Full wwPDB X-ray Structure Validation Report ⓘ

Jun 23, 2024 – 09:38 AM EDT

PDB ID	:	6E52
Title	:	Chimeric structure of Saccharomyces cerevisiae GCN4 leucine zipper fused to Staphylococcus aureus AgrC cytoplasmic histidine kinase module (dataset anisotropically truncated by STARANISO)
Authors	:	Xie, Q.; Jeffrey, P.D.; Muir, T.W.
Deposited on	:	2018-07-19
Resolution	:	1.93 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

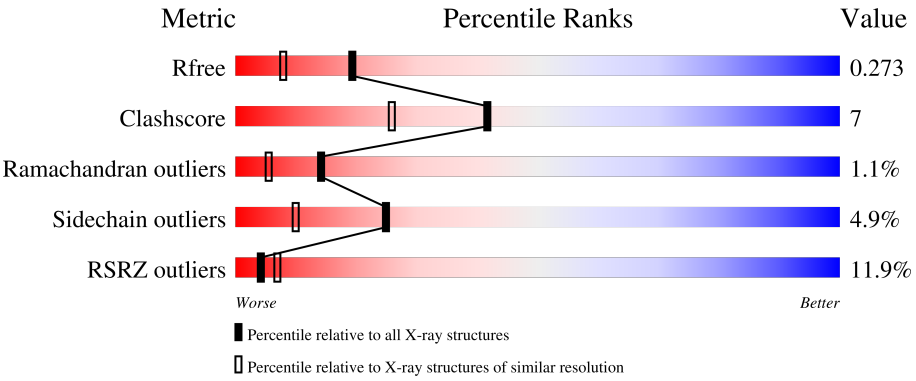
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.93 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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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 of 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	6% 80% 16%
1	B	250	16% 74% 16% 8%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4088 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	243	Total	C	N	O	S	0	1	0
			1977	1245	332	391	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.

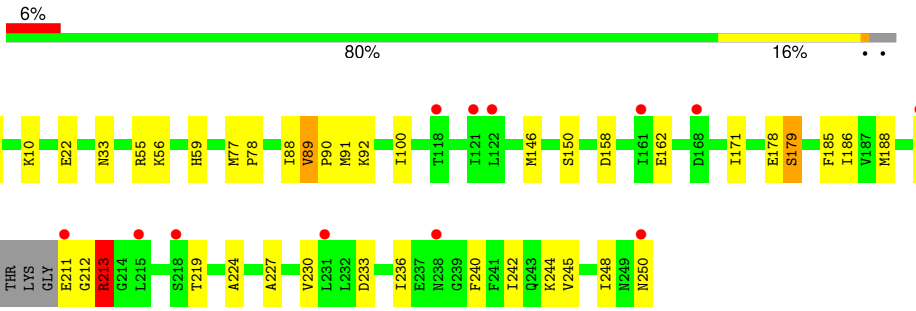
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	148	Total	O	0	0
			148	148		
2	B	102	Total	O	0	0
			102	102		

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

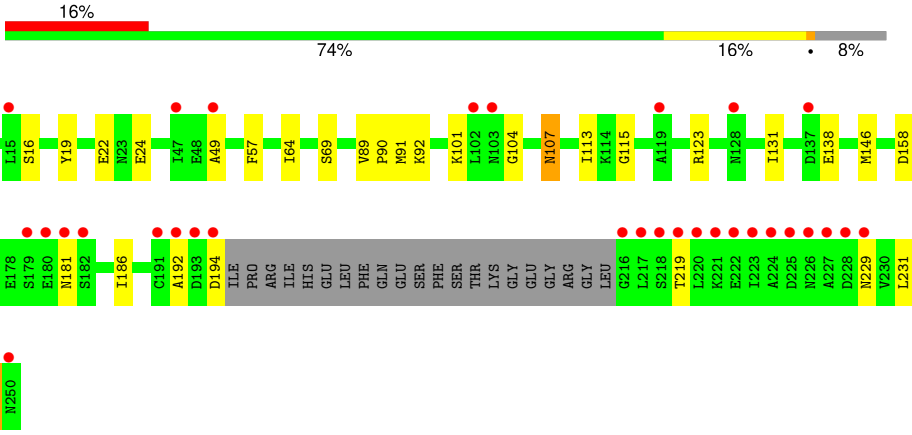
• Molecule 1: Staphylococcus aureus AgrC histidine kinase module fused to Saccharomyces cerevisiae GCN4 leucine zipper

