



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 12, 2017 – 06:33 AM EST

PDB ID : 4OC9  
Title : 2.35 Angstrom resolution crystal structure of putative O-acetylhomoserine (thiol)-lyase (metY) from Campylobacter jejuni subsp. jejuni NCTC 11168 with N'-Pyridoxyl-Lysine-5'-Monophosphate at position 205  
Authors : Halavaty, A.S.; Brunzelle, J.S.; Wawrzak, Z.; Onopriyenko, O.; Savchenko, A.; Anderson, W.F.; Center for Structural Genomics of Infectious Diseases (CSGID)  
Deposited on : unknown  
Resolution : 2.35 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : rb-20030345

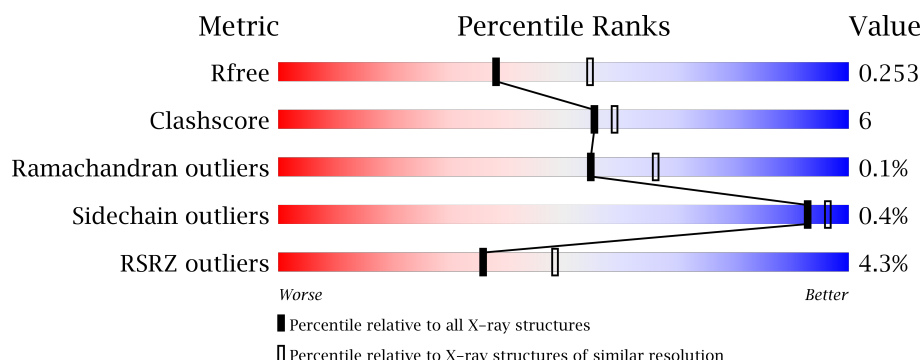
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.35 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))
$R_{free}$	100719	1522 (2.38-2.34)
Clashscore	112137	1626 (2.38-2.34)
Ramachandran outliers	110173	1605 (2.38-2.34)
Sidechain outliers	110143	1606 (2.38-2.34)
RSRZ outliers	101464	1528 (2.38-2.34)

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	424	<div> <div>5%</div> <div>88%</div> <div>10%</div> <div>..</div> </div>
1	B	424	<div> <div>4%</div> <div>87%</div> <div>10%</div> <div>..</div> </div>
1	C	424	<div> <div>4%</div> <div>91%</div> <div>8%</div> <div>..</div> </div>
1	D	424	<div> <div>4%</div> <div>91%</div> <div>7%</div> <div>.</div> </div>
1	E	424	<div> <div>4%</div> <div>91%</div> <div>7%</div> <div>..</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain	
1	F	424		•
1	G	424		•
1	H	424		•
1	I	424		• 5%
1	J	424		• •
1	K	424		• •
1	L	424		• •
1	M	424		• •
1	N	424		• •
1	O	424		• •
1	P	424		•

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	GOL	O	501	-	-	-	X

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 55287 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.

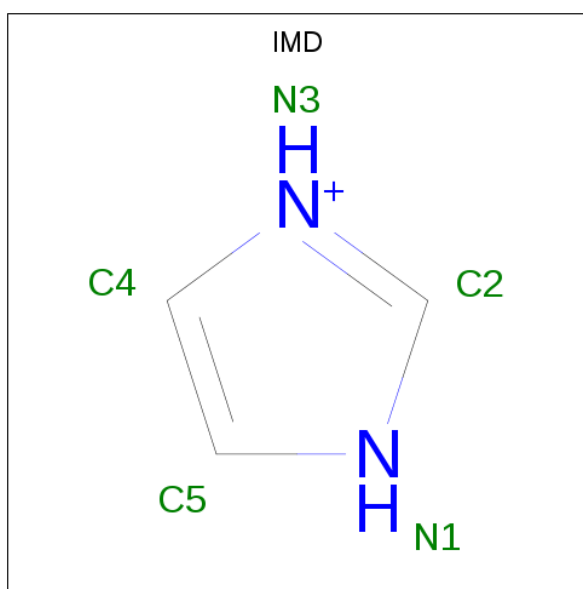
- Molecule 1 is a protein called Putative O-acetylhomoserine (Thiol)-lyase.

