



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 13, 2020 – 12:17 PM BST

PDB ID : 3PGY  
Title : Serine hydroxymethyltransferase from Staphylococcus aureus, S95P mutant.  
Authors : Osipiuk, J.; Makowska-Grzyska, M.; Kwon, K.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2010-11-02  
Resolution : 1.92 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

<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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.14.4.dev1  
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) : 2.14.4.dev1

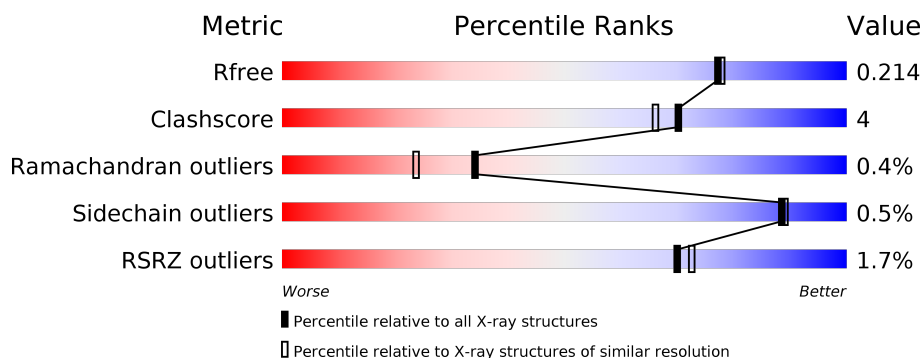
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.92 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	7937 (1.94-1.90)
Clashscore	141614	8644 (1.94-1.90)
Ramachandran outliers	138981	8530 (1.94-1.90)
Sidechain outliers	138945	8530 (1.94-1.90)
RSRZ outliers	127900	7793 (1.94-1.90)

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 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  $\leq 5\%$ . 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	415	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div></div> </div> <div>•</div> </div>
1	B	415	<div> <div>2%</div> <div> <div></div> <div>91%</div> <div>6%</div> <div></div> </div> <div>•</div> </div>
1	C	415	<div> <div>%</div> <div> <div></div> <div>89%</div> <div>6%</div> <div></div> </div> <div>•</div> </div>
1	D	415	<div> <div>2%</div> <div> <div></div> <div>88%</div> <div>8%</div> <div></div> </div> <div>• •</div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 13941 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 Serine hydroxymethyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	402	Total	C	N	O	S	0	11	0
			3172	2017	532	612	11			
1	B	404	Total	C	N	O	S	0	8	0
			3172	2016	532	613	11			
1	C	398	Total	C	N	O	S	0	22	0
			3201	2044	535	611	11			
1	D	402	Total	C	N	O	S	0	17	0
			3212	2046	537	618	11			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	EXPRESSION TAG	UNP Q5HE87
A	-1	ASN	-	EXPRESSION TAG	UNP Q5HE87
A	0	ALA	-	EXPRESSION TAG	UNP Q5HE87
A	95	PRO	SER	ENGINEERED MUTATION	UNP Q5HE87
B	-2	SER	-	EXPRESSION TAG	UNP Q5HE87
B	-1	ASN	-	EXPRESSION TAG	UNP Q5HE87
B	0	ALA	-	EXPRESSION TAG	UNP Q5HE87
B	95	PRO	SER	ENGINEERED MUTATION	UNP Q5HE87
C	-2	SER	-	EXPRESSION TAG	UNP Q5HE87
C	-1	ASN	-	EXPRESSION TAG	UNP Q5HE87
C	0	ALA	-	EXPRESSION TAG	UNP Q5HE87
C	95	PRO	SER	ENGINEERED MUTATION	UNP Q5HE87
D	-2	SER	-	EXPRESSION TAG	UNP Q5HE87
D	-1	ASN	-	EXPRESSION TAG	UNP Q5HE87
D	0	ALA	-	EXPRESSION TAG	UNP Q5HE87
D	95	PRO	SER	ENGINEERED MUTATION	UNP Q5HE87

