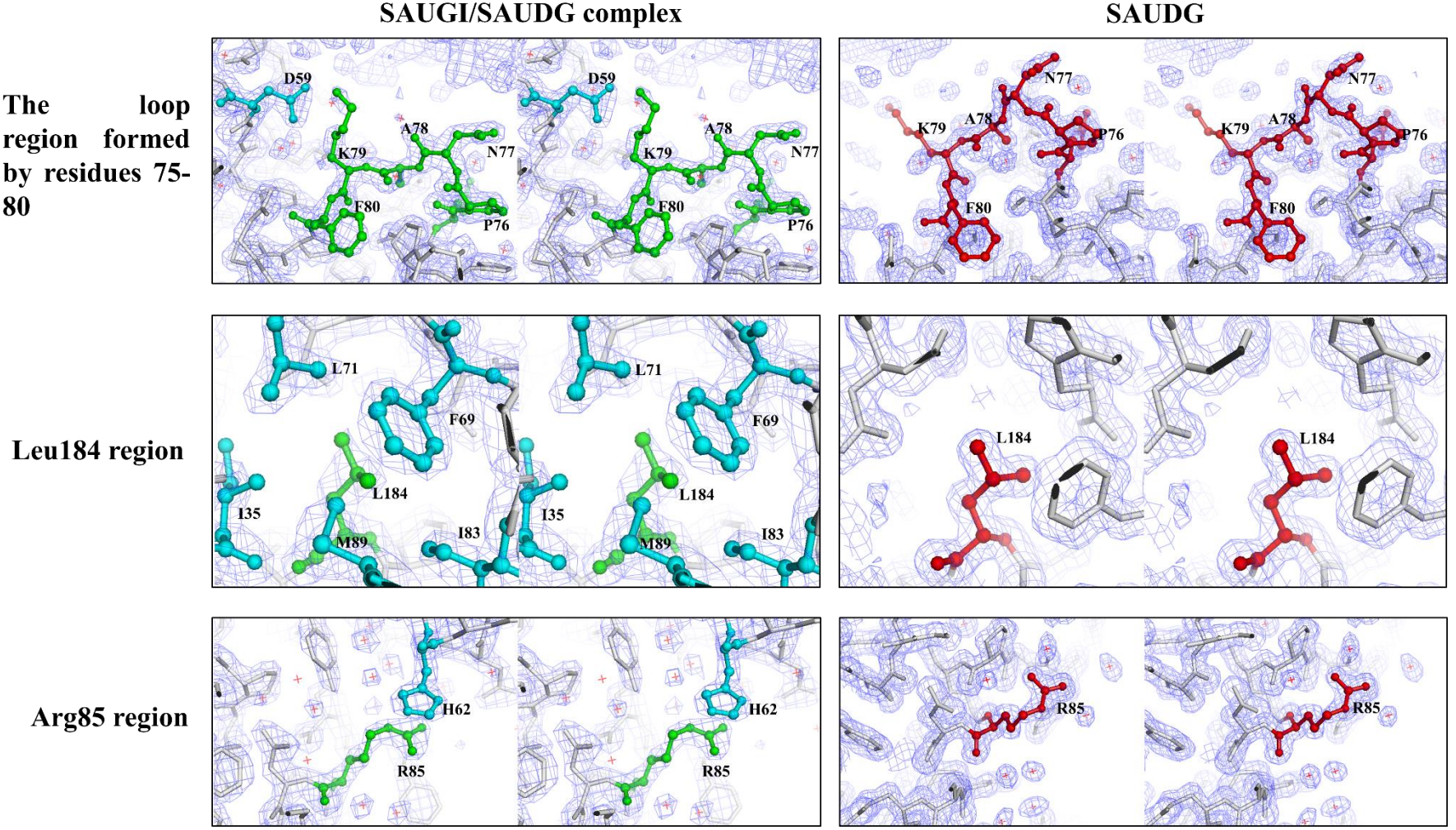
SAUGI/SAUDG complex
SAUDG
The loop region formed by residues 75-80
Leu184 region
Arg85 region
D59
A78
K79
N77
F80
P76
L71
F69
L184
I35
M89
I83
H62
R85

