Supplemental Information.

A variable DNA recognition site organization establishes the LiaR mediated cell envelope stress response of enterococci to daptomycin.

Authors: Milya Davlieva, Yiwen Shi, Paul G. Leonard, Troy A. Johnson, Michael R. Zianni, Cesar A. Arias, John E. Ladbury and Yousif Shamoo.

Table S1. Data collection and refinement statistics

Data set	$LiaR^{DBD}$	$LiaR^{(DBD)D191N}$	$LiaR^{DBD}$ /consensus *liaXYZ*	$LiaR^{(DBD)D191N}$ /consensus *liaXYZ*	$LiaR^{(DBD)D191N}$ /secondary *liaXYZ*
Data collection					
Wavelength (Å)	0.97872	0.97856	0.9788	0.9788	0.9788
Resolution (Å)[a]	30.0-1.78 (1.81-1.78)	30.0-1.48 (1.51-1.48)	30.0-2.3 (2.34-2.28)	30.0-2.3 (2.34-2.30)	30.0-2.5 (2.54-2.5)
Space group	$P22_12_1$	$P22_12_1$	$P2_12_12_1$	$P2_12_12_1$	$P6_2$
Cell dimensions a,b,c; (Å) α, β, ɣ; (°)	31.23, 77.02, 77.19 90, 90, 90	31.05, 76.93, 76.92 90, 90, 90	38.3, 77.9, 104.7 90, 90, 90	33.34, 77.25, 104.99 90, 90, 90	113.45, 113.45 48.11 90, 90, 120
Molecules per a.u.	2	2	2	2	2
Reflections Total Unique[a]	208834 18096	420527 30660	206189 14672	197972 14428	280738 12455
Average redundancy[a]	11.5 (11.5)	13.7 (9.7)	14.1 (14.8)	13.7 (11.7)	22.5 (22.3)
Completeness (%)[a]	97.7 (96.3)	96.9 (92.6)	100 (100)	99.9 (99.2)	100 (100)
R_{merge} (%)[a,b]	7.3 (34.3)	6.5 (66.9)	10.7 (43.1)	9.7 (90.4)	10 (73.1)
Output <I/sigI>[a]	31.64 (7.5)	34.8 (1.88)	25.0 (7.2)	26 (2.66)	33.7 (9.7)
Refinement					
R_{work} (%)[c]	16.39	18.20	21.5	20.84	21.85
R_{free} (%)[d]	20.81	24.64	25.9	26.2	27.55
r.m.s.d.[e] from ideality Bonds (Å) Angles (°)	0.007 0.91	0.008 0.93	0.004 0.64	0.003 0.60	0.004 0.68
Twin law Twin fraction	-h,l,k 0.49	-h,l,k 0.49			
Wilson B-factor (Å^2)	18.5	20.7	27.4	28.0	33.3
Ramachandran[f] Favored (%) Allowed (%) Outliers (%)	96.88 3.12 0	96.09 3.90 0	99.24 0.76 0	98.5 1.5 0	100 0 0
PDB ID	4WSZ	4WTO	4WUH	4WU4	4WUL

[a] Values for the last shell are in parenthesis

[b] $R_{merge} = \Sigma \left| I - \langle I \rangle \right| / \Sigma I$, where I is measured intensity for reflections with indices of hkl

[c] $R_{work} = \Sigma \left| F_o - F_c \right| / \Sigma \left| F_o \right|$ for all data with $F_o > 2\ \sigma\ (F_o)$ excluding data to calculate R_{free}

[d] $R_{free} = \Sigma \left| F_o - F_c \right| / \Sigma \left| F_o \right|$ for all data with $F_o > 2\ \sigma\ (F_o)$ excluded from refinement.

[e] Root mean square deviation

[f] Calculated by using MolProbity (1)

Supplemental Experimental Procedures

Expression and purification of *E. faecalis* response regulator S613 LiaR and variants

