



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 27, 2023 – 07:20 PM EDT

PDB ID : 3GVD  
Title : Crystal Structure of Serine Acetyltransferase CysE from *Yersinia pestis*  
Authors : Kim, Y.; Zhou, M.; Peterson, S.; Anderson, W.F.; Joachimiak, A.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : 2009-03-30  
Resolution : 2.40 Å (reported)

This is a wwPDB X-ray Structure Validation Summary 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 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  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.35  
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) : 2.35

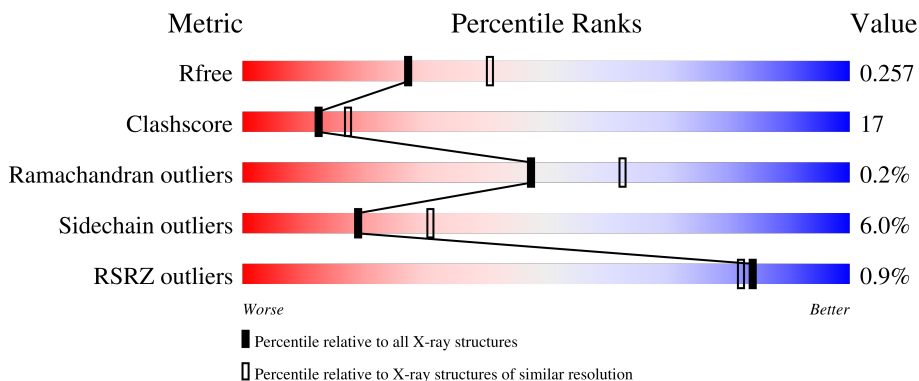
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*


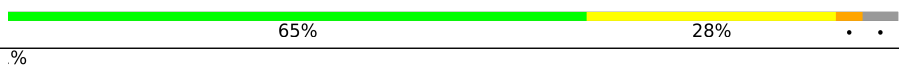
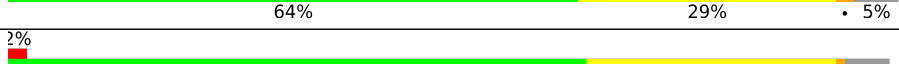
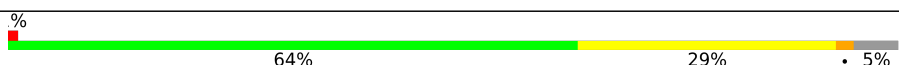

The reported resolution of this entry is 2.40 Å.

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	3907 (2.40-2.40)
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)
RSRZ outliers	127900	3811 (2.40-2.40)

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  $\geq 3$ , 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	276	 70% 26% ..
1	B	276	 65% 28% . .
1	C	276	 64% 29% . 5%
1	D	276	 65% 28% . 5%
1	E	276	 64% 29% . 5%

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	276	 67% 25% • 5%
1	G	276	 % 58% 35% • 5%
1	H	276	 % 64% 28% • 5%
1	I	276	 72% 24% • •
1	J	276	 % 59% 32% • 5%
1	K	276	 % 66% 26% • 6%
1	L	276	 % 66% 28% • 5%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	ACY	A	275	-	-	X	-
6	PG5	C	274	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 25003 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 acetyltransferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	271	Total 2039	C 1289	N 359	O 382	S 9	0	0	0
1	B	265	Total 1987	C 1254	N 353	O 371	S 9	0	0	0
1	C	262	Total 1969	C 1245	N 349	O 366	S 9	0	0	0
1	D	262	Total 1979	C 1250	N 353	O 367	S 9	0	1	0
1	E	263	Total 1984	C 1253	N 353	O 369	S 9	0	1	0
1	F	261	Total 1964	C 1242	N 348	O 365	S 9	0	0	0
1	G	263	Total 1985	C 1253	N 354	O 369	S 9	0	1	0
1	H	261	Total 1961	C 1241	N 347	O 364	S 9	0	0	0
1	I	272	Total 2054	C 1298	N 362	O 384	S 10	0	1	0
1	J	261	Total 1964	C 1242	N 348	O 365	S 9	0	0	0
1	K	260	Total 1956	C 1238	N 346	O 363	S 9	0	0	0
1	L	261	Total 1972	C 1247	N 349	O 366	S 10	0	1	0

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	SER	-	expression tag	UNP Q0WKM4
A	-1	ASN	-	expression tag	UNP Q0WKM4
A	0	ALA	-	expression tag	UNP Q0WKM4
B	-2	SER	-	expression tag	UNP Q0WKM4
B	-1	ASN	-	expression tag	UNP Q0WKM4