Figure S8

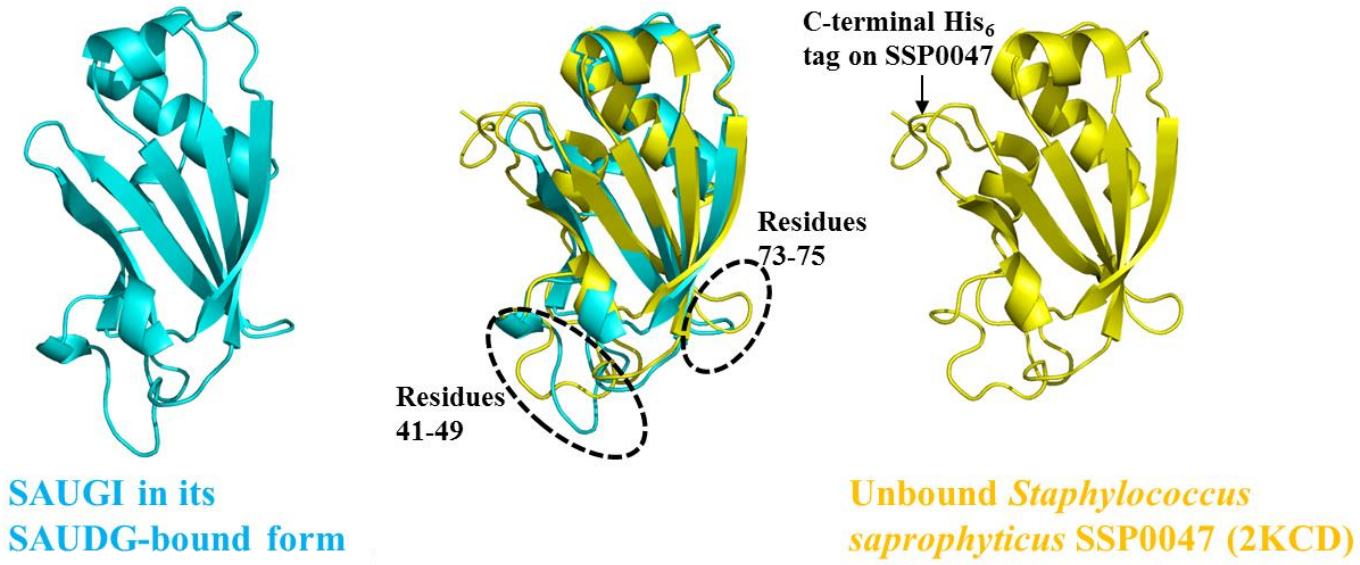
C-terminal His6
tag on SSP0047
Residues
73-75
Residues
41-49
SAUGI in its
SAUDG-bound form
Unbound Staphylococcus
saprophyticus SSP0047 (2KCD)