- Molecule 2 is CITRIC ACID (three-letter code: CIT) (formula: C<sub>6</sub>H<sub>8</sub>O<sub>7</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	B	1	Total	C	O	0	0
			13	6	7		
2	C	1	Total	C	O	0	0
			13	6	7		
2	D	1	Total	C	O	0	0
			13	6	7		
2	D	1	Total	C	O	0	0
			13	6	7		

- Molecule 3 is GLYCINE (three-letter code: GLY) (formula: C<sub>2</sub>H<sub>5</sub>NO<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			5	2	1	2		
3	B	1	Total	C	N	O	0	0
			5	2	1	2		
3	B	1	Total	C	N	O	0	0
			5	2	1	2		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	2	2		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	C	O	0	0
			4	2	2		

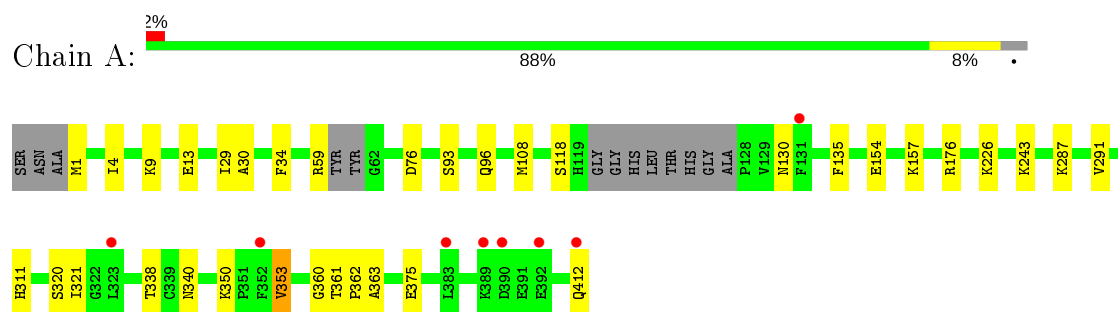
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	247	Total	O	0	0
			247	247		
5	B	289	Total	O	0	0
			289	289		
5	C	279	Total	O	0	0
			279	279		
5	D	242	Total	O	0	0
			242	242		

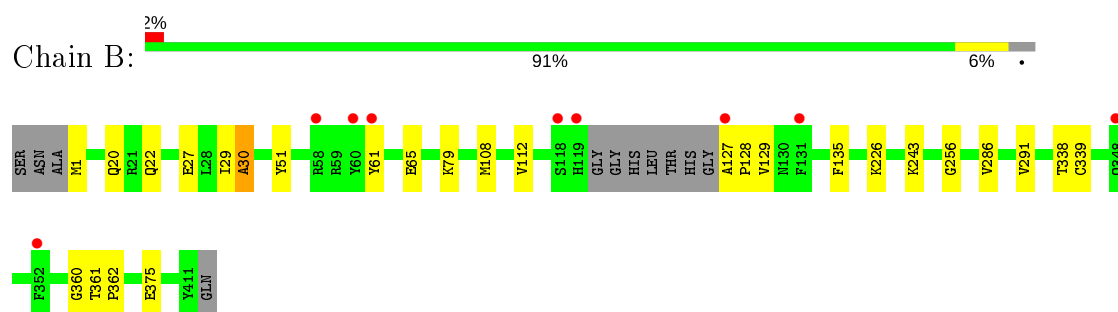
### 3 Residue-property plots [i](#)

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.

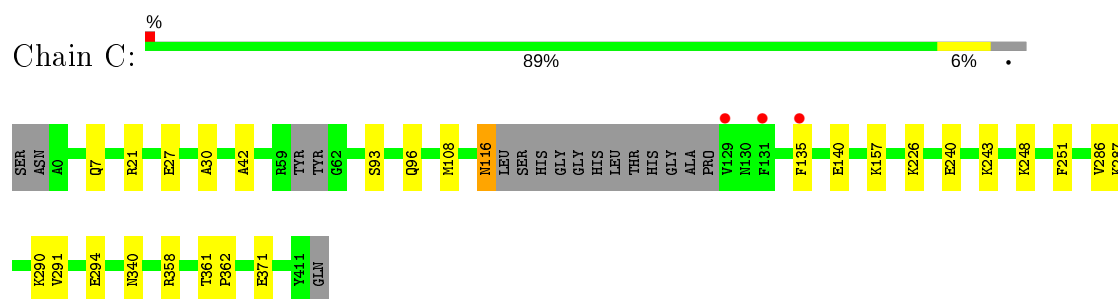
- Molecule 1: Serine hydroxymethyltransferase