The gene encoding LiaR S613 was cloned between the NcoI/HindIII sites of the modified pET-Duet vector (Novagen, NJ, USA) containing an N-terminal His tagged SUMO peptide sequence and expressed in *Escherichia coli* BL21 (DE3) cells. DNA sequence information for the adaptive mutant $LiaR^{D191N}$ was obtained from *E. faecalis* Turbidostat-derived DAP resistant (TDR4) strain (MIC=0.5 µg/ml) (Miller et al, 2013). Constructs encoding LiaR and variants were confirmed by Sanger sequencing (SeqWright). Biophysical and structural analysis of *E. faecalis* LiaR and its mutants (full length and DNA binding domain) including the adaptive variant $LiaR^{D191N}$ have been hindered by low expression and poor solubility. Expression of LiaR and LiaR variants in LB with varying temperatures proved unsatisfactory results. Using EnPresso B (Biosilta, Oulu, Finland) provided significant improvement in yield of soluble and folded protein. Overproduction of soluble $LiaR^{WT}$, $LiaR^{D50E}$ and $LiaR^{D50A}$ routinely produced about 1 mg/per 100 ml of cell culture for the full length of protein and about 2 mg/per 100 ml of cell culture for the DNA binding domain of LiaR. Adaptive mutant $LiaR^{D191N}$ could only be isolated at the scale of ~300 µg/per 100 ml of induced cell culture due to a smaller amount of total expressed protein. Yield of the $LiaR^{D50E/D191N}$ was extremely low (about 50 µg/per 100 ml of induced cell culture). Gel filtration chromatography (Superedex-200) suggested significant aggregation of $LiaR^{D50E/D191N}$ even at moderate ionic strength (300 mM NaCl) and glycerol (20%v/v) where other LiaR variants were well behaved.

For purification, the frozen cell pellet was re-suspended in Buffer A (50 mM Tris pH 7.4, 0.5 M NaCl, 20 mM Imidazole, 0.3 mM DDT, 0.2 mM PMSF, 20% (v/v) glycerol, 0.05 % (v/v) Tween-20) and Complete protease inhibitor cocktail tablet, EDTA-free (Roche Diagnostics Corp, Indianapolis, IN, US). Lysis was performed using Branson Sonifier 250 (VWR Scientific). The lysate was then centrifuged for 60 min at 24,000 rpm at 4°C. The supernatant was loaded onto a Hi Trap affinity (Ni2+) column (GE Healthcare Life Sciences). The column was washed with 10 column volumes of Buffer A, and eluted with a step elution gradient from 20 to 500 mM Imidazole (pH 7.5). The fractions containing the protein of interest were pooled, dialyzed against 50 mM Tris pH 7.5, 0.5 M NaCl, 0.3 mM DTT, 20 % (v/v) glycerol, 0.05 % (v/v) Tween-20 overnight at 4 °C. N-terminal SUMO peptide (within the 6xHis tag) was removed by treatment with His-tagged SUMO protease. LiaR without the 6xHis SUMO- tag was purified from the reaction mixture using the same chromatography strategy described above. The fractions containing LiaR were pooled, dialyzed against 50 mM Tris pH 7.5, 0.5 M NaCl, 0.3 mM DTT, 5mM $MgCl_2$, 20 % (v/v) glycerol, 0.05 % (v/v) Tween-20 and purified over a Q-XL Sepharose column (GE Healthcare) using a 0.1-1 M NaCl gradient. The peak fractions were pooled, concentrated and loaded onto a Superdex-200 column (GE Healthcare, HiLoad 16/60) (50 mM Tris pH 7.5, 0.3 M NaCl, 0.3 mM DTT, 10 mM $MgCl_2$, 20 % (v/v) glycerol, 0.05 % (v/v) Tween-20) for the final purification step. The purities of the expressed proteins $LiaR^{WT}$, $LiaR^{D50E}$, $LiaR^{D50A}$, $LiaR^{D191N}$, double mutant $LiaR^{D50E/D191N}$, $LiaR^{DBD}$ and $LiaR^{(DBD)D191N}$ proteins were assessed by SDS–PAGE to be greater than 95%.