Figure S9

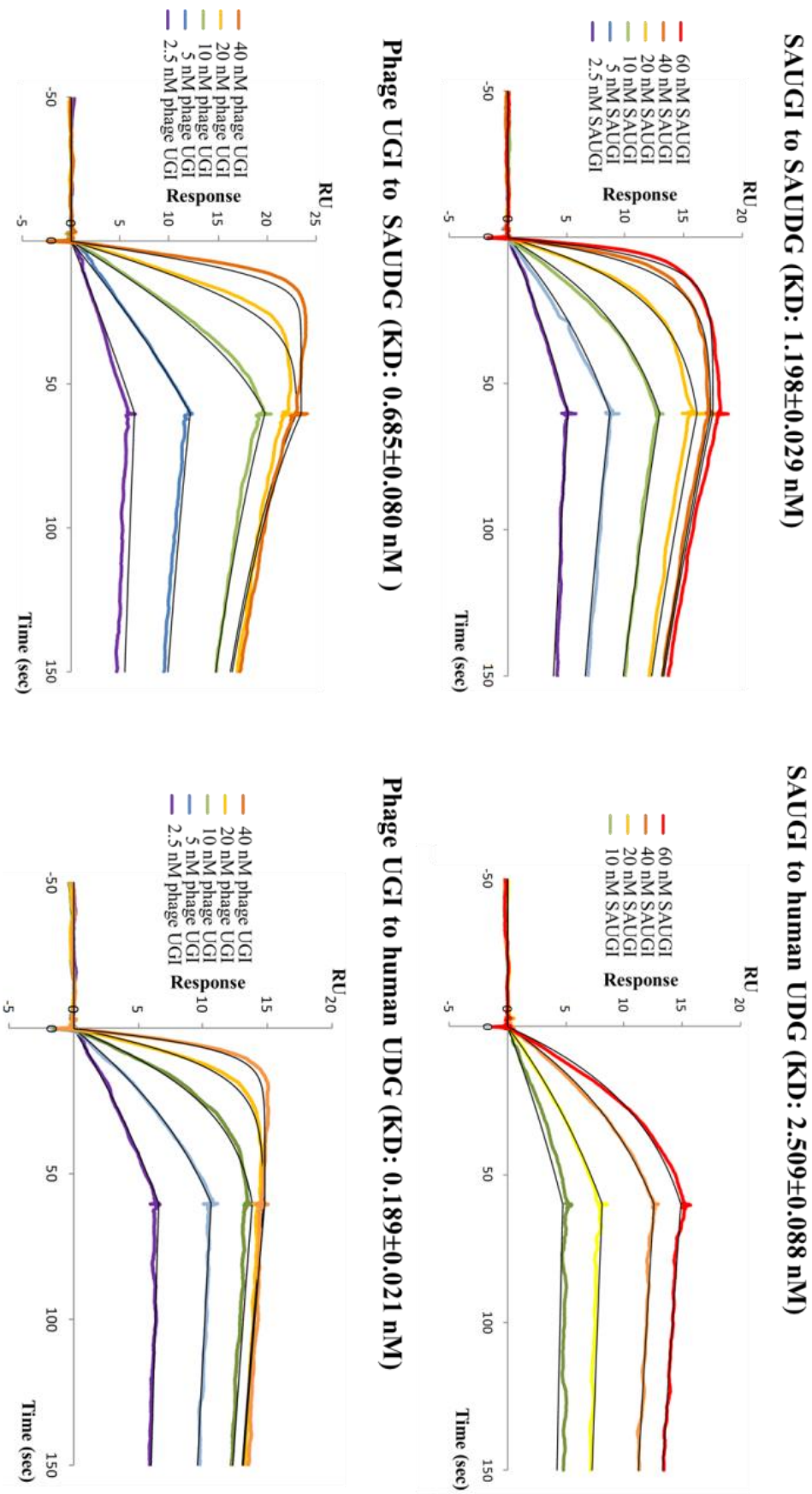
SAUGI to SAUDG (KD: 1.198±0.029 nM)
60 nM SAUGI
40 nM SAUGI
20 nM SAUGI
10 nM SAUGI
5 nM SAUGI
2.5 nM SAUGI
Response
RU
Time (sec)
Phage UGI to SAUDG (KD: 0.685±0.080 nM)
40 nM phage UGI
20 nM phage UGI
10 nM phage UGI
5 nM phage UGI
2.5 nM phage UGI
Response
RU
Time (sec)
SAUGI to human UDG (KD: 2.509±0.088 nM)
60 nM SAUGI
40 nM SAUGI
20 nM SAUGI
10 nM SAUGI
Response
RU
Time (sec)
Phage UGI to human UDG (KD: 0.189±0.021 nM)
40 nM phage UGI
20 nM phage UGI
10 nM phage UGI
5 nM phage UGI
2.5 nM phage UGI
Response
RU
Time (sec)