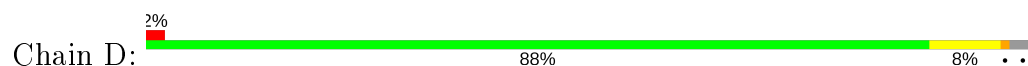
- Molecule 1: Serine hydroxymethyltransferase

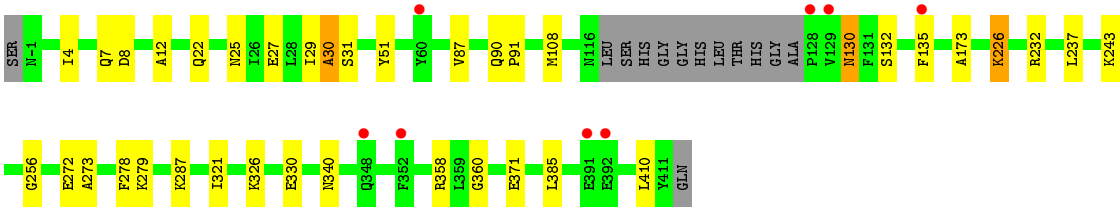


- Molecule 1: Serine hydroxymethyltransferase



- Molecule 1: Serine hydroxymethyltransferase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.08 Å 86.53 Å 301.81 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.70 – 1.92 47.57 – 1.92	Depositor EDS
% Data completeness (in resolution range)	97.8 (47.70-1.92) 97.7 (47.57-1.92)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.52 (at 1.92 Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.166 , 0.208 0.173 , 0.214	Depositor DCC
$R_{free}$ test set	6815 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.5	Xtriage
Anisotropy	0.592	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	13941	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.59% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

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

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, CIT

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	0.83	0/3257	0.76	4/4388 (0.1%)
1	B	0.82	0/3251	0.75	0/4386
1	C	0.84	0/3318	0.77	1/4466 (0.0%)
1	D	0.81	0/3318	0.75	0/4474
All	All	0.83	0/13144	0.76	5/17714 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	59	ARG	NE-CZ-NH2	-5.69	117.45	120.30
1	A	176	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	A	59	ARG	NE-CZ-NH1	5.54	123.07	120.30
1	A	76	ASP	CB-CG-OD2	-5.47	113.37	118.30
1	C	358	ARG	NE-CZ-NH1	-5.27	117.67	120.30

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	3172	0	3207	19	0
1	B	3172	0	3186	22	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3201	0	3281	22	1
1	D	3212	0	3253	30	0
2	A	13	0	5	1	0
2	B	39	0	15	8	0
2	C	13	0	5	1	0
2	D	39	0	15	5	0
3	A	5	0	2	0	0
3	B	10	0	4	0	0
4	C	8	0	12	0	0
5	A	247	0	0	1	1
5	B	289	0	0	7	0
5	C	279	0	0	5	0
5	D	242	0	0	4	0
All	All	13941	0	12985	93	1

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 4.

