Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 09:28 pm BST

PDB ID	:	6E95
Title	:	Chimeric structure of Saccharomyces cerevisiae GCN4 leucine zipper fused to Staphylococcus aureus AgrC cytoplasmic histidine kinase module (dataset isotropically truncated by HKL2000)
Authors	:	Xie, Q.; Jeffrey, P.D.; Muir, T.W.
Deposited on	:	2018-07-31
Resolution	:	2.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see references ①) were used in the production of this report:

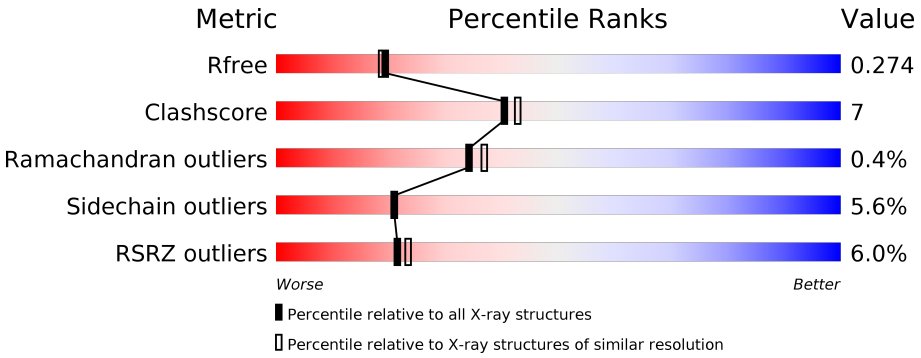
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1377 (2.26-2.26)
Clashscore	141614	1487 (2.26-2.26)
Ramachandran outliers	138981	1449 (2.26-2.26)
Sidechain outliers	138945	1450 (2.26-2.26)
RSRZ outliers	127900	1356 (2.26-2.26)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	250	3% 81% 13%
1	B	250	8% 77% 14% 8%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4071 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Staphylococcus aureus AgrC histidine kinase module fused to Saccharomyces cerevisiae GCN4 leucine zipper.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	240	Total	C	N	O	S	0	1	0
			1955	1233	329	384	9			
1	B	229	Total	C	N	O	S	0	3	0
			1861	1174	311	367	9			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	GLY	-	expression tag	UNP P03069
A	2	SER	-	expression tag	UNP P03069
A	3	HIS	-	expression tag	UNP P03069
B	1	GLY	-	expression tag	UNP P03069
B	2	SER	-	expression tag	UNP P03069
B	3	HIS	-	expression tag	UNP P03069

- Molecule 2 is water.

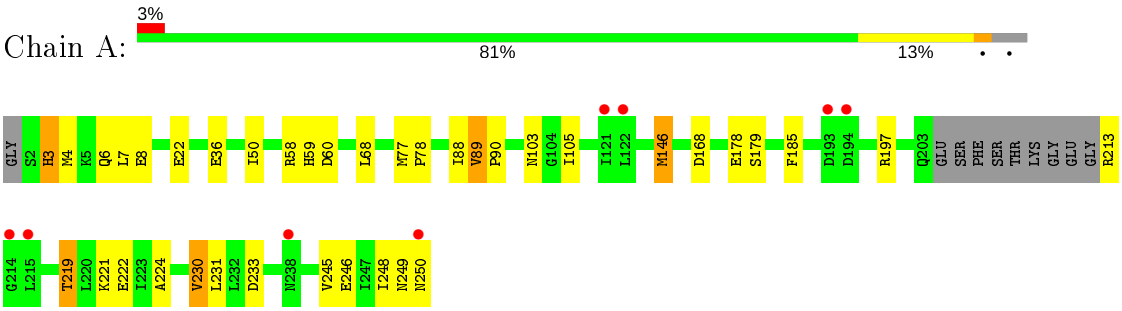
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	149	Total	O	0	0
			149	149		
2	B	106	Total	O	0	0
			106	106		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

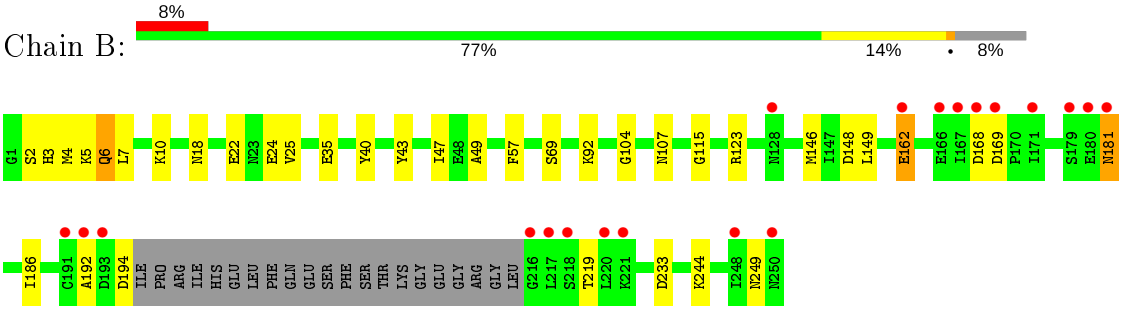
• Molecule 1: Staphylococcus aureus AgrC histidine kinase module fused to Saccharomyces cerevisiae GCN4 leucine zipper