*Continued on next page...*

*Continued from previous page...*

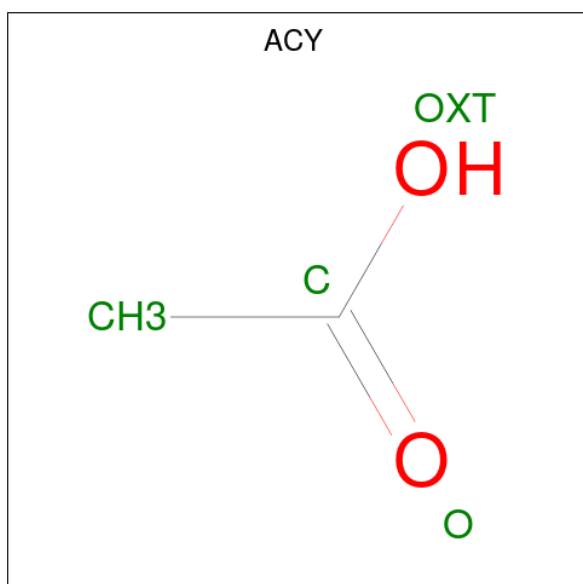
Chain	Residue	Modelled	Actual	Comment	Reference
B	0	ALA	-	expression tag	UNP Q0WKM4
C	-2	SER	-	expression tag	UNP Q0WKM4
C	-1	ASN	-	expression tag	UNP Q0WKM4
C	0	ALA	-	expression tag	UNP Q0WKM4
D	-2	SER	-	expression tag	UNP Q0WKM4
D	-1	ASN	-	expression tag	UNP Q0WKM4
D	0	ALA	-	expression tag	UNP Q0WKM4
E	-2	SER	-	expression tag	UNP Q0WKM4
E	-1	ASN	-	expression tag	UNP Q0WKM4
E	0	ALA	-	expression tag	UNP Q0WKM4
F	-2	SER	-	expression tag	UNP Q0WKM4
F	-1	ASN	-	expression tag	UNP Q0WKM4
F	0	ALA	-	expression tag	UNP Q0WKM4
G	-2	SER	-	expression tag	UNP Q0WKM4
G	-1	ASN	-	expression tag	UNP Q0WKM4
G	0	ALA	-	expression tag	UNP Q0WKM4
H	-2	SER	-	expression tag	UNP Q0WKM4
H	-1	ASN	-	expression tag	UNP Q0WKM4
H	0	ALA	-	expression tag	UNP Q0WKM4
I	-2	SER	-	expression tag	UNP Q0WKM4
I	-1	ASN	-	expression tag	UNP Q0WKM4
I	0	ALA	-	expression tag	UNP Q0WKM4
J	-2	SER	-	expression tag	UNP Q0WKM4
J	-1	ASN	-	expression tag	UNP Q0WKM4
J	0	ALA	-	expression tag	UNP Q0WKM4
K	-2	SER	-	expression tag	UNP Q0WKM4
K	-1	ASN	-	expression tag	UNP Q0WKM4
K	0	ALA	-	expression tag	UNP Q0WKM4
L	-2	SER	-	expression tag	UNP Q0WKM4
L	-1	ASN	-	expression tag	UNP Q0WKM4
L	0	ALA	-	expression tag	UNP Q0WKM4

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



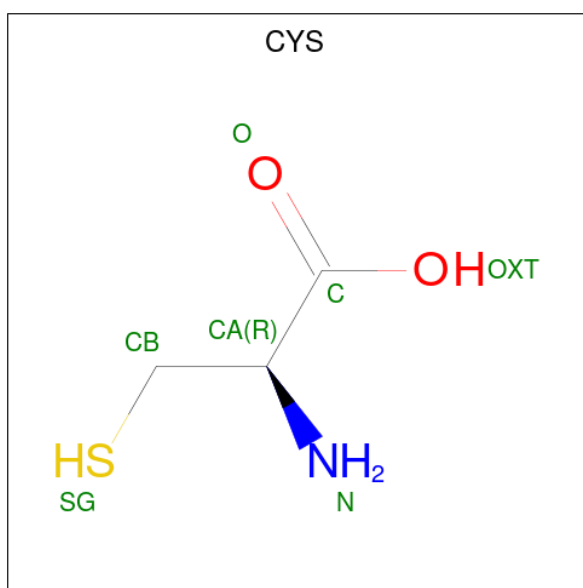
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	F	1	Total O S 5 4 1	0	0
2	H	1	Total O S 5 4 1	0	0
2	K	1	Total O S 5 4 1	0	0
2	L	1	Total O S 5 4 1	0	0

- Molecule 3 is ACETIC ACID (three-letter code: ACY) (formula: C<sub>2</sub>H<sub>4</sub>O<sub>2</sub>).



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

- Molecule 4 is CYSTEINE (three-letter code: CYS) (formula: C<sub>3</sub>H<sub>7</sub>NO<sub>2</sub>S).