All (93) 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)
2:B:506:CIT:H41	2:B:506:CIT:O2	1.51	0.96
1:C:290:LYS:HE3	5:C:819:HOH:O	1.82	0.78
2:D:507:CIT:O5	5:D:800:HOH:O	2.06	0.74
1:D:4:ILE:HG22	1:D:12:ALA:HB2	1.69	0.72
1:B:286:VAL:HG13	5:B:481:HOH:O	1.88	0.72
1:D:25[A]:ASN:ND2	5:D:774:HOH:O	2.23	0.71
1:D:173:ALA:HB3	1:D:358[B]:ARG:CZ	2.22	0.69
1:C:21:ARG:NH1	1:C:27:GLU:OE2	2.26	0.69
1:B:22:GLN:O	1:B:27[B]:GLU:OE2	2.11	0.68
1:C:240:GLU:OE1	1:C:243[A]:LYS:HD2	1.95	0.67
1:B:29:ILE:HD11	1:B:338:THR:HG22	1.77	0.66
1:C:286:VAL:HG13	5:C:460:HOH:O	1.99	0.63
1:D:29:ILE:HG23	1:D:340:ASN:HD21	1.64	0.62
2:B:501:CIT:O1	5:B:424:HOH:O	2.16	0.61
1:D:226:LYS:NZ	2:D:507:CIT:O3	2.31	0.61
1:A:9:LYS:HE2	1:A:13[A]:GLU:OE2	2.00	0.61
1:C:243[A]:LYS:NZ	2:C:508:CIT:O7	2.33	0.61
1:A:154[B]:GLU:OE2	1:A:157[B]:LYS:HD2	2.02	0.60
1:B:129:VAL:HG23	5:B:728:HOH:O	2.01	0.60
1:A:118:SER:HB3	1:A:130:ASN:ND2	2.18	0.58

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:256:GLY:HA3	2:B:506:CIT:H42	1.84	0.58
1:A:243:LYS:HE2	2:A:505:CIT:O7	2.07	0.55
1:B:20:GLN:NE2	5:B:644:HOH:O	2.40	0.54
1:A:108:MET:SD	1:A:135:PHE:HB3	2.48	0.54
1:D:31:SER:HB3	1:D:226:LYS:HB3	1.90	0.53
1:D:287[B]:LYS:HD2	1:D:371:GLU:OE2	2.09	0.52
1:D:51:TYR:HE1	2:D:503:CIT:H41	1.74	0.52
1:C:287[B]:LYS:HD2	5:C:818:HOH:O	2.09	0.51
1:D:22:GLN:O	1:D:27[A]:GLU:OE2	2.28	0.51
1:B:1:MET:CE	5:B:547:HOH:O	2.58	0.50
1:C:157[A]:LYS:HD2	5:C:1054:HOH:O	2.11	0.50
1:D:232:ARG:NH1	2:D:507:CIT:O4	2.45	0.50
1:C:42:ALA:HB2	1:D:4:ILE:HD11	1.95	0.48
1:B:1:MET:HE3	5:B:547:HOH:O	2.14	0.48
1:C:93:SER:OG	1:C:96:GLN:NE2	2.46	0.48
1:D:108:MET:HE1	1:D:135:PHE:CD2	2.49	0.48
1:D:243:LYS:HD2	5:D:1055:HOH:O	2.12	0.47
1:D:130:ASN:C	1:D:130:ASN:OD1	2.53	0.47
1:D:30:ALA:O	1:D:360:GLY:HA3	2.15	0.47
1:B:127:ALA:HB1	1:B:128:PRO:CD	2.44	0.47
1:D:256:GLY:HA3	2:D:503:CIT:H22	1.97	0.46
1:D:279:LYS:HE3	5:D:455:HOH:O	2.14	0.46
1:D:87:VAL:HG22	1:D:237:LEU:HD22	1.98	0.46
1:A:287[A]:LYS:HG3	5:A:539:HOH:O	2.16	0.46
1:B:30:ALA:O	1:B:360:GLY:HA3	2.16	0.46
1:B:79:LYS:NZ	2:B:502:CIT:O5	2.49	0.46
1:B:361:THR:N	1:B:362:PRO:CD	2.79	0.46
1:C:7[B]:GLN:HE21	1:C:7[B]:GLN:HB2	1.46	0.46
1:B:361:THR:N	1:B:362:PRO:HD3	2.31	0.45
1:D:7[B]:GLN:HG3	1:D:8:ASP:HB2	1.99	0.45
1:B:243:LYS:HE2	2:B:502:CIT:O7	2.17	0.45
1:D:273:ALA:HA	1:D:278:PHE:CG	2.53	0.45
1:C:108:MET:HE2	1:D:135:PHE:CE2	2.53	0.44
1:A:30:ALA:O	1:A:360:GLY:HA3	2.17	0.44
1:A:29[B]:ILE:HD12	1:A:340:ASN:ND2	2.33	0.44
1:A:311:HIS:HB2	1:A:362:PRO:HD3	1.99	0.44
1:C:116:ASN:HB3	1:C:140:GLU:OE2	2.18	0.43
1:B:51:TYR:HE1	2:B:506:CIT:H22	1.84	0.43
1:D:410:LEU:HD23	1:D:410:LEU:HA	1.91	0.43
1:C:240:GLU:OE1	1:C:243[A]:LYS:CD	2.64	0.43
1:C:361:THR:N	1:C:362:PRO:CD	2.82	0.43