Analytical Ultracentrifugation

All samples were prepared in 50 mM Tris-HCl, 300 mM NaCl, 10 mM $MgCl_2$, 0.3 mM DTT, 20% (v/v) glycerol and 0.05% (v/v) Tween-20 at pH 7.5 and loaded into sample chambers with Epon double sector centerpieces and sapphire windows. SEQ scans were recorded using either the absorbance at 280 nm (LiaR, $LiaR^{D50E}$ and $LiaR^{D191N}$) or interference optics ($LiaR^{DBD}$, $LiaR^{(DBD)D191N}$), after 72 h incubation at each rotor speed. The protein partial specific volume and solvent density were calculated using Sednterp 1.09 (2) Rotor speeds for $LiaR^{wt}$ = 14,000 r.p.m. (abs data),$LiaR^{D50E}$ = 14,000 r.p.m. (abs data), $LiaR^{D191N}$ = 11,000, 14000 r.p.m., 19,000 (abs data), DBD LiaR = 22000 r.p.m., 36,000 r.p.m., 42,000 r.p.m. (interference data), DBD $LiaR^{D191N}$ = 36,000 r.p.m. and 42,000 r.p.m. (interference data). A F-statistics error mapping approach was used to determine the 95% confidence intervals for the dissociation constant. (14,000 r.pm. for $LiaR^{WT}$, $LiaR^{D50E}$ and $LiaR^{D191N}$ and 36,000 r.p.m. for $LiaR^{DBD}$ and $LiaR^{(DBD)\ D191N}$).

DNA footprint analysis by automated capillary electrophoresis (DFACE)

The 397-bp fluorescent labeled DNA probe for the S613 *E. faecalis liaFSR* operon and 350-bp fluorescent labeled DNA probe for the S613 *E. faecalis liaXYZ* operon were generated with PCR amplification with the primers 5'-(VIC)-CATCGGTAAAACAGTTACTTTCCA -3' and 5'-(FAM)-GACGATAAAAAAGCGCCAAGGGTT-3' for *liaFSR* operon and 5'-(VIC)-CAATACTTGGAAAGAATTGGCGAC -3' and 5'-(FAM)-TAATTCTAATACGCG TTCTCTTTC-3' *liaXYZ* operon from Applied Biosystems (Grand Island, NY). Full length of $LiaR^{D191N}$ protein at 0.5 µM and 5.0 µM was then incubated with respective fluorescent labeled DNA probe in Binding Buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM $MgCl_2$, and 0.3 mM DTT, 20% (v/v) glycerol, 0.05% (v/v) Tween-20) for 10 min. After this incubation, DNase I (New England Biolabs Inc., Ipswich, MA, 2 U/µl) digestion was performed with diluted DNase I (1/10) for 1 min at room temperature (final volume, 20 µl). The reaction was then stopped by adding EDTA to a final concentration of 5 mM following by incubation at 75°C for 10 min. Control digestions with the probe were performed in the absence of the protein. The DNA fragments were purified with a QIA-quick PCR purification kit (Qiagen, Valencia, CA) and then sent to the Plant-Microbe Genomics Facility for further analysis. Fragments from 2 µl of each sample were separated and detected on a 3730 DNA analyzer (Applied Biosystems; AB) with a 50 cm capillary. Each sample was combined with 9 µl HiDi/0.1µl GS600LIZ size standard (AB) and injected with 3kV for 30s to maximize signal. The resulting electropherograms from the DFACE assay were analyzed with the software GeneMapper 4.0 set to normalize all samples to the sum of the signal of all samples. In addition, the histograms were further normalized by subtracting the height of the equivalent peak (same Bin) in the negative control from the sample, i.e. each bar is the difference between sample and negative control. Data reported utilized fragments/strand labeled with VIC dye.

Determination of LiaR and LiaR variants DNA binding activity by Microscale Thermophoresis

The 5'-ends of the specified oligonucleotides labeled with fluorescein (Flc) were purchased from Sigma-Aldrich. The annealing of the two complementary strands was performed in 50 mM Tris-HCl (pH 7.5), 300 mM NaCl, and 10 mM $MgCl_2$ by heating at 95°C for 10

min and slowly cooling to the room temperature. A solution of unlabeled LiaR (wild type or mutants) was serially diluted in reaction buffer (50 mM Tris pH 7.5, 300 mM NaCl, 10 mM $MgCl_2$, 0.3 mM DTT, 20% (v/v) glycerol, 0.05% (v/v) Tween-20) to which an equal volume of fluorescein labeled DNA was added to a final concentration of 40 nM. The samples were loaded into standard treated capillaries (NanoTemper).
MST Data Fitting: The raw MST traces for each individual experiment were transformed and fit according to published methods (3). Briefly, each raw fluorescence trace is fit to the law of mass action equation $F=(1/2c_A)(c_T+c_A+K_d-sqrt((c_T+c_A+K_d)\text{^}2-(4* c_T*c_A))$, where K_d is the dissociation constant, c_A is the concentration of the fluorescently labeled molecule (titrant), and c_T is the concentration of titrant. Note, when the markers were increased in size for readability the error bars became covered in some cases. Attempts to fit the data with a cooperativity parameter were not satisfactory suggesting that at these protein concentrations the tetrameric oligomeric state was already predominant. This was in good agreement with our AUC data.

