



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:59 PM GMT

PDB ID : 4QLO  
Title : Crystal Structure of homoserine o-acetyltransferase from *Staphylococcus aureus*  
Authors : Thangavelu, B.; Pavlovsky, A.G.; Viola, R.E.  
Deposited on : 2014-06-12  
Resolution : 2.45 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](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

MolProbity : 4.02b-467  
Mogul : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

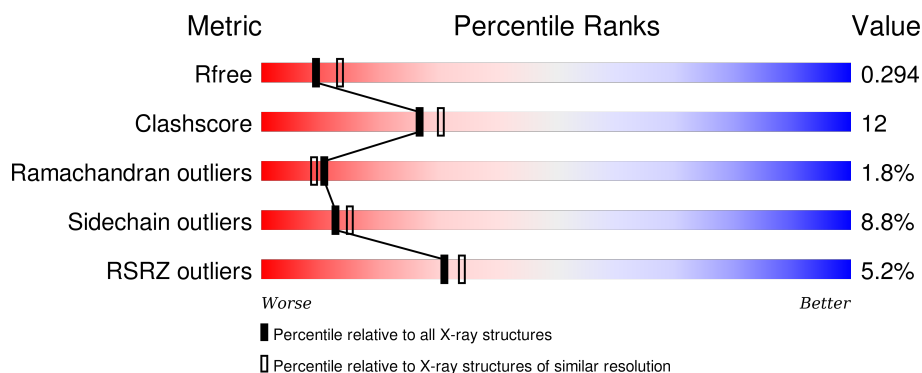
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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))
$R_{free}$	91344	4776 (2.50-2.42)
Clashscore	102246	1030 (2.48-2.44)
Ramachandran outliers	100387	1024 (2.48-2.44)
Sidechain outliers	100360	1024 (2.48-2.44)
RSRZ outliers	91569	4787 (2.50-2.42)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	360	<div> <div>5%</div> <div>69%</div> <div>18%</div> <div>• • 9%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2688 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 homoserine O-acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	327	Total	C	N	O	S	0	1	0
			2660	1706	461	488	5			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	323	ASP	-	EXPRESSION TAG	UNP A8YYT5
A	324	PRO	-	EXPRESSION TAG	UNP A8YYT5
A	325	ALA	-	EXPRESSION TAG	UNP A8YYT5
A	326	PHE	-	EXPRESSION TAG	UNP A8YYT5
A	327	LEU	-	EXPRESSION TAG	UNP A8YYT5
A	328	TYR	-	EXPRESSION TAG	UNP A8YYT5
A	329	LYS	-	EXPRESSION TAG	UNP A8YYT5
A	330	VAL	-	EXPRESSION TAG	UNP A8YYT5
A	331	VAL	-	EXPRESSION TAG	UNP A8YYT5
A	332	ILE	-	EXPRESSION TAG	UNP A8YYT5
A	333	ASN	-	EXPRESSION TAG	UNP A8YYT5
A	334	SER	-	EXPRESSION TAG	UNP A8YYT5
A	335	LYS	-	EXPRESSION TAG	UNP A8YYT5
A	336	LEU	-	EXPRESSION TAG	UNP A8YYT5
A	337	GLU	-	EXPRESSION TAG	UNP A8YYT5
A	338	GLY	-	EXPRESSION TAG	UNP A8YYT5
A	339	LYS	-	EXPRESSION TAG	UNP A8YYT5
A	340	PRO	-	EXPRESSION TAG	UNP A8YYT5
A	341	ILE	-	EXPRESSION TAG	UNP A8YYT5
A	342	PRO	-	EXPRESSION TAG	UNP A8YYT5
A	343	ASN	-	EXPRESSION TAG	UNP A8YYT5
A	344	PRO	-	EXPRESSION TAG	UNP A8YYT5
A	345	LEU	-	EXPRESSION TAG	UNP A8YYT5
A	346	LEU	-	EXPRESSION TAG	UNP A8YYT5
A	347	GLY	-	EXPRESSION TAG	UNP A8YYT5
A	348	LEU	-	EXPRESSION TAG	UNP A8YYT5
A	349	ASP	-	EXPRESSION TAG	UNP A8YYT5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	350	SER	-	EXPRESSION TAG	UNP A8YYT5
A	351	THR	-	EXPRESSION TAG	UNP A8YYT5
A	352	ARG	-	EXPRESSION TAG	UNP A8YYT5
A	353	THR	-	EXPRESSION TAG	UNP A8YYT5
A	354	GLY	-	EXPRESSION TAG	UNP A8YYT5
A	355	HIS	-	EXPRESSION TAG	UNP A8YYT5
A	356	HIS	-	EXPRESSION TAG	UNP A8YYT5
A	357	HIS	-	EXPRESSION TAG	UNP A8YYT5
A	358	HIS	-	EXPRESSION TAG	UNP A8YYT5
A	359	HIS	-	EXPRESSION TAG	UNP A8YYT5
A	360	HIS	-	EXPRESSION TAG	UNP A8YYT5