Mol	Chain	Residues	Atoms							ZeroOcc	AltConf	Trace
1	A	421	Total	C	N	O	P	S	Se	0	5	0
			3335	2112	570	648	1	3	1			
1	B	421	Total	C	N	O	P	S	Se	0	4	0
			3329	2108	571	645	1	3	1			
1	C	421	Total	C	N	O	P	S	Se	0	3	0
			3311	2097	566	643	1	3	1			
1	D	422	Total	C	N	O	P	S	Se	0	3	0
			3327	2107	570	645	1	3	1			
1	E	421	Total	C	N	O	P	S	Se	0	3	0
			3320	2102	568	645	1	3	1			
1	F	422	Total	C	N	O	P	S	Se	0	2	0
			3318	2101	568	644	1	3	1			
1	G	422	Total	C	N	O	P	S	Se	0	3	0
			3326	2105	570	646	1	3	1			
1	H	422	Total	C	N	O	P	S	Se	0	3	0
			3326	2107	569	645	1	3	1			
1	I	403	Total	C	N	O	P	S	Se	0	1	0
			3166	2012	541	608	1	3	1			
1	J	407	Total	C	N	O	P	S	Se	0	1	0
			3195	2029	547	614	1	3	1			
1	K	406	Total	C	N	O	P	S	Se	0	0	0
			3184	2023	545	611	1	3	1			
1	L	407	Total	C	N	O	P	S	Se	0	3	0
			3211	2039	549	618	1	3	1			
1	M	407	Total	C	N	O	P	S	Se	0	1	0
			3197	2032	547	613	1	3	1			
1	N	405	Total	C	N	O	P	S	Se	0	1	0
			3188	2025	545	613	1	3	1			
1	O	405	Total	C	N	O	P	S	Se	0	1	0
			3189	2026	545	613	1	3	1			
1	P	412	Total	C	N	O	P	S	Se	0	2	0
			3252	2063	559	625	1	3	1			

There are 16 discrepancies between the modelled and reference sequences:

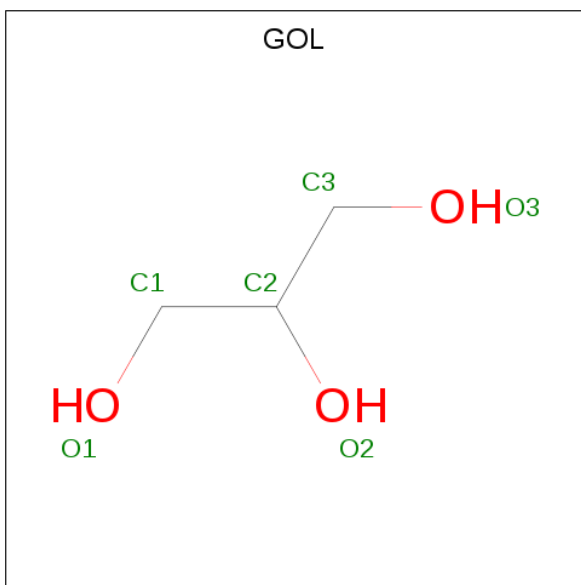
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
B	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
C	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
D	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
E	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
F	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
G	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
H	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
I	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
J	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
K	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
L	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
M	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
N	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
O	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4
P	0	GLY	-	EXPRESSION TAG	UNP Q0P7Q4

- Molecule 2 is IMIDAZOLE (three-letter code: IMD) (formula:  $C_3H_5N_2$ ).



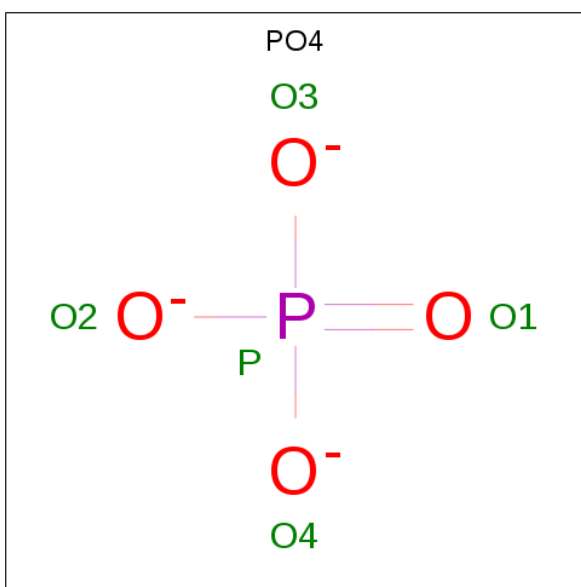
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C N 5 3 2	0	0
2	D	1	Total C N 5 3 2	0	0

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	1
			12	6	6		
3	O	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula:  $\text{O}_4\text{P}$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	C	1	Total	O	P	0	0
			5	4	1		

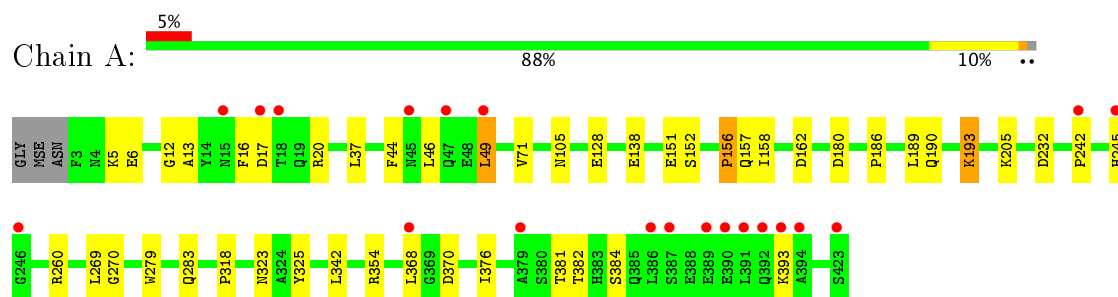
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	187	Total 194	O 194	0	7
5	B	214	Total 220	O 220	0	6
5	C	206	Total 212	O 212	0	6
5	D	186	Total 195	O 195	0	9
5	E	210	Total 217	O 217	0	8
5	F	195	Total 197	O 197	0	2
5	G	199	Total 207	O 207	0	8
5	H	187	Total 191	O 191	0	4
5	I	192	Total 195	O 195	0	3
5	J	185	Total 192	O 192	0	7
5	K	177	Total 181	O 181	0	5
5	L	154	Total 159	O 159	0	6
5	M	161	Total 166	O 166	0	5
5	N	162	Total 168	O 168	0	6
5	O	188	Total 193	O 193	0	5
5	P	180	Total 187	O 187	0	7