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)	30.00 – 1.93 31.76 – 1.93	Depositor EDS
% Data completeness (in resolution range)	62.4 (30.00-1.93) 62.4 (31.76-1.93)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$< I/\sigma(I) >$ [1]	1.12 (at 1.94Å)	Xtriage
Refinement program	REFMAC 5.8.0238	Depositor
R, R_{free}	0.218 , 0.268 0.224 , 0.273	Depositor DCC
R_{free} test set	1830 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å^2)	34.8	Xtriage
Anisotropy	0.157	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 58.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	4088	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.11% 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	0/2003	0.85	0/2699
1	B	0.71	0/1891	0.79	0/2548
All	All	0.70	0/3894	0.82	0/5247

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	1977	0	1987	27	0
1	B	1861	0	1880	30	0
2	A	148	0	0	8	3
2	B	102	0	0	5	2
All	All	4088	0	3867	54	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 (54) 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:B:181:ASN:C	1:B:249:ASN:HD21	1.71	0.94

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:181:ASN:O	1:B:249:ASN:ND2	2.08	0.85
1:A:55:ARG:NH1	2:A:302:HOH:O	2.11	0.81
1:B:146:MET:HE3	1:B:146:MET:HA	1.67	0.77
1:B:181:ASN:C	1:B:249:ASN:ND2	2.40	0.73
1:A:59:HIS:HD2	2:A:313:HOH:O	1.81	0.63
1:A:56:LYS:HD2	2:A:313:HOH:O	1.98	0.62
1:A:211:GLU:O	1:A:213:ARG:N	2.34	0.61
1:B:3:HIS:O	1:B:6:GLN:HG3	2.01	0.60
1:A:59:HIS:CD2	2:A:313:HOH:O	2.56	0.59
1:B:113:ILE:HD11	1:B:146:MET:HE2	1.84	0.59
1:A:3:HIS:NE2	2:A:301:HOH:O	2.07	0.58
1:B:249:ASN:ND2	1:B:249:ASN:H	2.01	0.58
1:B:113:ILE:HD11	1:B:146:MET:CE	2.33	0.58
1:B:4:MET:O	1:B:5:LYS:C	2.43	0.57
1:A:185:PHE:HB3	1:A:245:VAL:HG22	1.87	0.56
1:B:16:SER:HB3	2:B:383:HOH:O	2.07	0.54
1:B:22:GLU:HG3	2:B:316:HOH:O	2.07	0.54
1:A:224:ALA:HA	1:A:230:VAL:HG23	1.90	0.53
1:A:186:ILE:HG12	1:A:244:LYS:HG3	1.91	0.52
1:A:224:ALA:HA	1:A:230:VAL:CG2	2.41	0.51
1:B:16:SER:O	1:B:19:TYR:HB3	2.11	0.51
1:A:227:ALA:HB3	1:A:230:VAL:HG22	1.91	0.51
1:B:104:GLY:N	2:B:308:HOH:O	2.45	0.50
1:B:231:LEU:HD11	1:B:248:ILE:HD11	1.94	0.50
1:B:89:VAL:N	1:B:90:PRO:HD2	2.28	0.49
1:B:24:GLU:HG3	2:B:370:HOH:O	2.12	0.49
1:A:4:MET:HA	1:B:4:MET:HE1	1.95	0.48
1:B:167:ILE:HG22	1:B:168:ASP:O	2.13	0.48
1:A:55:ARG:HD3	1:A:55:ARG:C	2.34	0.47
1:A:89:VAL:N	1:A:90:PRO:CD	2.77	0.47
1:A:4:MET:HB2	1:B:4:MET:HE2	1.95	0.47
1:B:186:ILE:HG12	1:B:244:LYS:HG3	1.96	0.47
1:B:123:ARG:NH1	1:B:158:ASP:OD1	2.48	0.47
1:A:197:ARG:NH2	2:A:318:HOH:O	2.49	0.46
1:A:178:GLU:HG2	1:A:179:SER:O	2.16	0.46
1:A:77:MET:N	1:A:78:PRO:HD2	2.32	0.45
1:A:248:ILE:HG22	1:A:250:ASN:H	1.82	0.45
1:A:188:MET:HG3	1:A:242:ILE:HG12	1.97	0.45
1:A:100:ILE:HD13	2:A:364:HOH:O	2.17	0.44
1:A:88:ILE:O	1:A:91:MET:HG3	2.18	0.43
1:B:229:ASN:O	1:B:248:ILE:HD12	2.17	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:113:ILE:CG1	1:B:146:MET:HE1	2.49	0.43
1:B:131:ILE:HA	1:B:172:ILE:O	2.18	0.43
1:A:236:ILE:HA	1:A:240:PHE:O	2.19	0.43
1:A:4:MET:CB	1:B:4:MET:HE2	2.49	0.42
1:B:57:PHE:HA	2:B:340:HOH:O	2.20	0.42
1:B:64:ILE:HD11	1:B:91:MET:HE1	2.01	0.42
1:B:49:ALA:HA	1:B:115:GLY:HA3	2.01	0.42
1:B:146:MET:HA	1:B:146:MET:CE	2.42	0.42
1:A:158:ASP:O	1:A:162:GLU:HG2	2.20	0.41
1:A:33:ASN:CB	2:A:351:HOH:O	2.68	0.41
1:A:7:LEU:O	1:A:10:LYS:N	2.54	0.41
1:B:107:ASN:ND2	1:B:138[A]:GLU:HG2	2.36	0.41