Figure S10

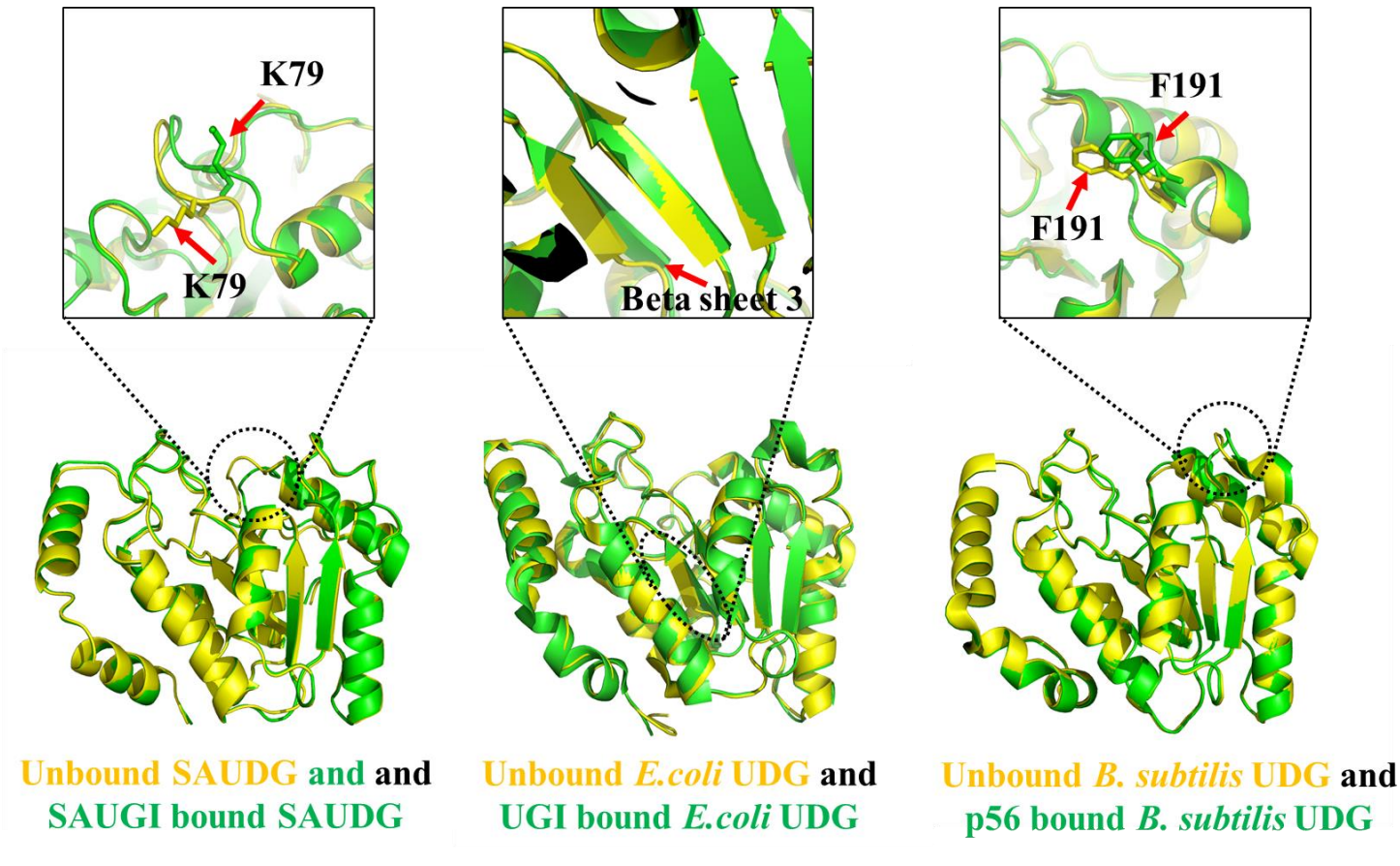
K79
K79
Beta sheet 3
F191
F191
Unbound SAUDG and and
SAUGI bound SAUDG
Unbound E.coli UDG and
UGI bound E.coli UDG
Unbound B. subtilis UDG and
p56 bound B. subtilis UDG

Figure S11

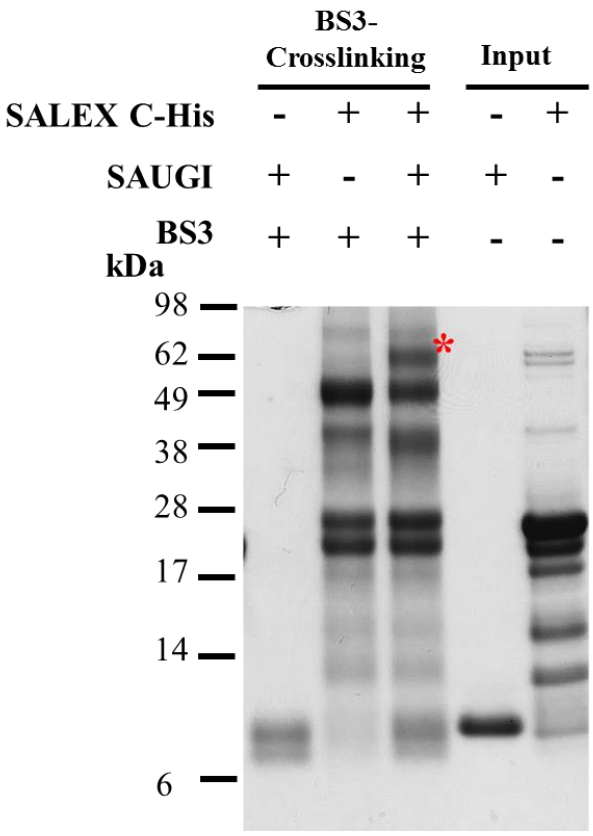
BS3-
Crosslinking
Input
SALEX C-His - + + - +
SAUGI + - + + -
BS3 + + + - -
kDa
98
62
49
38
28
17
14
6
*

Figure S12