Crystallization, data collection and structure determination of the $LiaR^{DBD}$, $LiaR^{(DBD)D191N}$.

Crystals were briefly soaked in mother liquor plus 20% (v/v) glycerol and were flash frozen in a liquid-nitrogen stream. X-ray diffraction data were processed using HKL2000 (4). The crystal belonged to the space group $P22_12_1$ with the unit cell parameters a=31.23, b=77.02, c=77.19, α=β=ɣ=90°. Crystals of DNA binding $LiaR^{D191N}$ suitable for data collection were obtained in 0.1 M Tris pH 8.5, 0.2 M $LiSO_4$, 25% PEG 3,350 (w/v), 0.05% (v/v) Tween-20, 20% (v/v) glycerol, 10 mM Praseodymium (III) acetate hydrate using native microcrystal seeds of DNA binding LiaR. Prior to data collection, crystals were briefly transferred to 25% (v/v) glycerol plus mother liquor and flash cooled in liquid nitrogen. The diffraction data set was collected to 1.5 Å resolution at Argonne National Laboratory's Advanced Photon Source beamline 21-ID-G on a MarMosaic 300 CCD detector. The crystal belonged to the space group $P22_12_1$ with the unit cell parameters a=31.05, b=76.92, c=76.92, α=β=ɣ=90°.

Crystallization, data collection and structure determination of the $LiaR^{DBD}$and $LiaR^{(DBD)D191N}$/DNA complexes.

The DNA used in the present structures are double-stranded DNA 26-bp in length with blunt ends (*liaXYZ_* updated from DFACE experiment: 5'-CTAGTCCTTACTAATGAGAAGAAAT) and 23-mer oligonucleotide duplex of 22-bp with one G nucleotide overhang on 5'-end (*liaXYZ_* predicted consensus sequence: 5'-GAAATCGTTCTTAAGTCCTATGA). The DBD LiaR and $LiaR^{D191N}$–DNA complex was prepared by mixing 0.5 mM of protein with 0.5 mM DNA duplex. Crystals of the DBD LiaR–22bp DNA complex were grown in 0.2 M Magnesium formate dihydrate, 20%w/v PEG 3.350, 0.012 M Spermine tetrachloride. Before data collection the crystals were gradually transferred to a cryo-protectant solution containing 20% (v/v) glycerol (v/v) and flash-frozen in liquid nitrogen. Diffraction data were processed and scaled with the HKL2000. These crystals belong to space group $P6_2$ with the unit cell parameters a=113.58, b=113.58, c=48.22, α=β=120°, ɣ=90° and $P2_12_12_1$ with the unit cell parameters a=38.34, b=77.25, c=104.98, α=β=ɣ=90°. The molecular replacement method was used to

determine the structures of DNA–protein complexes using isolated DBD structures of S613 *E. faecalis* LiaR^{D191N} solved previously (see above) in Phaser-MR (5), which successfully placed two molecules of the respective domain in the asymmetric unit. After several cycles of rigid-body refinement and restrained refinement using phenix.refine (6), a 2mFo-DFc map revealed electron density of the ds-DNA. The double-stranded DNA was built manually using the 2mFo-Fc and mFo-DFc electron density map as guide in COOT (7). Each round of refinement was followed by manual rebuilding and placement of additional nucleotides as the electron density map improved. Structure refinement was carried out iteratively using PHENIX and included simulated annealing, group B factor refinement, model building and density modification. Water molecules were added using the update water option in phenix.refine and by manual inspection of 2Fo-Fc electron density maps. In the LiaR^{D191N}/22-bp DNA complex structure only 17-bp could be modeled. All structures were validated using a structure validation program MolProbity (1).

CURVES+ analysis.

The scripts used were:

```
/home/kakhanip/curves+/Cur+ <<!
&inp
file=138_renumbered_correct_order,lis=curves_138_renumbered_correct_order,lib=/ho
me/kakhanip/curves+/standard,
&end
2 1 -1 0 0
1:17
18:34
!
/home/kakhanip/curves+/Cur+ <<!
&inp file=127,lis=curves_127,lib=/home/kakhanip/curves+/standard,
&end
2 1 -1 0 0
1:26
52:27
!
```

Overall bending is measured between the ends of the helical axis. To address concerns that the DNA ends might be more distorted and exaggerating the predicted bend we deleted bases (-86), (-100), (-101), (-102) in case of consensus sequence, and region -99-(-101) in case of secondary site, since those bases had stronger evidence for crystal packing effects. Either way however the overall bend did not change dramatically suggesting that the protein on the DNA induces the DNA bending bound to the LiaR surface and is not a crystal packing artifact.