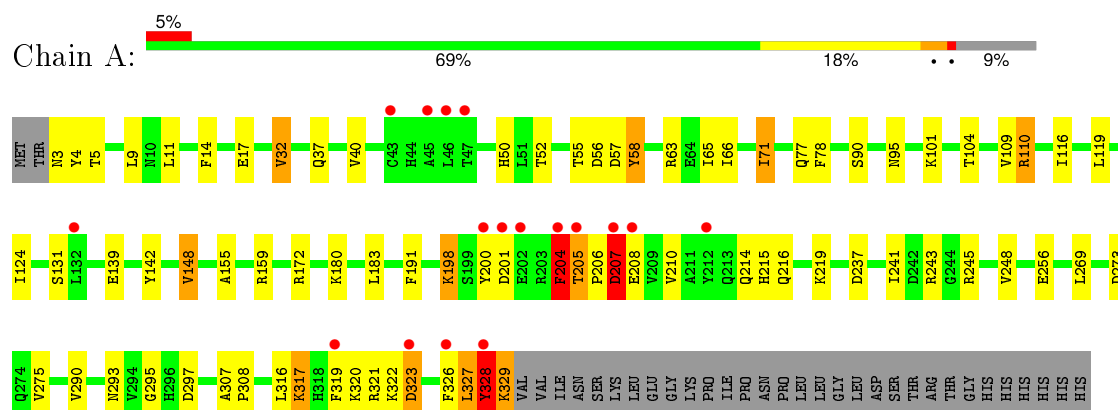
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	28	Total O 28 28	0	0

### 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 errors 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: homoserine O-acetyltransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.16 Å 49.16 Å 481.42 Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	50.00 – 2.45 38.93 – 2.45	Depositor EDS
% Data completeness (in resolution range)	91.3 (50.00-2.45) 91.3 (38.93-2.45)	Depositor EDS
$R_{merge}$	0.15	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.19 (at 2.45 Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, $R_{free}$	0.240 , 0.291 0.243 , 0.294	Depositor DCC
$R_{free}$ test set	635 reflections (5.22%)	DCC
Wilson B-factor (Å <sup>2</sup> )	43.8	Xtriage
Anisotropy	0.830	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.30 , 31.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 12790 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2688	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.20% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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	0.62	0/2735	0.72	1/3709 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
1	A	58	TYR	C-N-CD	5.96	140.91	128.40

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	2660	0	2561	63	0
2	A	28	0	0	0	0
All	All	2688	0	2561	63	0

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 12.

All (63) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance ( $\text{\AA}$ )	Clash overlap ( $\text{\AA}$ )
1:A:205:THR:CG2	1:A:208:GLU:OE1	1.71	1.39

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:205:THR:HG22	1:A:208:GLU:OE1	1.03	1.18
1:A:206:PRO:O	1:A:207:ASP:HB2	1.54	1.07
1:A:205:THR:HG22	1:A:208:GLU:CD	1.81	1.00
1:A:326:PHE:O	1:A:329:LYS:O	1.78	1.00
1:A:215:HIS:NE2	1:A:219:LYS:HE3	1.78	0.97
1:A:205:THR:HB	1:A:208:GLU:HB2	1.49	0.93
1:A:205:THR:HB	1:A:208:GLU:CB	2.04	0.85
1:A:205:THR:HG21	1:A:208:GLU:OE1	1.80	0.81
1:A:241:ILE:HD12	1:A:248:VAL:HG11	1.62	0.81
1:A:327:LEU:O	1:A:328:TYR:CD2	2.34	0.80
1:A:205:THR:CB	1:A:208:GLU:HB2	2.13	0.79
1:A:55:THR:HG22	1:A:57:ASP:H	1.47	0.78
1:A:329:LYS:NZ	1:A:329:LYS:HB2	2.05	0.71
1:A:55:THR:HG22	1:A:57:ASP:N	2.07	0.70
1:A:206:PRO:O	1:A:207:ASP:CB	2.36	0.67
1:A:329:LYS:HZ2	1:A:329:LYS:HB2	1.60	0.66
1:A:116:ILE:O	1:A:119:LEU:O	2.13	0.66
1:A:328:TYR:N	1:A:329:LYS:HA	2.10	0.65
1:A:17:GLU:OE1	1:A:104:THR:HG21	2.00	0.61
1:A:9:LEU:HB3	1:A:11:LEU:HD21	1.83	0.61
1:A:4:TYR:HD1	1:A:32:VAL:HG13	1.65	0.61
1:A:172:ARG:NH2	1:A:237:ASP:OD2	2.34	0.60
1:A:205:THR:HB	1:A:208:GLU:HB3	1.85	0.58
1:A:37:GLN:HB2	1:A:77:GLN:HB2	1.86	0.56
1:A:55:THR:CG2	1:A:56:ASP:N	2.69	0.55
1:A:326:PHE:O	1:A:328:TYR:N	2.40	0.55
1:A:205:THR:O	1:A:208:GLU:N	2.36	0.54
1:A:198:LYS:O	1:A:201:ASP:HB3	2.07	0.54
1:A:317:LYS:HG3	1:A:317:LYS:O	2.07	0.53
1:A:322:LYS:O	1:A:323:ASP:HB2	2.09	0.52
1:A:210:VAL:O	1:A:214:GLN:HG2	2.09	0.52
1:A:32:VAL:HG23	1:A:78:PHE:HB2	1.90	0.52
1:A:256:GLU:HG3	1:A:320:LYS:HE3	1.92	0.52
1:A:243:ARG:O	1:A:245:ARG:HG3	2.10	0.51
1:A:65:ILE:H	1:A:65:ILE:HD12	1.76	0.51
1:A:328:TYR:N	1:A:329:LYS:CA	2.73	0.50
1:A:124:ILE:O	1:A:148:VAL:HA	2.13	0.48
1:A:327:LEU:O	1:A:328:TYR:CG	2.67	0.47
1:A:55:THR:HG22	1:A:58:TYR:H	1.79	0.47
1:A:205:THR:CB	1:A:208:GLU:CB	2.82	0.47
1:A:293:ASN:OD1	1:A:295:GLY:N	2.42	0.47