Chain A:



• Molecule 1: Staphylococcus aureus AgrC histidine kinase module fused to Saccharomyces cerevisiae GCN4 leucine zipper

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 2 21 21	Depositor
Cell constants a, b, c, α, β, γ	63.14Å 83.38Å 146.97Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	34.81 – 2.25 34.79 – 2.25	Depositor EDS
% Data completeness (in resolution range)	99.7 (34.81-2.25) 99.8 (34.79-2.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.86 (at 2.24Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.219 , 0.275 0.223 , 0.274	Depositor DCC
R_{free} test set	1869 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å^2)	46.0	Xtriage
Anisotropy	0.817	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 62.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4071	wwPDB-VP
Average B, all atoms (Å^2)	72.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.49% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with $|Z| > 5$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\#\|Z\| > 5$	RMSZ	$\#\|Z\| > 5$
1	A	0.69	0/1981	0.85	0/2670
1	B	0.70	0/1891	0.82	0/2548
All	All	0.69	0/3872	0.84	0/5218

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1955	0	1972	35	0
1	B	1861	0	1880	27	0
2	A	149	0	0	14	3
2	B	106	0	0	7	2
All	All	4071	0	3852	55	3

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (55) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:4:MET:O	2:A:301:HOH:O	2.01	0.79

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:60:ASP:OD1	2:A:302:HOH:O	2.04	0.76
1:A:3:HIS:NE2	2:A:305:HOH:O	2.24	0.70
1:A:7:LEU:HD12	2:B:331:HOH:O	1.93	0.68
1:A:178:GLU:O	2:A:303:HOH:O	2.14	0.66
1:A:224:ALA:HA	1:A:230:VAL:CG2	2.26	0.64
1:B:104:GLY:N	2:B:305:HOH:O	2.30	0.63
1:B:148:ASP:OD1	2:B:301:HOH:O	2.15	0.63
1:B:10:LYS:NZ	2:B:302:HOH:O	2.20	0.62
1:B:24:GLU:HG3	2:B:368:HOH:O	2.00	0.62
1:B:3:HIS:O	1:B:6:GLN:HG3	2.02	0.60
1:A:59:HIS:CD2	2:A:302:HOH:O	2.58	0.56
1:B:146:MET:HA	1:B:146:MET:HE3	1.88	0.55
1:B:123:ARG:NH2	1:B:162:GLU:OE2	2.38	0.55
1:B:181:ASN:C	1:B:249:ASN:ND2	2.60	0.55
1:B:22:GLU:HG3	2:B:304:HOH:O	2.07	0.55
1:A:36:GLU:HB3	2:A:390:HOH:O	2.07	0.54
1:B:146:MET:CE	1:B:149:LEU:HD23	2.38	0.54
1:B:181:ASN:C	1:B:249:ASN:HD22	2.10	0.54
1:B:181:ASN:O	1:B:249:ASN:ND2	2.40	0.52
1:A:4:MET:HA	2:A:301:HOH:O	2.10	0.51
1:A:3:HIS:O	1:A:6:GLN:HB2	2.11	0.50
1:B:146:MET:HA	1:B:146:MET:CE	2.41	0.50
1:A:7:LEU:HD23	1:A:7:LEU:N	2.27	0.50
1:B:4:MET:O	1:B:5:LYS:C	2.48	0.50
1:A:7:LEU:HB2	2:A:301:HOH:O	2.12	0.49
1:A:4:MET:HG3	1:B:4:MET:CE	2.44	0.48
1:A:89:VAL:HG12	1:A:90:PRO:HD3	1.96	0.48
1:A:50:ILE:HD13	1:B:47:ILE:HD11	1.98	0.46
1:A:59:HIS:HD2	2:A:302:HOH:O	1.96	0.46
2:A:329:HOH:O	1:B:25:VAL:HG11	2.15	0.45
1:A:68:LEU:HD21	1:A:88:ILE:CD1	2.48	0.44
1:A:89:VAL:N	1:A:90:PRO:CD	2.80	0.44
1:A:185:PHE:HB3	1:A:245:VAL:HG22	1.98	0.44
1:B:49:ALA:HA	1:B:115:GLY:HA3	1.99	0.44
1:B:146:MET:HE3	1:B:149:LEU:HD23	2.00	0.44
1:A:230:VAL:HA	1:A:246:GLU:O	2.18	0.44
1:A:103:ASN:ND2	2:A:321:HOH:O	2.51	0.43
1:A:58:ARG:HG2	1:B:57:PHE:CE1	2.54	0.43
1:A:4:MET:CA	2:A:301:HOH:O	2.67	0.43
1:A:231:LEU:N	1:A:231:LEU:HD12	2.34	0.43
1:A:219:THR:HA	1:A:222:GLU:HB2	1.99	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:249:ASN:O	2:A:304:HOH:O	2.21	0.42
1:A:197:ARG:NH2	2:A:320:HOH:O	2.51	0.42
1:A:146:MET:HG2	1:B:40:TYR:CD2	2.55	0.42
1:A:77:MET:HB3	1:A:78:PRO:HD3	2.01	0.41
1:A:219:THR:OG1	1:A:219:THR:O	2.32	0.41
1:A:4:MET:HA	1:B:4:MET:HE1	2.03	0.41
1:A:248:ILE:HG22	1:A:250:ASN:H	1.85	0.41
1:B:186:ILE:HG12	1:B:244:LYS:HG3	2.02	0.41
1:B:18:ASN:O	1:B:22:GLU:HB2	2.21	0.41
1:A:178:GLU:HG2	1:A:179:SER:O	2.22	0.40
1:A:8:GLU:CG	1:B:7:LEU:HD11	2.50	0.40
1:B:104:GLY:CA	2:B:305:HOH:O	2.68	0.40
1:A:4:MET:HG3	1:B:4:MET:HE2	2.03	0.40