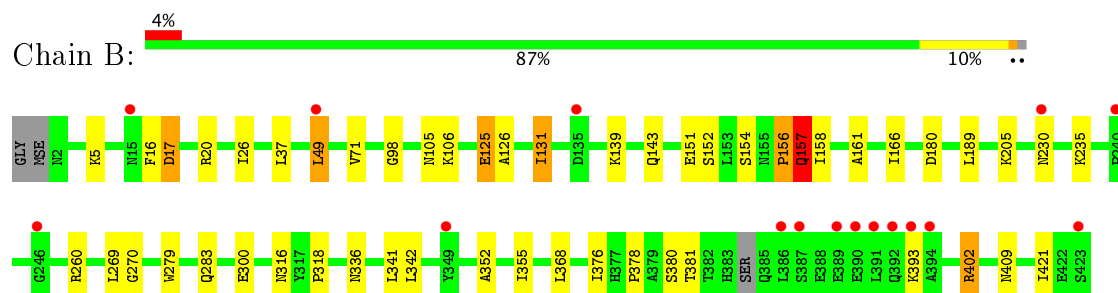
### 3 Residue-property plots [i](#)

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.

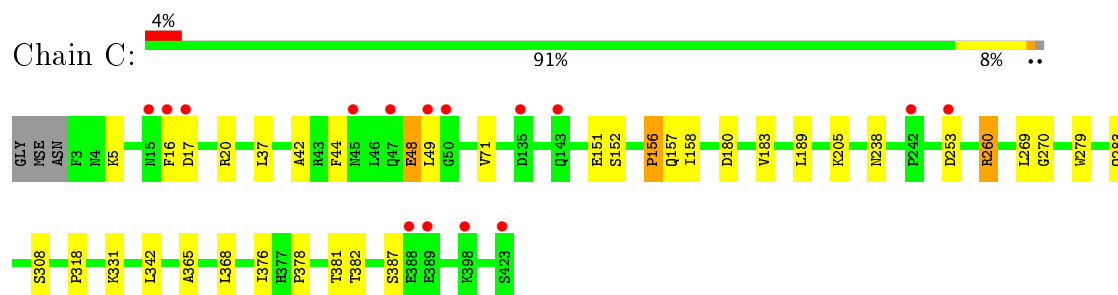
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



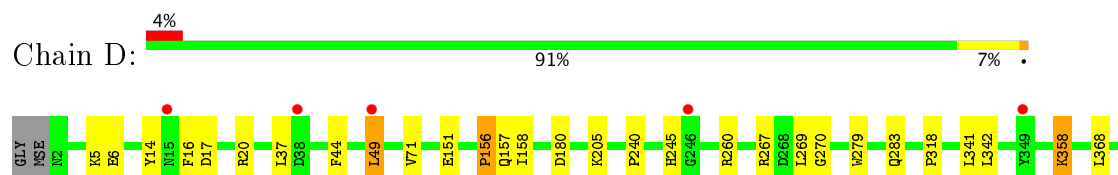
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



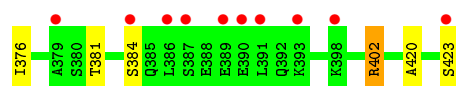
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



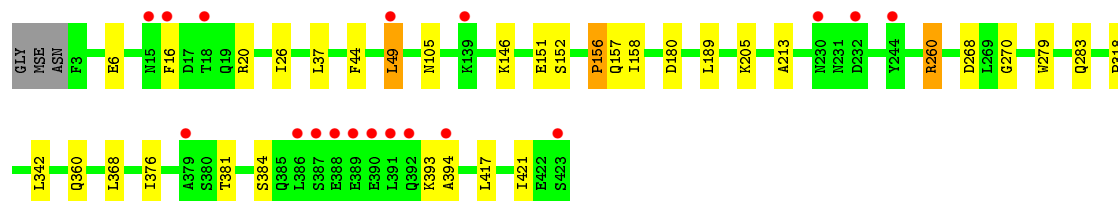
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