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:327:LEU:O	1:A:328:TYR:HD2	1.93	0.46
1:A:205:THR:O	1:A:208:GLU:HB2	2.16	0.46
1:A:131:SER:CB	1:A:155:ALA:O	2.64	0.46
1:A:215:HIS:O	1:A:219:LYS:HD2	2.17	0.45
1:A:58:TYR:N	1:A:58:TYR:CD1	2.84	0.45
1:A:328:TYR:H	1:A:329:LYS:CA	2.30	0.44
1:A:55:THR:CG2	1:A:57:ASP:H	2.24	0.44
1:A:55:THR:HG22	1:A:56:ASP:N	2.32	0.44
1:A:180:LYS:O	1:A:183:LEU:HB2	2.17	0.44
1:A:14:PHE:HE1	1:A:110:ARG:HG3	1.83	0.43
1:A:109:VAL:HG21	1:A:139:GLU:HB2	2.01	0.42
1:A:210:VAL:O	1:A:214:GLN:CG	2.68	0.42
1:A:191:PHE:CD1	1:A:191:PHE:O	2.73	0.42
1:A:139:GLU:O	1:A:142:TYR:N	2.53	0.41
1:A:90:SER:O	1:A:95:ASN:ND2	2.53	0.41
1:A:215:HIS:HD2	1:A:216:GLN:NE2	2.19	0.41
1:A:66:ILE:HG23	1:A:71:ILE:O	2.21	0.41
1:A:307:ALA:N	1:A:308:PRO:CD	2.84	0.40
1:A:204:PHE:C	1:A:206:PRO:HD3	2.41	0.40
1:A:327:LEU:O	1:A:328:TYR:CB	2.69	0.40
1:A:50:HIS:O	1:A:52:THR:N	2.54	0.40

There are no symmetry-related clashes.

## 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	326/360 (91%)	302 (93%)	18 (6%)	6 (2%)	<b>11</b> <b>9</b>

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	204	PHE
1	A	323	ASP
1	A	328	TYR
1	A	198	LYS
1	A	207	ASP
1	A	317	LYS

### 5.3.2 Protein sidechains ⓘ

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	284/315 (90%)	258 (91%)	26 (9%)	11	13

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	3	ASN
1	A	5	THR
1	A	32	VAL
1	A	40	VAL
1	A	63	ARG
1	A	71	ILE
1	A	101	LYS
1	A	110	ARG
1	A	148	VAL
1	A	159[A]	ARG
1	A	159[B]	ARG
1	A	200	TYR
1	A	204	PHE
1	A	205	THR
1	A	207	ASP
1	A	269	LEU
1	A	273	ASP
1	A	275	VAL
1	A	290	VAL
1	A	297	ASP
1	A	316	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	319	PHE
1	A	321	ARG
1	A	327	LEU
1	A	328	TYR
1	A	329	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	31	HIS
1	A	95	ASN
1	A	125	ASN
1	A	168	ASN
1	A	214	GLN
1	A	215	HIS
1	A	216	GLN
1	A	296	HIS
1	A	312	HIS

### 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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	327/360 (90%)	0.00	17 (5%) 31 34	39, 54, 93, 129	0

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	326	PHE	4.8
1	A	328	TYR	3.9
1	A	204	PHE	3.8
1	A	319	PHE	3.6
1	A	212	TYR	3.6
1	A	207	ASP	3.2
1	A	208	GLU	2.9
1	A	202	GLU	2.7
1	A	47	THR	2.5
1	A	201	ASP	2.4
1	A	323	ASP	2.3
1	A	46	LEU	2.1
1	A	45	ALA	2.1
1	A	43	CYS	2.1
1	A	132	LEU	2.1
1	A	200	TYR	2.0
1	A	205	THR	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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.