The scripts used were:

```
/home/kakhanip/curves+/Cur+ <<!
```

&inp
file=138,lis=curves_138_smaller_boundary_1_start_3_end,lib=/home/kakhanip/curves+/
standard,
&end
2 1 -1 0 0
2:14
19:31
!
/home/kakhanip/curves+/Cur+ <<!
&inp
file=127,lis=curves_127_smallerboundary_3_from_end,lib=/home/kakhanip/curves+/sta
ndard,
&end
2 1 -1 0 0
1:23
52:30
!

FIGURES LEGENDS.

S1.

A schematic view of the secondary structure of the LiaR$^{(DBD)}$ with the residues shown for each helix. The figure was created using a custom generated pictorial web-based database (PDBsum) (8)**.**

S2.

Closed up view of the binding interface of the LiaR$^{(DBD)D191N}$ bound to DNA sequences derived from the *liaXYZ* consensus and secondary sites.

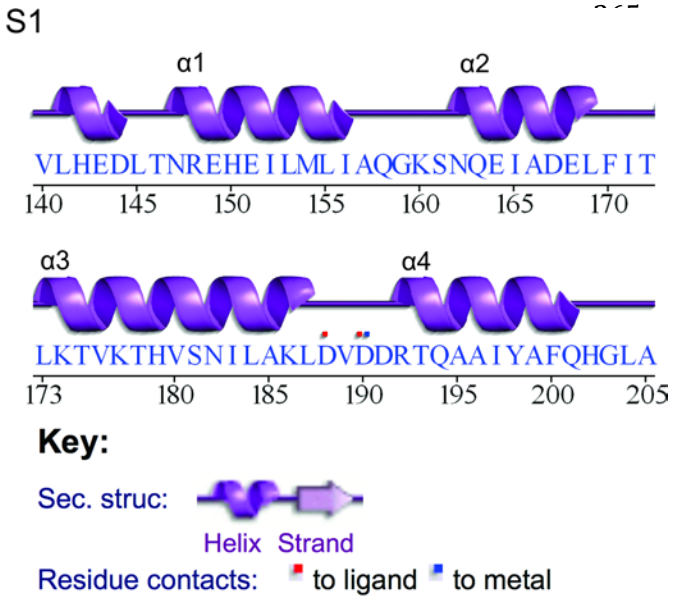
The FEMs (Feature Enhanced Maps) are modified $2mF_{obs}$-DF_{model} σA-weighted maps computed using phenix to reduce the model bias and retain the existing features (6). The (FEM–PHIFEM) electron density map is contoured at 0.6 absolute value of electrons/$Å^3$ to show how Lys174, Lys177 and Thr178 interact with DNA. The consensus sequence bases are indicated as red color.