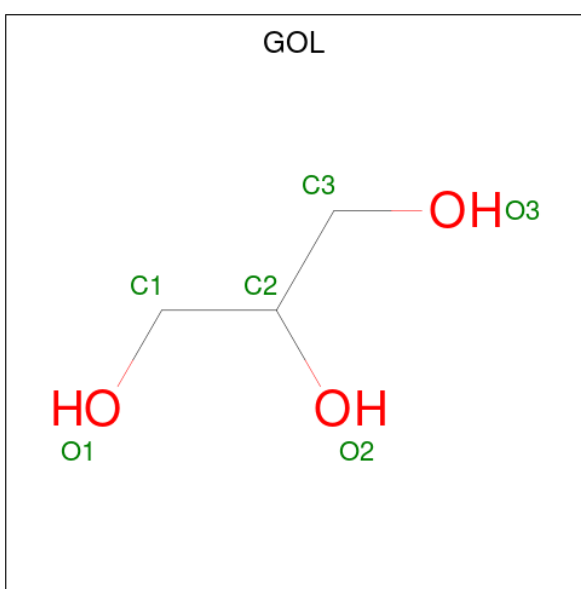
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O S 7 3 1 2 1	0	0
4	A	1	Total C N O S 7 3 1 2 1	0	0
4	B	1	Total C N O S 7 3 1 2 1	0	0
4	D	1	Total C N O S 7 3 1 2 1	0	0
4	D	1	Total C N O S 7 3 1 2 1	0	0
4	E	1	Total C N O S 7 3 1 2 1	0	0
4	G	1	Total C N O S 7 3 1 2 1	0	0
4	H	1	Total C N O S 7 3 1 2 1	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	I	1	Total	C	N	O	S	0	0
			7	3	1	2	1		
4	J	1	Total	C	N	O	S	0	0
			7	3	1	2	1		
4	L	1	Total	C	N	O	S	0	0
			7	3	1	2	1		
4	L	1	Total	C	N	O	S	0	0
			7	3	1	2	1		

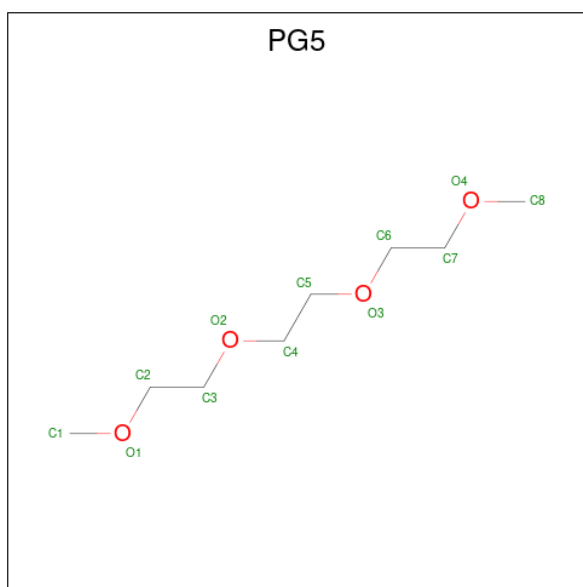
- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	C O	0	0
			6	3 3		
5	E	1	Total	C O	0	0
			6	3 3		
5	F	1	Total	C O	0	0
			6	3 3		
5	H	1	Total	C O	0	0
			6	3 3		
5	J	1	Total	C O	0	0
			6	3 3		

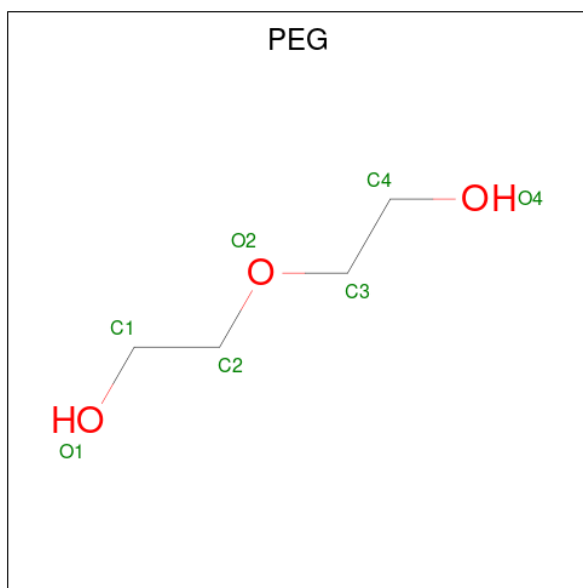
- Molecule 6 is 1-METHOXY-2-[2-(2-METHOXY-ETHOXY)]-ETHANE (three-letter code: PG5) (formula: C<sub>8</sub>H<sub>18</sub>O<sub>4</sub>).





Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total	C	O	0	0
			12	8	4		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula:  $C_4H_{10}O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	L	1	Total	C	O	0	0
			7	4	3		

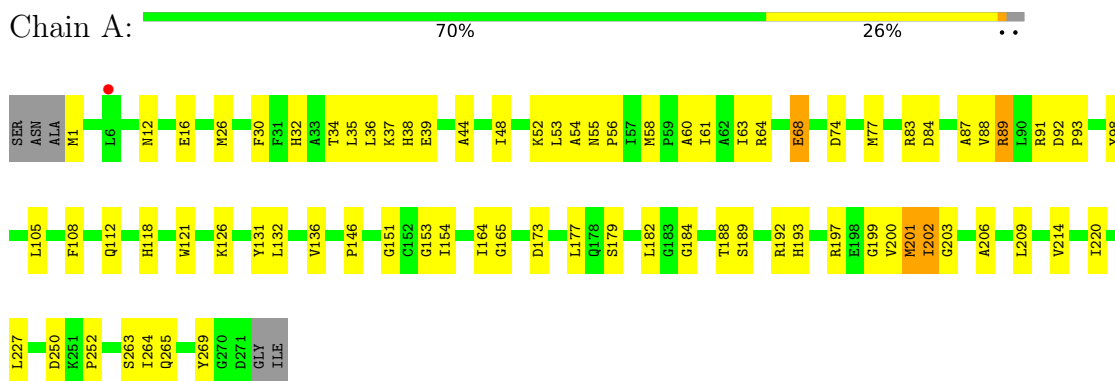
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	100	Total 100	O 100	0	0
8	B	111	Total 111	O 111	0	0
8	C	90	Total 90	O 90	0	0
8	D	73	Total 73	O 73	0	0
8	E	78	Total 78	O 78	0	0
8	F	73	Total 73	O 73	0	0
8	G	71	Total 71	O 71	0	0
8	H	69	Total 69	O 69	0	0
8	I	113	Total 113	O 113	0	0
8	J	82	Total 82	O 82	0	0
8	K	87	Total 87	O 87	0	0
8	L	68	Total 68	O 68	0	0