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:29:ILE:HG23	1:D:340:ASN:ND2	2.33	0.43
1:D:326:LYS:O	1:D:330:GLU:HG3	2.19	0.43
1:A:29[B]:ILE:HD11	1:A:338:THR:HG22	2.01	0.42
1:C:108:MET:SD	1:C:135:PHE:HB3	2.59	0.42
1:D:4:ILE:HD13	1:D:4:ILE:HG21	1.64	0.42
1:A:320:SER:O	1:A:321[A]:ILE:HD13	2.19	0.42
1:B:108:MET:HE1	1:B:135:PHE:CD1	2.54	0.42
1:C:287[B]:LYS:CD	5:C:818:HOH:O	2.66	0.42
1:C:291:VAL:HG21	1:C:371[B]:GLU:HG2	2.00	0.42
1:D:130:ASN:OD1	1:D:132:SER:N	2.52	0.42
1:A:34:PHE:CE2	1:A:363:ALA:HB1	2.55	0.42
1:A:108:MET:HE3	1:B:135:PHE:CE1	2.55	0.41
1:A:361:THR:N	1:A:362:PRO:CD	2.83	0.41
1:B:51:TYR:HE1	2:B:506:CIT:C2	2.33	0.41
1:D:4:ILE:HG22	1:D:4:ILE:O	2.18	0.41
1:B:51:TYR:CD2	1:B:61:TYR:CD2	3.08	0.41
1:C:7[B]:GLN:OE1	1:D:272:GLU:OE2	2.38	0.41
1:A:291:VAL:CG1	1:A:375[B]:GLU:HG2	2.50	0.41
1:D:90:GLN:N	1:D:91:PRO:CD	2.83	0.41
1:C:248:LYS:HA	1:C:251:PHE:O	2.20	0.41
2:B:506:CIT:O7	2:B:506:CIT:O3	2.32	0.41
1:A:1:MET:HE3	1:A:4:ILE:HG21	2.03	0.41
1:A:93:SER:OG	1:A:96:GLN:NE2	2.54	0.41
1:A:350:LYS:O	1:A:353:VAL:HG13	2.20	0.40
1:B:1:MET:HE2	5:B:547:HOH:O	2.20	0.40
1:D:321:ILE:HD11	1:D:385:LEU:HB3	2.03	0.40
1:B:29:ILE:HD13	1:B:339:CYS:HA	2.03	0.40
1:B:291:VAL:HG13	1:B:375[A]:GLU:HG3	2.04	0.40

All (1) 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)
1:C:294[B]:GLU:OE2	5:A:859:HOH:O[3_555]	2.13	0.07

## 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	407/415 (98%)	397 (98%)	9 (2%)	1 (0%)	47	38
1	B	408/415 (98%)	397 (97%)	9 (2%)	2 (0%)	29	18
1	C	414/415 (100%)	402 (97%)	10 (2%)	2 (0%)	29	18
1	D	415/415 (100%)	402 (97%)	11 (3%)	2 (0%)	29	18
All	All	1644/1660 (99%)	1598 (97%)	39 (2%)	7 (0%)	34	24

All (7) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	226	LYS
1	B	30	ALA
1	C	226	LYS
1	D	30	ALA
1	D	226	LYS
1	B	226	LYS
1	C	30	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	340/337 (101%)	338 (99%)	2 (1%)	86	86
1	B	338/337 (100%)	335 (99%)	3 (1%)	78	78
1	C	346/337 (103%)	344 (99%)	2 (1%)	86	86