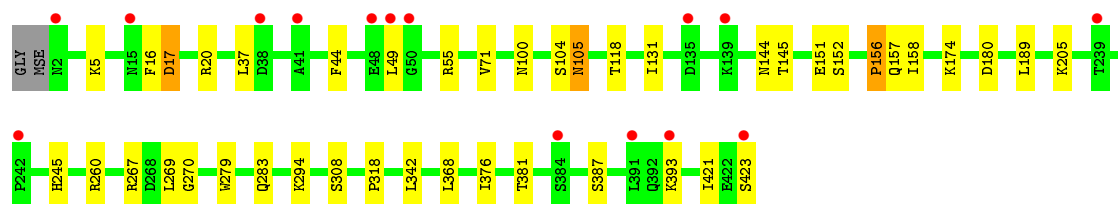
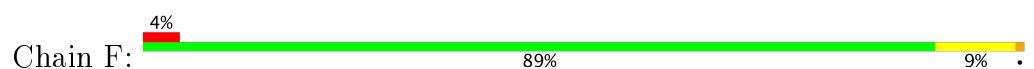




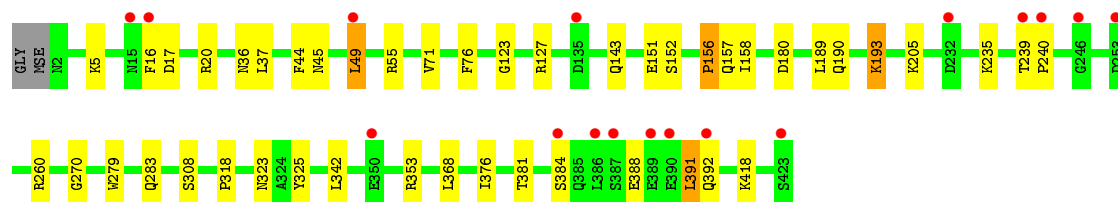
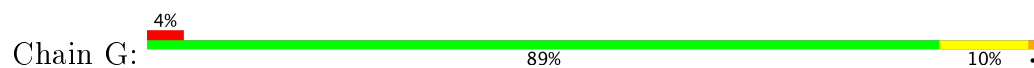
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



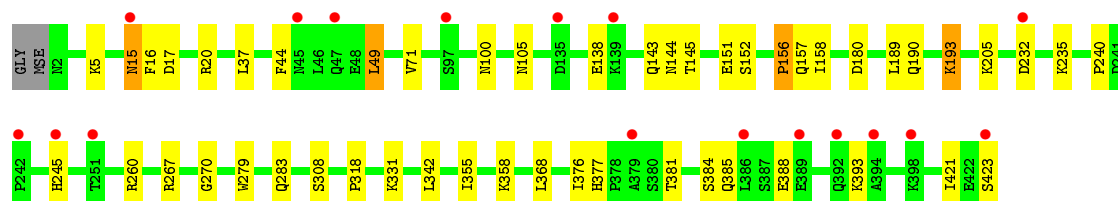
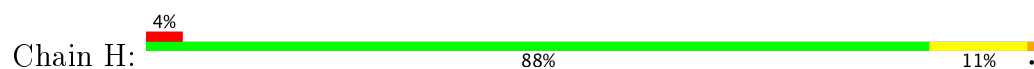
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



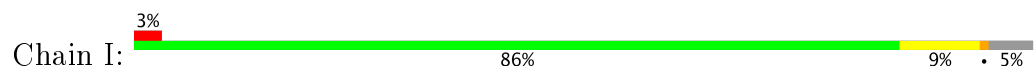
- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

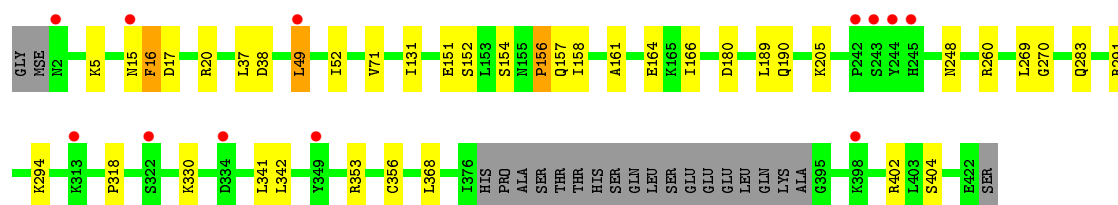


- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

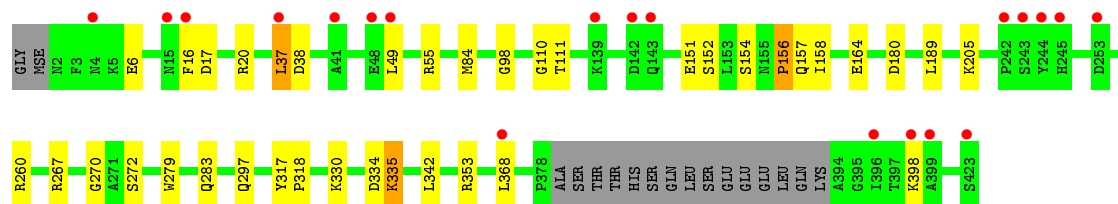
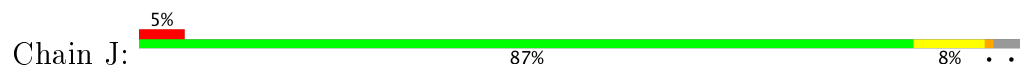


- Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

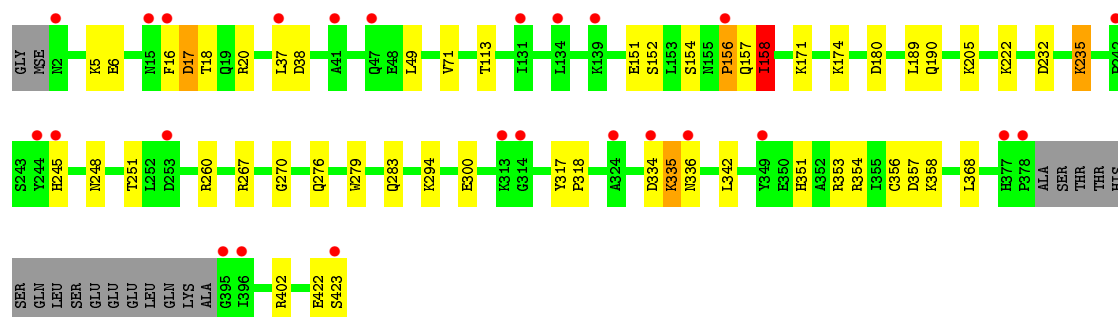
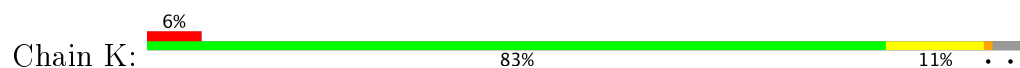




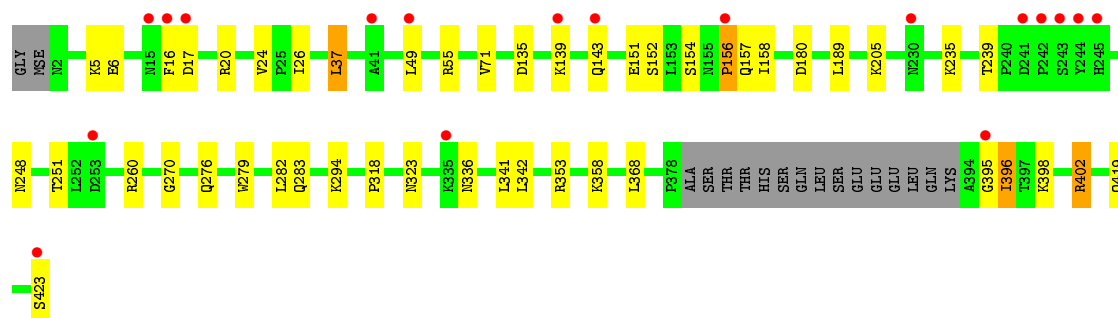
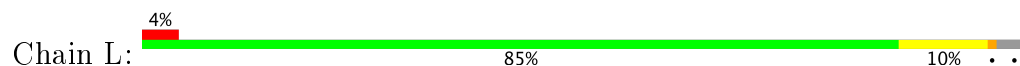
• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



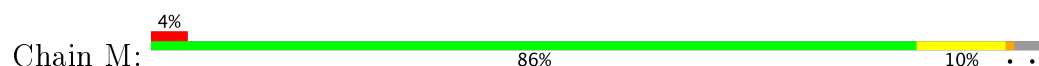
• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

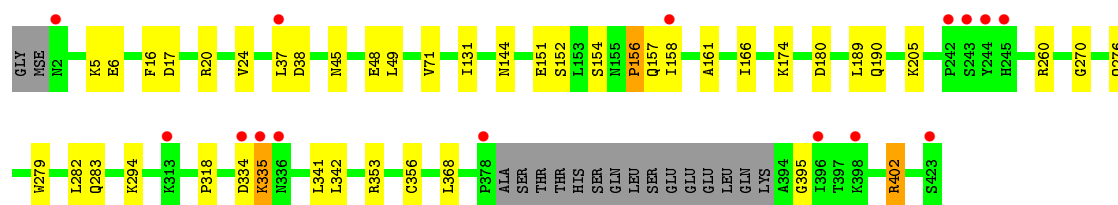


• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

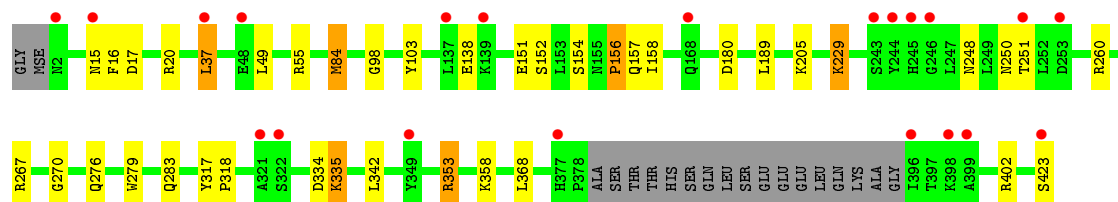
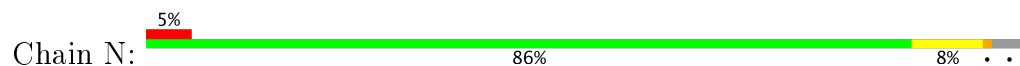


• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase

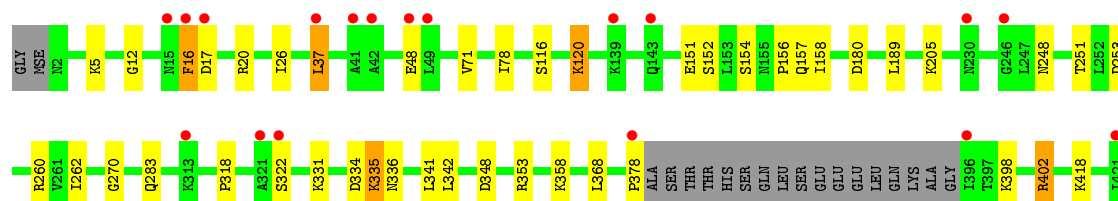
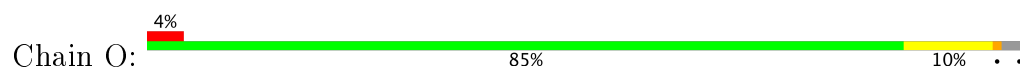




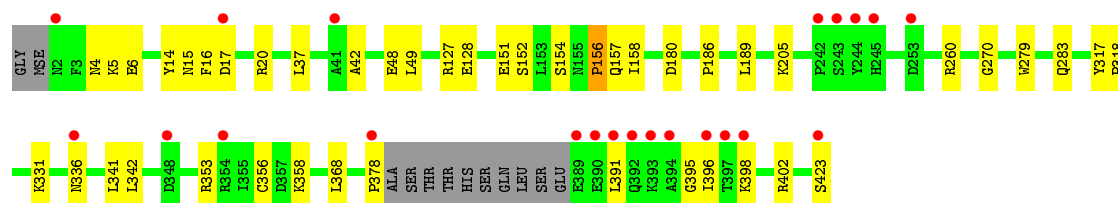
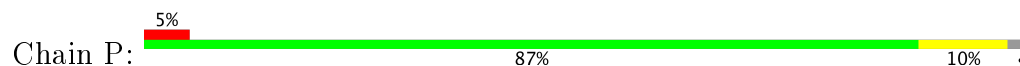
• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



• Molecule 1: Putative O-acetylhomoserine (Thiol)-lyase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	60.43 Å   149.79 Å   186.59 Å 100.58°   92.47°   90.10°	Depositor
Resolution (Å)	29.65 – 2.35 29.63 – 2.35	Depositor EDS
% Data completeness (in resolution range)	73.4 (29.65-2.35) 69.0 (29.63-2.35)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.81 (at 2.36 Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.202 , 0.249 0.209 , 0.253	Depositor DCC
$R_{free}$ test set	9880 reflections (5.30%)	DCC
Wilson B-factor (Å <sup>2</sup> )	18.9	Xtriage
Anisotropy	0.438	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 45.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.046 for h,-k,-l	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	55287	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 39.62 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 3.0867e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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 ⓘ

### 5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, PO4, LLP, IMD

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	2/3370 (0.1%)	0.98	15/4564 (0.3%)
1	B	0.62	2/3363 (0.1%)	0.94	7/4552 (0.2%)
1	C	0.60	0/3353	0.94	8/4541 (0.2%)
1	D	0.60	0/3362	0.97	10/4552 (0.2%)
1	E	0.60	0/3355	0.92	5/4544 (0.1%)
1	F	0.64	1/3353 (0.0%)	0.95	8/4541 (0.2%)
1	G	0.58	0/3361	0.94	8/4552 (0.2%)
1	H	0.59	0/3361	0.93	6/4552 (0.1%)
1	I	0.59	0/3197	0.96	9/4329 (0.2%)
1	J	0.57	0/3228	0.95	9/4371 (0.2%)
1	K	0.59	0/3218	0.96	10/4359 (0.2%)
1	L	0.60	2/3244 (0.1%)	1.33	10/4393 (0.2%)
1	M	0.63	5/3231 (0.2%)	1.34	12/4377 (0.3%)
1	N	0.60	1/3222 (0.0%)	0.99	13/4365 (0.3%)
1	O	0.63	3/3223 (0.1%)	0.99	15/4366 (0.3%)
1	P	0.60	1/3286 (0.0%)	0.96	7/4449 (0.2%)
All	All	0.60	17/52727 (0.0%)	1.01	152/71407 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	N	0	1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	F	308	SER	CA-CB	9.15	1.66	1.52
1	M	260	ARG	CZ-NH1	-8.92	1.21	1.33

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	L	260	ARG	CZ-NH1	-8.68	1.21	1.33
1	A	128	GLU	CD-OE1	-7.53	1.17	1.25
1	M	48	GLU	CD-OE1	-7.14	1.17	1.25