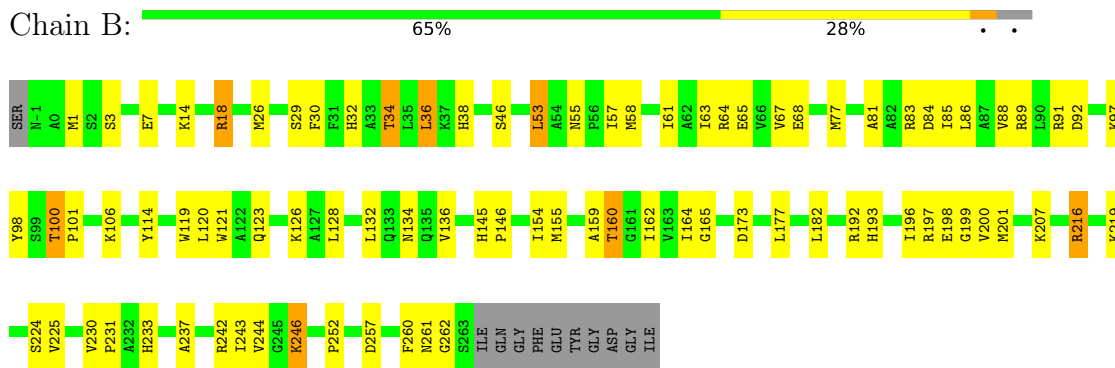
### 3 Residue-property plots [i](#)

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.

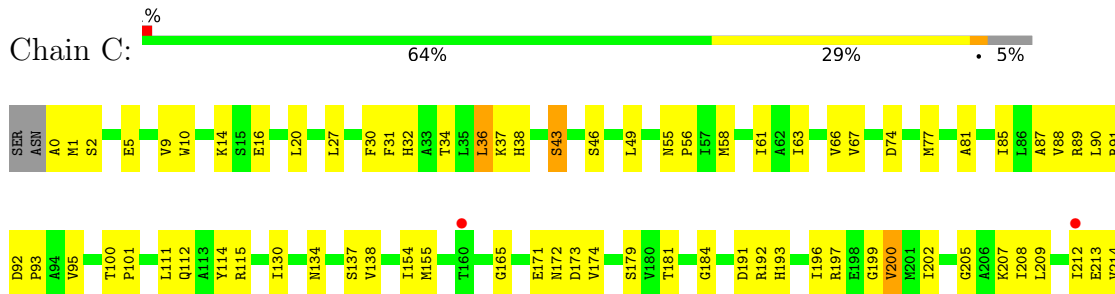
- Molecule 1: Serine acetyltransferase



- Molecule 1: Serine acetyltransferase

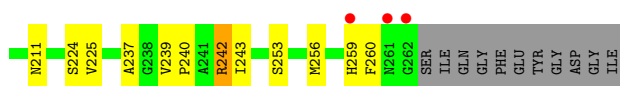
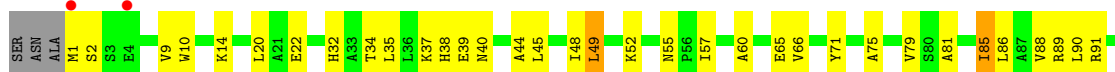


- Molecule 1: Serine acetyltransferase

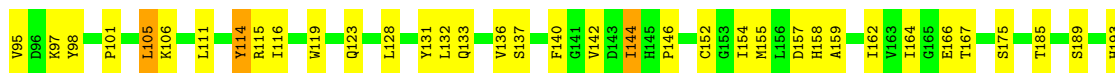
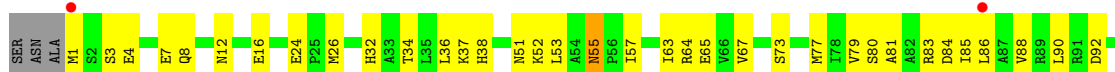