Supplementary References

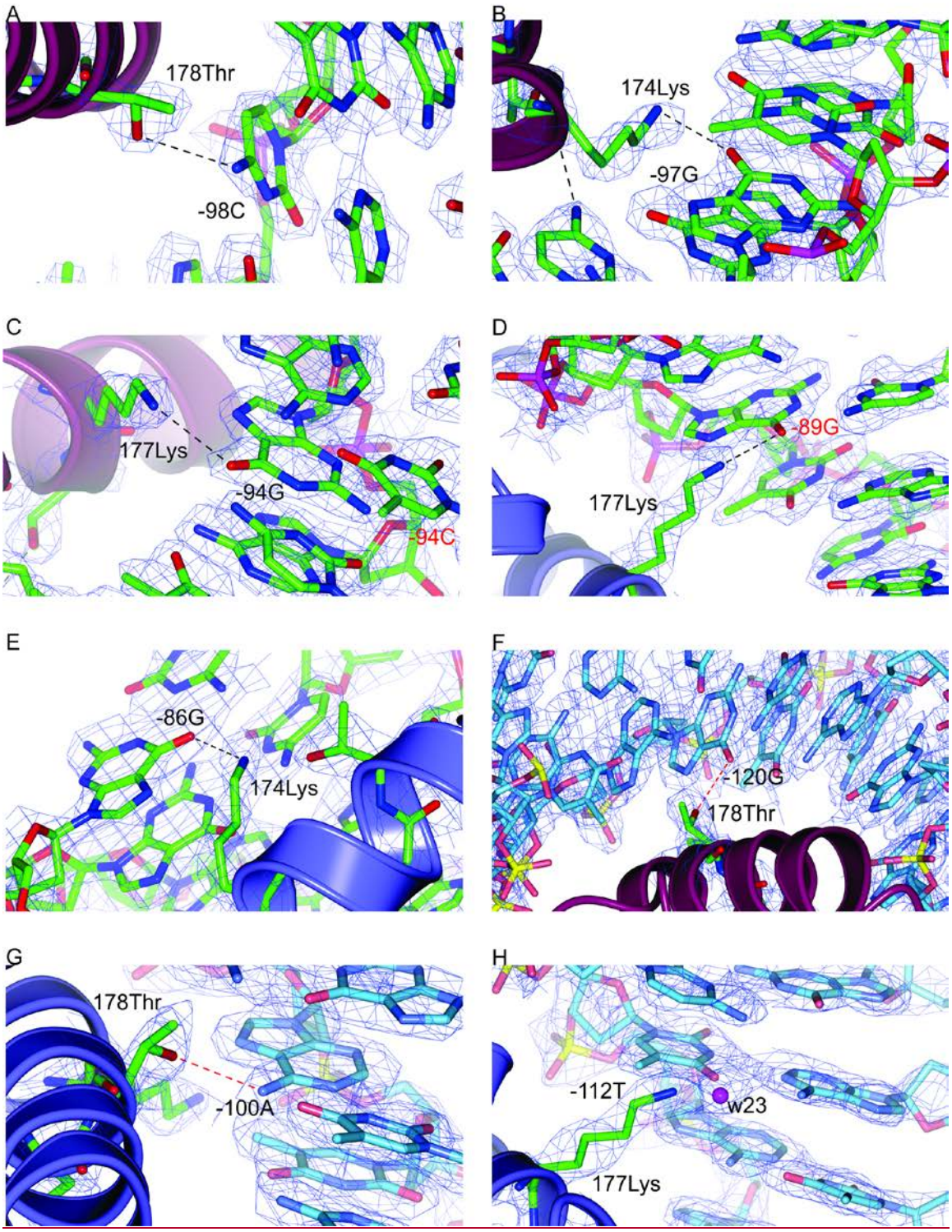
References

1. Chen, V.B., Arendall, W.B., 3rd, Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D Biol Crystallogr*, **66**, 12-21.
2. Laue, T.M., Shah, B.D., Ridgeway, T.M. and Pelletier, S.L. (1992) *Computer-aided interpretation of analytical sedimentation data for proteins. Analytical Ultracentrifugation in Biochemistry and Polymer Science*. Royal Society of Chemistry, Cambridge.
3. Seidel, S.A., Dijkman, P.M., Lea, W.A., van den Bogaart, G., Jerabek-Willemsen, M., Lazic, A., Joseph, J.S., Srinivasan, P., Baaske, P., Simeonov, A. *et al.* (2013) Microscale

thermophoresis quantifies biomolecular interactions under previously challenging conditions. *Methods*, **59**, 301-315.
4. Otwinowski, Z. and Minor, W. (1997) In Charles W. Carter, Jr. (ed.), *Methods in Enzymology*. Academic Press, Vol. Volume 276, pp. 307-326.
5. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C. and Read, R.J. (2007) Phaser crystallographic software. *Journal of applied crystallography*, **40**, 658-674.
6. Adams, P.D., Afonine, P.V., Bunkoczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta crystallographica. Section D, Biological crystallography*, **66**, 213-221.
7. Emsley, P. and Cowtan, K. (2004) Coot: model-building tools for molecular graphics. *Acta crystallographica. Section D, Biological crystallography*, **60**, 2126-2132.
8. Laskowski, R.A. (2001) PDBsum: summaries and analyses of PDB structures. *Nucleic Acids Res*, **29**, 221-222.



S2


A
178Thr
-98C
B
174Lys
-97G
C
177Lys
-94G
-94C
D
-89G
177Lys
E
-86G
174Lys
F
-120G
178Thr
G
178Thr
-100A
H
-112T
w23
177Lys

282