The worst 5 of 152 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	L	260	ARG	NE-CZ-NH1	-44.91	97.85	120.30
1	M	260	ARG	NE-CZ-NH1	-44.42	98.09	120.30
1	M	260	ARG	NE-CZ-NH2	41.49	141.04	120.30
1	L	260	ARG	NE-CZ-NH2	41.17	140.89	120.30
1	N	55	ARG	NE-CZ-NH1	-11.49	114.56	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	N	353	ARG	Sidechain

## 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	3335	0	3297	37	0
1	B	3329	0	3294	48	0
1	C	3311	0	3277	30	0
1	D	3327	0	3295	38	0
1	E	3320	0	3281	32	0
1	F	3318	0	3283	38	0
1	G	3326	0	3288	43	0
1	H	3326	0	3293	60	0
1	I	3166	0	3144	41	0
1	J	3195	0	3168	41	0
1	K	3184	0	3158	53	0
1	L	3211	0	3181	62	0
1	M	3197	0	3173	38	0
1	N	3188	0	3158	43	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	O	3189	0	3160	39	0
1	P	3252	0	3224	40	0
2	A	5	0	5	0	0
2	D	5	0	5	0	0
3	A	6	0	8	1	0
3	B	12	0	16	1	0
3	O	6	0	8	0	0
4	C	5	0	0	0	0
5	A	194	0	0	8	0
5	B	220	0	0	17	0
5	C	212	0	0	8	0
5	D	195	0	0	11	0
5	E	217	0	0	11	0
5	F	197	0	0	8	0
5	G	207	0	0	10	0
5	H	191	0	0	11	0
5	I	195	0	0	9	0
5	J	192	0	0	12	0
5	K	181	0	0	19	0
5	L	159	0	0	5	0
5	M	166	0	0	7	0
5	N	168	0	0	5	0
5	O	193	0	0	15	0
5	P	187	0	0	7	0
All	All	55287	0	51716	617	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 6.

The worst 5 of 617 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:K:158:ILE:CD1	1:K:318:PRO:HD3	1.79	1.11
1:J:84:MSE:HE1	1:J:110:GLY:C	1.70	1.11
1:H:235:LYS:CE	1:L:139:LYS:HE3	1.81	1.10
1:K:158:ILE:HD11	1:K:318:PRO:CD	1.81	1.09
1:N:84:MSE:HE3	1:N:205:LLP:H5'2	1.30	1.09

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone

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	423/424 (100%)	407 (96%)	16 (4%)	0	100	100
1	B	420/424 (99%)	406 (97%)	13 (3%)	1 (0%)	51	61
1	C	421/424 (99%)	406 (96%)	14 (3%)	1 (0%)	51	61
1	D	422/424 (100%)	408 (97%)	14 (3%)	0	100	100
1	E	421/424 (99%)	405 (96%)	16 (4%)	0	100	100
1	F	421/424 (99%)	404 (96%)	16 (4%)	1 (0%)	51	61
1	G	422/424 (100%)	406 (96%)	16 (4%)	0	100	100
1	H	422/424 (100%)	407 (96%)	15 (4%)	0	100	100
1	I	399/424 (94%)	386 (97%)	12 (3%)	1 (0%)	44	53
1	J	403/424 (95%)	389 (96%)	14 (4%)	0	100	100
1	K	401/424 (95%)	387 (96%)	13 (3%)	1 (0%)	51	61
1	L	405/424 (96%)	387 (96%)	17 (4%)	1 (0%)	51	61
1	M	403/424 (95%)	389 (96%)	13 (3%)	1 (0%)	51	61
1	N	401/424 (95%)	385 (96%)	16 (4%)	0	100	100
1	O	401/424 (95%)	385 (96%)	15 (4%)	1 (0%)	51	61
1	P	409/424 (96%)	393 (96%)	16 (4%)	0	100	100
All	All	6594/6784 (97%)	6350 (96%)	236 (4%)	8 (0%)	55	66

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	I	16	PHE
1	L	396	ILE
1	O	16	PHE
1	C	17	ASP
1	B	17	ASP



### 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	360/355 (101%)	359 (100%)	1 (0%)	94	97
1	B	359/355 (101%)	357 (99%)	2 (1%)	89	94
1	C	358/355 (101%)	357 (100%)	1 (0%)	94	97
1	D	359/355 (101%)	358 (100%)	1 (0%)	94	97
1	E	358/355 (101%)	357 (100%)	1 (0%)	94	97
1	F	358/355 (101%)	356 (99%)	2 (1%)	89	94
1	G	359/355 (101%)	358 (100%)	1 (0%)	94	97
1	H	359/355 (101%)	357 (99%)	2 (1%)	89	94
1	I	340/355 (96%)	339 (100%)	1 (0%)	94	97
1	J	343/355 (97%)	342 (100%)	1 (0%)	94	97
1	K	342/355 (96%)	340 (99%)	2 (1%)	89	94
1	L	345/355 (97%)	344 (100%)	1 (0%)	94	97
1	M	343/355 (97%)	342 (100%)	1 (0%)	94	97
1	N	343/355 (97%)	341 (99%)	2 (1%)	89	94
1	O	343/355 (97%)	342 (100%)	1 (0%)	94	97
1	P	349/355 (98%)	347 (99%)	2 (1%)	89	94
All	All	5618/5680 (99%)	5596 (100%)	22 (0%)	93	96