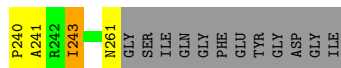
- Molecule 1: Serine acetyltransferase



- Molecule 1: Serine acetyltransferase

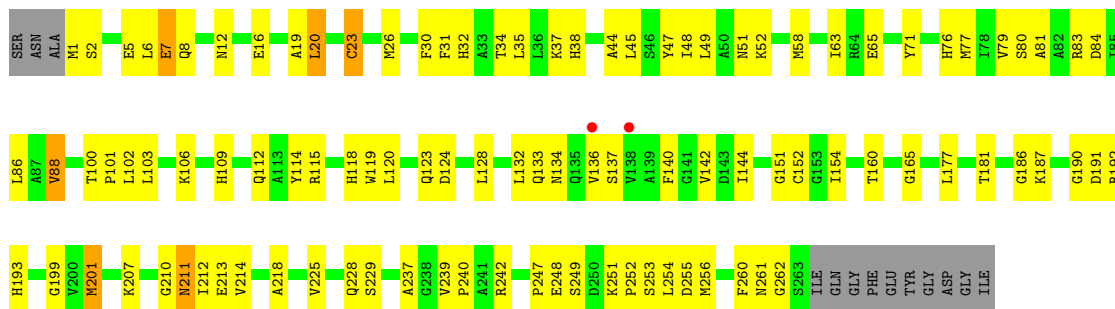


- Molecule 1: Serine acetyltransferase

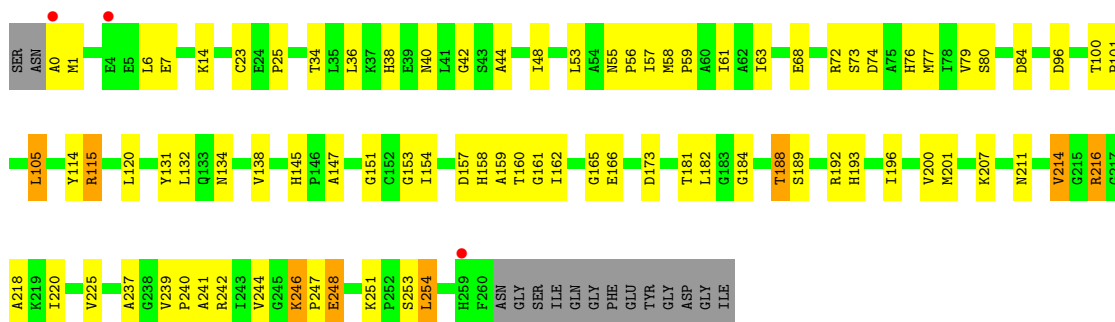


- Molecule 1: Serine acetyltransferase

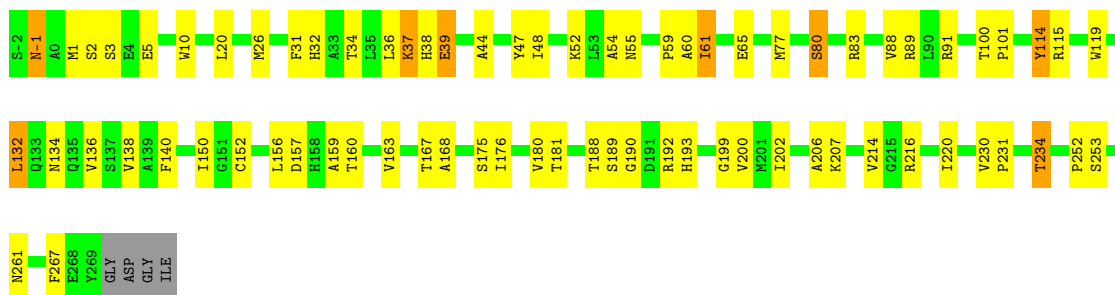




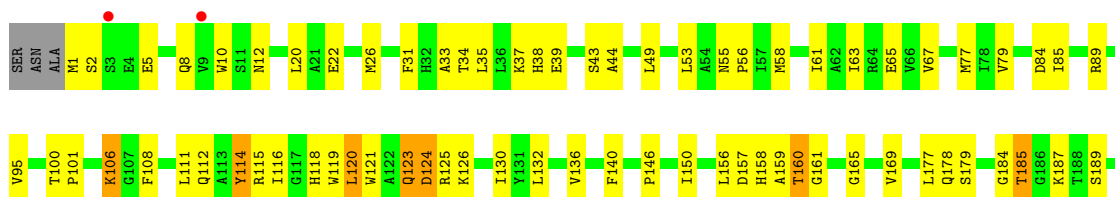
• Molecule 1: Serine acetyltransferase

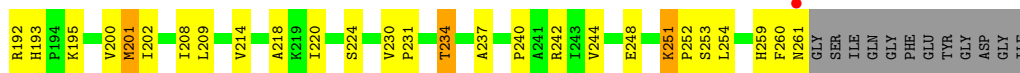


• Molecule 1: Serine acetyltransferase

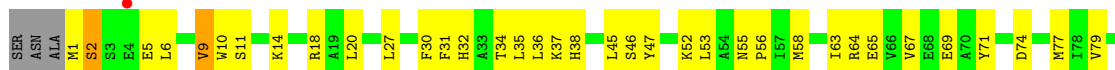


• Molecule 1: Serine acetyltransferase

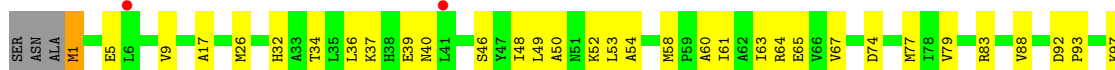




• Molecule 1: Serine acetyltransferase



• Molecule 1: Serine acetyltransferase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.47Å 81.23Å 139.41Å 97.07° 95.32° 94.72°	Depositor
Resolution (Å)	45.99 – 2.40 45.98 – 2.40	Depositor EDS
% Data completeness (in resolution range)	97.0 (45.99-2.40) 97.1 (45.98-2.40)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.33 (at 2.39Å)	Xtrriage
Refinement program	PHENIX (phenix.refine)	Depositor
R, $R_{free}$	0.199 , 0.265 0.191 , 0.257	Depositor DCC
$R_{free}$ test set	5663 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	32.7	Xtrriage
Anisotropy	0.754	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 35.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.36$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	25003	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	46.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.09% 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: PG5, ACY, GOL, PEG, SO4