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:423:HOH:O	2:A:436:HOH:O[2_565]	1.86	0.34
2:A:443:HOH:O	2:B:335:HOH:O[4_655]	2.11	0.09
2:A:427:HOH:O	2:B:398:HOH:O[3_645]	2.16	0.04

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	240/250 (96%)	225 (94%)	11 (5%)	4 (2%)	9 2
1	B	228/250 (91%)	216 (95%)	11 (5%)	1 (0%)	34 24
All	All	468/500 (94%)	441 (94%)	22 (5%)	5 (1%)	14 5

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	212	GLY
1	B	192	ALA
1	A	146	MET
1	A	213	ARG
1	A	179	SER

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	224/229 (98%)	215 (96%)	9 (4%)	31 16
1	B	211/229 (92%)	199 (94%)	12 (6%)	20 8
All	All	435/458 (95%)	414 (95%)	21 (5%)	25 10

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	22	GLU
1	A	89	VAL
1	A	92	LYS
1	A	150	SER
1	A	171	ILE
1	A	194	ASP
1	A	213	ARG
1	A	219	THR
1	A	233	ASP
1	B	2	SER
1	B	69	SER
1	B	92	LYS
1	B	101	LYS
1	B	107	ASN
1	B	162	GLU
1	B	168	ASP
1	B	169	ASP
1	B	194	ASP
1	B	219	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	234	THR
1	B	249	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	143	ASN
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 monosaccharides 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	243/250 (97%)	0.53	16 (6%) 18 24	21, 42, 84, 110	0
1	B	229/250 (91%)	0.87	40 (17%) 1 2	23, 53, 93, 107	0
All	All	472/500 (94%)	0.69	56 (11%) 4 7	21, 47, 91, 110	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	223	ILE	5.6
1	B	217	LEU	5.5
1	B	167	ILE	5.5
1	B	221	LYS	5.3
1	B	168	ASP	5.2
1	B	227	ALA	4.7
1	B	250	ASN	4.6
1	B	218	SER	4.5
1	B	166	GLU	4.4
1	B	226	ASN	4.3
1	B	216	GLY	4.3
1	B	224	ALA	4.0
1	B	192	ALA	3.7
1	A	250	ASN	3.6
1	B	248	ILE	3.6
1	B	225	ASP	3.6
1	B	228	ASP	3.4
1	B	169	ASP	3.3
1	B	177	ILE	3.3
1	A	194	ASP	3.2
1	B	220	LEU	3.2
1	B	180	GLU	3.2
1	B	103	ASN	3.2
1	A	215	LEU	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	222	GLU	3.1
1	B	219	THR	3.0
1	A	204	GLU	3.0
1	B	181	ASN	2.9
1	A	231	LEU	2.9
1	B	47	ILE	2.8
1	A	168	ASP	2.8
1	B	229	ASN	2.7
1	A	200	GLU	2.7
1	B	128	ASN	2.7
1	B	102	LEU	2.6
1	B	233	ASP	2.6
1	B	179	SER	2.5
1	B	182	SER	2.5
1	A	199	HIS	2.5
1	A	122	LEU	2.5
1	A	218	SER	2.4
1	A	211	GLU	2.4
1	B	194	ASP	2.4
1	A	6	GLN	2.4
1	B	193	ASP	2.3
1	B	119	ALA	2.3
1	A	161	ILE	2.3
1	A	238	ASN	2.3
1	B	191	CYS	2.3
1	B	15	LEU	2.3
1	B	137	ASP	2.2
1	B	247	ILE	2.2
1	A	118	THR	2.1
1	A	121	ILE	2.1
1	B	242	ILE	2.1
1	B	49	ALA	2.1

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.