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:421:HOH:O	2:A:434:HOH:O[2_565]	1.88	0.32
2:A:444:HOH:O	2:B:333:HOH:O[4_655]	2.08	0.12
2:A:431:HOH:O	2:B:401:HOH:O[3_645]	2.11	0.09

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	237/250 (95%)	223 (94%)	13 (6%)	1 (0%)	34	37
1	B	228/250 (91%)	214 (94%)	13 (6%)	1 (0%)	34	37
All	All	465/500 (93%)	437 (94%)	26 (6%)	2 (0%)	34	37

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	146	MET
1	B	192	ALA

5.3.2 Protein sidechains (i)

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	222/229 (97%)	212 (96%)	10 (4%)	27 31
1	B	211/229 (92%)	197 (93%)	14 (7%)	16 15
All	All	433/458 (94%)	409 (94%)	24 (6%)	21 21

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	HIS
1	A	22	GLU
1	A	89	VAL
1	A	105	ILE
1	A	168	ASP
1	A	213	ARG
1	A	219	THR
1	A	221	LYS
1	A	230	VAL
1	A	233	ASP
1	B	2	SER
1	B	6	GLN
1	B	35	GLU
1	B	43	TYR
1	B	69	SER
1	B	92	LYS
1	B	107	ASN
1	B	162	GLU
1	B	168	ASP
1	B	169	ASP
1	B	181	ASN
1	B	194	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	219	THR
1	B	233	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	181	ASN
1	B	103	ASN
1	B	107	ASN
1	B	143	ASN
1	B	249	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

There are no ligands in this entry.

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å^2)	Q<0.9
1	A	240/250 (96%)	0.18	8 (3%) 46 48	34, 62, 113, 139	0
1	B	229/250 (91%)	0.48	20 (8%) 10 11	38, 72, 120, 138	0
All	All	469/500 (93%)	0.33	28 (5%) 21 23	34, 67, 117, 139	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	216	GLY	5.8
1	B	168	ASP	5.7
1	B	250	ASN	4.9
1	B	166	GLU	4.5
1	B	167	ILE	3.8
1	B	217	LEU	3.8
1	B	181	ASN	3.8
1	B	192	ALA	3.7
1	B	191	CYS	3.6
1	B	248	ILE	3.5
1	B	169	ASP	3.5
1	B	180	GLU	3.5
1	A	193	ASP	2.9
1	A	194	ASP	2.8
1	B	218	SER	2.8
1	B	221	LYS	2.7
1	B	162	GLU	2.5
1	A	250	ASN	2.5
1	A	215	LEU	2.4
1	B	128	ASN	2.4
1	B	193	ASP	2.3
1	A	122	LEU	2.3
1	B	220	LEU	2.2
1	A	238	ASN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	121	ILE	2.1
1	A	214	GLY	2.1
1	B	179	SER	2.1
1	B	171	ILE	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates (i)

There are no carbohydrates in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.