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	0.42	0/2078	0.57	0/2818
1	B	0.44	0/2024	0.59	0/2746
1	C	0.43	0/2006	0.59	0/2722
1	D	0.40	0/2016	0.57	0/2734
1	E	0.41	0/2022	0.58	0/2743
1	F	0.39	0/2001	0.57	0/2715
1	G	0.39	0/2022	0.56	0/2742
1	H	0.36	0/1998	0.56	0/2711
1	I	0.42	0/2093	0.59	0/2838
1	J	0.39	0/2001	0.57	0/2715
1	K	0.38	0/1993	0.60	0/2704
1	L	0.38	0/2009	0.57	0/2725
All	All	0.40	0/24263	0.58	0/32913

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 [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	2039	0	2054	57	1

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	B	1987	0	2012	79	0
1	C	1969	0	1998	83	0
1	D	1979	0	2008	78	0
1	E	1984	0	2007	84	0
1	F	1964	0	1993	64	0
1	G	1985	0	2013	87	0
1	H	1961	0	1992	68	0
1	I	2054	0	2071	63	0
1	J	1964	0	1993	82	0
1	K	1956	0	1987	60	0
1	L	1972	0	2001	76	0
2	A	5	0	0	1	0
2	F	5	0	0	0	0
2	H	5	0	0	0	0
2	K	5	0	0	0	0
2	L	5	0	0	1	0
3	A	4	0	3	2	0
3	C	4	0	3	0	0
3	J	8	0	6	1	0
4	A	14	0	8	1	0
4	B	7	0	4	0	0
4	D	14	0	8	0	0
4	E	7	0	4	0	0
4	G	7	0	4	0	0
4	H	7	0	4	0	0
4	I	7	0	4	1	0
4	J	7	0	4	0	0
4	L	14	0	8	2	0
5	B	6	0	8	2	0
5	E	6	0	8	2	0
5	F	6	0	8	0	0
5	H	6	0	8	1	0
5	J	6	0	8	0	0
6	C	12	0	18	9	0
7	L	7	0	10	3	0
8	A	100	0	0	1	0
8	B	111	0	0	5	0
8	C	90	0	0	4	0
8	D	73	0	0	6	0
8	E	78	0	0	6	0
8	F	73	0	0	3	0
8	G	71	0	0	3	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	H	69	0	0	3	1
8	I	113	0	0	3	0
8	J	82	0	0	5	0
8	K	87	0	0	6	0
8	L	68	0	0	3	0
All	All	25003	0	24257	809	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 17.

The worst 5 of 809 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:D:57:ILE:HD11	1:F:30:PHE:HB2	1.34	1.07
1:E:114:TYR:CD1	1:E:144:ILE:HD11	1.92	1.05
1:L:120:LEU:HD22	1:L:125:ARG:HD2	1.42	1.01
1:D:91:ARG:HB2	1:D:155:MET:HE3	1.40	1.00
1:G:34:THR:HG23	1:G:35:LEU:HG	1.44	0.96

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:A:269:TYR:OH	8:H:729:HOH:O[1_444]	2.19	0.01

## 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	269/276 (98%)	256 (95%)	13 (5%)	0	100   100
1	B	263/276 (95%)	253 (96%)	9 (3%)	1 (0%)	34   48

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	C	260/276 (94%)	249 (96%)	11 (4%)	0	100	100
1	D	261/276 (95%)	249 (95%)	11 (4%)	1 (0%)	34	48
1	E	262/276 (95%)	253 (97%)	9 (3%)	0	100	100
1	F	259/276 (94%)	246 (95%)	12 (5%)	1 (0%)	34	48
1	G	262/276 (95%)	250 (95%)	11 (4%)	1 (0%)	34	48
1	H	259/276 (94%)	246 (95%)	13 (5%)	0	100	100
1	I	271/276 (98%)	256 (94%)	14 (5%)	1 (0%)	34	48
1	J	259/276 (94%)	246 (95%)	11 (4%)	2 (1%)	19	29
1	K	258/276 (94%)	248 (96%)	10 (4%)	0	100	100
1	L	260/276 (94%)	247 (95%)	13 (5%)	0	100	100
All	All	3143/3312 (95%)	2999 (95%)	137 (4%)	7 (0%)	47	62

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	192	ARG
1	D	2	SER
1	F	198	GLU
1	G	211	ASN
1	J	158	HIS