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	345/337 (102%)	344 (100%)	1 (0%)	92	93
All	All	1369/1348 (102%)	1361 (99%)	8 (1%)	88	86

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	353	VAL
1	A	412	GLN
1	B	65[A]	GLU
1	B	65[B]	GLU
1	B	112	VAL
1	C	116	ASN
1	C	340	ASN
1	D	130	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	96	GLN
1	B	7	GLN
1	B	20	GLN
1	B	25	ASN
1	C	96	GLN
1	C	116	ASN
1	D	162	HIS
1	D	340	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

13 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	CIT	B	502	-	3,12,12	1.15	0	3,17,17	2.53	2 (66%)
4	EDO	C	513	-	3,3,3	0.92	0	2,2,2	0.39	0
4	EDO	C	512	-	3,3,3	0.95	0	2,2,2	0.71	0
2	CIT	B	501	-	3,12,12	2.92	2 (66%)	3,17,17	3.26	3 (100%)
2	CIT	C	508	-	3,12,12	1.51	0	3,17,17	3.89	2 (66%)
2	CIT	B	506	-	3,12,12	4.19	2 (66%)	3,17,17	6.06	3 (100%)
3	GLY	B	509	-	1,4,4	0.06	0	0,4,4	0.00	-
3	GLY	A	510	-	1,4,4	0.09	0	0,4,4	0.00	-
2	CIT	D	504	-	3,12,12	1.10	0	3,17,17	2.32	1 (33%)
2	CIT	A	505	-	3,12,12	0.75	0	3,17,17	1.75	1 (33%)
3	GLY	B	511	-	1,4,4	0.09	0	0,4,4	0.00	-
2	CIT	D	503	-	3,12,12	3.39	1 (33%)	3,17,17	2.78	1 (33%)
2	CIT	D	507	-	3,12,12	2.07	1 (33%)	3,17,17	4.85	2 (66%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CIT	B	502	-	-	0/6/16/16	-
4	EDO	C	513	-	-	0/1/1/1	-
4	EDO	C	512	-	-	1/1/1/1	-
2	CIT	B	501	-	-	3/6/16/16	-
2	CIT	C	508	-	-	0/6/16/16	-
2	CIT	B	506	-	-	3/6/16/16	-
3	GLY	B	509	-	-	0/0/2/2	-
3	GLY	A	510	-	-	0/0/2/2	-
2	CIT	D	504	-	-	0/6/16/16	-

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	CIT	A	505	-	-	0/6/16/16	-
3	GLY	B	511	-	-	0/0/2/2	-
2	CIT	D	503	-	-	1/6/16/16	-
2	CIT	D	507	-	-	3/6/16/16	-

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	506	CIT	C4-C3	-5.78	1.46	1.54
2	D	503	CIT	C2-C3	-5.63	1.47	1.54
2	B	501	CIT	C2-C3	-4.43	1.48	1.54
2	B	506	CIT	C2-C3	-4.24	1.48	1.54
2	D	507	CIT	O7-C3	2.66	1.47	1.43
2	B	501	CIT	C4-C3	-2.30	1.51	1.54

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	507	CIT	C3-C4-C5	-8.07	102.06	114.98
2	B	506	CIT	C4-C3-C2	-7.36	89.65	109.33
2	C	508	CIT	C3-C4-C5	-6.25	104.98	114.98
2	B	506	CIT	C3-C2-C1	-6.20	105.06	114.98
2	D	503	CIT	C3-C2-C1	-4.39	107.96	114.98
2	B	501	CIT	C3-C4-C5	-4.33	108.06	114.98
2	B	506	CIT	C3-C4-C5	-4.20	108.27	114.98
2	D	504	CIT	C3-C2-C1	-3.94	108.67	114.98
2	B	502	CIT	C3-C4-C5	-3.59	109.23	114.98
2	A	505	CIT	C3-C2-C1	-2.98	110.22	114.98
2	B	501	CIT	C3-C2-C1	2.76	119.40	114.98
2	B	501	CIT	C4-C3-C2	-2.37	102.99	109.33
2	C	508	CIT	C3-C2-C1	-2.36	111.20	114.98
2	D	507	CIT	C3-C2-C1	-2.21	111.44	114.98
2	B	502	CIT	C3-C2-C1	-2.03	111.73	114.98

There are no chirality outliers.