### 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	212/215 (99%)	201 (95%)	11 (5%)	23	38
1	B	207/215 (96%)	191 (92%)	16 (8%)	13	20
1	C	205/215 (95%)	197 (96%)	8 (4%)	32	50
1	D	206/215 (96%)	195 (95%)	11 (5%)	22	37
1	E	207/215 (96%)	196 (95%)	11 (5%)	22	37
1	F	205/215 (95%)	190 (93%)	15 (7%)	14	22

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	207/215 (96%)	197 (95%)	10 (5%)	25	41
1	H	204/215 (95%)	190 (93%)	14 (7%)	15	25
1	I	214/215 (100%)	204 (95%)	10 (5%)	26	42
1	J	205/215 (95%)	188 (92%)	17 (8%)	11	17
1	K	204/215 (95%)	190 (93%)	14 (7%)	15	25
1	L	206/215 (96%)	193 (94%)	13 (6%)	18	28
All	All	2482/2580 (96%)	2332 (94%)	150 (6%)	19	31

5 of 150 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	J	160	THR
1	L	116	ILE
1	J	201	MET
1	K	177	LEU
1	E	105	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 57 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	38	HIS
1	L	228	GLN
1	H	40	ASN
1	L	193	HIS
1	K	172	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

28 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	SO4	L	274	-	4,4,4	0.17	0	6,6,6	0.29	0
5	GOL	B	274	-	5,5,5	0.31	0	5,5,5	0.48	0
2	SO4	H	274	-	4,4,4	0.15	0	6,6,6	0.28	0
4	CYS	D	501	-	5,6,6	0.94	0	5,7,7	1.83	2 (40%)
4	CYS	A	501	-	5,6,6	1.31	1 (20%)	5,7,7	1.41	2 (40%)
4	CYS	L	276	-	5,6,6	1.09	1 (20%)	5,7,7	1.78	2 (40%)
3	ACY	A	275	-	3,3,3	0.63	0	3,3,3	1.41	0
4	CYS	J	501	-	5,6,6	1.11	1 (20%)	5,7,7	1.77	2 (40%)
6	PG5	C	274	-	11,11,11	0.47	0	10,10,10	1.48	1 (10%)
5	GOL	H	275	-	5,5,5	0.41	0	5,5,5	0.14	0
4	CYS	A	276	-	5,6,6	1.01	0	5,7,7	1.58	2 (40%)
4	CYS	L	501	-	5,6,6	1.04	0	5,7,7	1.45	1 (20%)
2	SO4	F	274	-	4,4,4	0.18	0	6,6,6	0.22	0
3	ACY	J	276	-	3,3,3	0.77	0	3,3,3	0.82	0
5	GOL	F	275	-	5,5,5	0.37	0	5,5,5	0.24	0
2	SO4	K	274	-	4,4,4	0.13	0	6,6,6	0.15	0
3	ACY	J	275	-	3,3,3	0.70	0	3,3,3	1.15	0
5	GOL	J	274	-	5,5,5	0.40	0	5,5,5	0.35	0
7	PEG	L	275	-	6,6,6	0.50	0	5,5,5	1.54	0
4	CYS	G	501	-	5,6,6	1.07	1 (20%)	5,7,7	1.80	1 (20%)
4	CYS	H	501	-	5,6,6	0.99	1 (20%)	5,7,7	1.74	2 (40%)
5	GOL	E	274	-	5,5,5	0.43	0	5,5,5	0.44	0
4	CYS	D	274	-	5,6,6	1.14	0	5,7,7	1.53	2 (40%)
2	SO4	A	274	-	4,4,4	0.17	0	6,6,6	0.46	0
4	CYS	E	501	-	5,6,6	0.93	0	5,7,7	1.31	1 (20%)
3	ACY	C	275	-	3,3,3	0.89	0	3,3,3	0.58	0
4	CYS	B	501	-	5,6,6	1.08	1 (20%)	5,7,7	1.23	0
4	CYS	I	501	-	5,6,6	1.16	1 (20%)	5,7,7	1.93	2 (40%)

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
5	GOL	B	274	-	-	2/4/4/4	-
4	CYS	D	501	-	-	0/6/6/6	-
4	CYS	A	501	-	-	0/6/6/6	-
4	CYS	L	276	-	-	0/6/6/6	-
4	CYS	J	501	-	-	1/6/6/6	-
6	PG5	C	274	-	-	5/9/9/9	-
5	GOL	H	275	-	-	4/4/4/4	-
4	CYS	A	276	-	-	0/6/6/6	-
4	CYS	L	501	-	-	0/6/6/6	-
5	GOL	F	275	-	-	4/4/4/4	-
5	GOL	J	274	-	-	4/4/4/4	-
7	PEG	L	275	-	-	2/4/4/4	-
4	CYS	G	501	-	-	0/6/6/6	-
4	CYS	H	501	-	-	0/6/6/6	-
5	GOL	E	274	-	-	0/4/4/4	-
4	CYS	D	274	-	-	0/6/6/6	-
4	CYS	E	501	-	-	0/6/6/6	-
4	CYS	B	501	-	-	0/6/6/6	-
4	CYS	I	501	-	-	0/6/6/6	-

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	J	501	CYS	OXT-C	-2.34	1.22	1.30
4	A	501	CYS	OXT-C	-2.26	1.23	1.30
4	I	501	CYS	OXT-C	-2.22	1.23	1.30
4	B	501	CYS	OXT-C	-2.14	1.23	1.30
4	G	501	CYS	OXT-C	-2.09	1.23	1.30

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	I	501	CYS	OXT-C-O	-3.46	116.22	124.09
4	J	501	CYS	OXT-C-O	-3.22	116.79	124.09
4	G	501	CYS	CB-CA-C	-2.90	107.00	109.89
4	H	501	CYS	OXT-C-O	-2.79	117.75	124.09
4	D	501	CYS	OXT-C-O	-2.74	117.87	124.09

There are no chirality outliers.

5 of 22 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	B	274	GOL	O1-C1-C2-C3
5	F	275	GOL	O1-C1-C2-C3
5	H	275	GOL	O1-C1-C2-C3
5	J	274	GOL	O1-C1-C2-C3
6	C	274	PG5	O1-C2-C3-O2

There are no ring outliers.

13 monomers are involved in 26 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	L	274	SO4	1	0
5	B	274	GOL	2	0
4	L	276	CYS	1	0
3	A	275	ACY	2	0
6	C	274	PG5	9	0
5	H	275	GOL	1	0
4	A	276	CYS	1	0
4	L	501	CYS	1	0
3	J	276	ACY	1	0
7	L	275	PEG	3	0
5	E	274	GOL	2	0
2	A	274	SO4	1	0
4	I	501	CYS	1	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 [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	271/276 (98%)	-0.30	1 (0%) 92 91	29, 41, 65, 89	0
1	B	265/276 (96%)	-0.39	0 100 100	29, 40, 55, 84	0
1	C	262/276 (94%)	-0.34	3 (1%) 80 79	26, 39, 59, 97	0
1	D	262/276 (94%)	-0.23	6 (2%) 60 58	33, 43, 68, 104	0
1	E	263/276 (95%)	-0.14	2 (0%) 86 84	31, 43, 68, 102	0
1	F	261/276 (94%)	-0.24	1 (0%) 92 91	31, 42, 71, 97	0
1	G	263/276 (95%)	-0.24	2 (0%) 86 84	33, 47, 67, 93	0
1	H	261/276 (94%)	-0.16	3 (1%) 80 79	33, 48, 72, 120	0
1	I	272/276 (98%)	-0.27	0 100 100	31, 43, 59, 83	0
1	J	261/276 (94%)	-0.20	3 (1%) 80 79	32, 45, 68, 88	0
1	K	260/276 (94%)	-0.03	3 (1%) 79 77	36, 49, 72, 97	0
1	L	261/276 (94%)	-0.17	3 (1%) 80 79	32, 47, 72, 90	0
All	All	3162/3312 (95%)	-0.23	27 (0%) 84 82	26, 44, 68, 120	0

The worst 5 of 27 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	J	9	VAL	3.2
1	D	261	ASN	3.1
1	E	1	MET	3.0
1	H	0	ALA	3.0
1	D	4	GLU	2.9

### 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
6	PG5	C	274	12/12	0.71	0.25	63,65,68,68	0
5	GOL	J	274	6/6	0.74	0.13	66,66,66,66	0
3	ACY	C	275	4/4	0.75	0.21	47,50,50,50	0
7	PEG	L	275	7/7	0.83	0.21	40,42,43,43	0
2	SO4	K	274	5/5	0.84	0.40	200,200,200,200	0
3	ACY	J	276	4/4	0.87	0.15	55,55,56,57	0
5	GOL	B	274	6/6	0.90	0.20	43,44,44,45	0
3	ACY	J	275	4/4	0.92	0.16	51,51,52,53	0
4	CYS	E	501	7/7	0.92	0.14	32,34,35,37	0
4	CYS	L	501	7/7	0.93	0.12	39,40,41,43	0
5	GOL	F	275	6/6	0.93	0.13	55,55,56,57	0
5	GOL	H	275	6/6	0.93	0.17	54,55,55,55	0
4	CYS	J	501	7/7	0.94	0.16	35,36,39,47	0
4	CYS	L	276	7/7	0.94	0.11	40,42,43,45	0
2	SO4	L	274	5/5	0.95	0.18	53,55,55,56	0
4	CYS	A	276	7/7	0.95	0.14	35,35,37,38	0
5	GOL	E	274	6/6	0.95	0.12	49,49,49,49	0
4	CYS	D	274	7/7	0.95	0.12	27,30,31,32	0
3	ACY	A	275	4/4	0.96	0.17	43,45,45,45	0
4	CYS	B	501	7/7	0.97	0.14	37,39,39,39	0
4	CYS	D	501	7/7	0.97	0.11	37,38,39,40	0
4	CYS	A	501	7/7	0.97	0.08	28,29,31,43	0
2	SO4	A	274	5/5	0.97	0.22	50,52,53,53	0
4	CYS	H	501	7/7	0.97	0.10	42,48,50,50	0
4	CYS	G	501	7/7	0.98	0.09	35,37,38,39	0
2	SO4	F	274	5/5	0.98	0.16	47,48,49,50	0
2	SO4	H	274	5/5	0.99	0.14	49,50,51,52	0
4	CYS	I	501	7/7	0.99	0.08	31,33,34,39	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.