All (11) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	501	CIT	C1-C2-C3-O7
2	B	501	CIT	C1-C2-C3-C4
2	B	501	CIT	C1-C2-C3-C6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	D	507	CIT	C2-C3-C4-C5
2	D	507	CIT	O7-C3-C4-C5
2	D	507	CIT	C6-C3-C4-C5
2	B	506	CIT	C2-C3-C4-C5
2	B	506	CIT	C6-C3-C4-C5
2	B	506	CIT	O7-C3-C4-C5
2	D	503	CIT	C2-C3-C4-C5
4	C	512	EDO	O1-C1-C2-O2

There are no ring outliers.

7 monomers are involved in 15 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	502	CIT	2	0
2	B	501	CIT	1	0
2	C	508	CIT	1	0
2	B	506	CIT	5	0
2	A	505	CIT	1	0
2	D	503	CIT	2	0
2	D	507	CIT	3	0

## 5.7 Other polymers [i](#)

There are no such residues in this entry.

## 5.8 Polymer linkage issues [i](#)

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<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	402/415 (96%)	-0.01	8 (1%) 65 68	16, 26, 45, 63	0
1	B	404/415 (97%)	-0.17	9 (2%) 62 65	19, 28, 40, 56	0
1	C	398/415 (95%)	-0.20	3 (0%) 86 87	17, 26, 41, 51	1 (0%)
1	D	402/415 (96%)	-0.03	8 (1%) 65 68	18, 30, 46, 58	0
All	All	1606/1660 (96%)	-0.10	28 (1%) 70 72	16, 28, 44, 63	1 (0%)

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	127	ALA	5.6
1	C	129	VAL	5.0
1	B	352	PHE	4.5
1	B	60	TYR	4.1
1	D	352[A]	PHE	3.9
1	D	129	VAL	3.6
1	D	60	TYR	3.4
1	A	412	GLN	3.3
1	A	392	GLU	3.1
1	B	61	TYR	3.0
1	B	58	ARG	2.8
1	D	348[A]	GLN	2.6
1	B	348	GLN	2.5
1	D	135	PHE	2.5
1	A	352	PHE	2.5
1	A	389	LYS	2.4
1	B	119	HIS	2.4
1	D	128	PRO	2.3
1	A	131[A]	PHE	2.3
1	C	131[A]	PHE	2.3
1	C	135	PHE	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	390	ASP	2.3
1	D	391	GLU	2.2
1	B	131[A]	PHE	2.2
1	A	323	LEU	2.2
1	A	383	LEU	2.2
1	B	118	SER	2.2
1	D	392	GLU	2.2

## 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 [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
3	GLY	B	509	5/5	0.80	0.23	55,55,56,56	0
2	CIT	D	507	13/13	0.86	0.16	48,51,59,61	0
2	CIT	D	503	13/13	0.87	0.20	38,40,57,59	0
2	CIT	B	506	13/13	0.87	0.19	37,43,49,56	0
3	GLY	B	511	5/5	0.88	0.17	67,67,68,68	0
2	CIT	B	501	13/13	0.88	0.16	41,45,51,52	0
2	CIT	D	504	13/13	0.88	0.16	37,46,54,57	0
4	EDO	C	512	4/4	0.91	0.21	30,36,38,38	0
2	CIT	C	508	13/13	0.91	0.18	39,48,61,63	0
2	CIT	A	505	13/13	0.92	0.11	28,37,55,58	0
2	CIT	B	502	13/13	0.93	0.15	42,48,60,62	0
3	GLY	A	510	5/5	0.94	0.17	51,52,54,54	0
4	EDO	C	513	4/4	0.94	0.11	31,35,35,36	0

## 6.5 Other polymers

There are no such residues in this entry.