5 of 22 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	15	ASN
1	J	156	PRO
1	P	15	ASN
1	H	156	PRO
1	I	156	PRO

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	H	190	GLN
1	O	190	GLN
1	J	190	GLN
1	F	105	ASN
1	M	181	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

16 non-standard protein/DNA/RNA residues 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$
1	LLP	A	205	1	24,24,25	1.89	3 (12%)	28,32,34	1.73	7 (25%)
1	LLP	B	205	1	24,24,25	1.74	3 (12%)	28,32,34	1.65	5 (17%)
1	LLP	C	205	1	24,24,25	1.84	3 (12%)	28,32,34	1.73	4 (14%)
1	LLP	D	205	1	24,24,25	1.87	4 (16%)	28,32,34	1.63	6 (21%)
1	LLP	E	205	1	24,24,25	1.92	3 (12%)	28,32,34	1.58	6 (21%)
1	LLP	F	205	1	24,24,25	1.72	3 (12%)	28,32,34	1.77	7 (25%)
1	LLP	G	205	1	24,24,25	1.51	3 (12%)	28,32,34	1.75	9 (32%)
1	LLP	H	205	1	24,24,25	1.88	4 (16%)	28,32,34	1.73	6 (21%)
1	LLP	I	205	1	24,24,25	1.98	4 (16%)	28,32,34	3.00	12 (42%)
1	LLP	J	205	1	24,24,25	2.17	2 (8%)	28,32,34	1.75	7 (25%)
1	LLP	K	205	1	24,24,25	1.66	3 (12%)	28,32,34	1.70	5 (17%)
1	LLP	L	205	1	24,24,25	2.24	4 (16%)	28,32,34	2.99	12 (42%)
1	LLP	M	205	1	24,24,25	1.71	3 (12%)	28,32,34	1.64	6 (21%)
1	LLP	N	205	1	24,24,25	2.06	4 (16%)	28,32,34	1.74	6 (21%)
1	LLP	O	205	1	24,24,25	2.69	5 (20%)	28,32,34	3.80	11 (39%)
1	LLP	P	205	1	24,24,25	1.81	2 (8%)	28,32,34	1.73	5 (17%)

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
1	LLP	A	205	1	-	0/15/17/19	0/1/1/1
1	LLP	B	205	1	-	0/15/17/19	0/1/1/1
1	LLP	C	205	1	-	0/15/17/19	0/1/1/1
1	LLP	D	205	1	-	0/15/17/19	0/1/1/1
1	LLP	E	205	1	-	0/15/17/19	0/1/1/1
1	LLP	F	205	1	-	0/15/17/19	0/1/1/1
1	LLP	G	205	1	-	0/15/17/19	0/1/1/1
1	LLP	H	205	1	-	0/15/17/19	0/1/1/1
1	LLP	I	205	1	-	1/15/17/19	0/1/1/1
1	LLP	J	205	1	-	0/15/17/19	0/1/1/1
1	LLP	K	205	1	-	0/15/17/19	0/1/1/1
1	LLP	L	205	1	-	1/15/17/19	0/1/1/1
1	LLP	M	205	1	-	0/15/17/19	0/1/1/1
1	LLP	N	205	1	-	0/15/17/19	0/1/1/1
1	LLP	O	205	1	-	0/15/17/19	0/1/1/1
1	LLP	P	205	1	-	0/15/17/19	0/1/1/1

The worst 5 of 53 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	O	205	LLP	C3-C2	-11.02	1.33	1.40
1	J	205	LLP	C3-C2	-9.22	1.34	1.40
1	L	205	LLP	C3-C2	-8.64	1.34	1.40
1	N	205	LLP	C3-C2	-8.19	1.35	1.40
1	P	205	LLP	C3-C2	-7.19	1.35	1.40

The worst 5 of 114 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	O	205	LLP	C3-C4-C4'	-9.69	101.98	120.52
1	L	205	LLP	CD-CE-NZ	-5.89	97.87	110.88
1	I	205	LLP	CD-CE-NZ	-5.76	98.14	110.88
1	L	205	LLP	CG-CD-CE	-5.74	91.55	113.59
1	I	205	LLP	CG-CD-CE	-5.56	92.26	113.59

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	I	205	LLP	C4-C4'-NZ-CE
1	L	205	LLP	C4-C4'-NZ-CE

There are no ring outliers.

16 monomers are involved in 22 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	205	LLP	1	0
1	B	205	LLP	1	0
1	C	205	LLP	2	0
1	D	205	LLP	1	0
1	E	205	LLP	1	0
1	F	205	LLP	1	0
1	G	205	LLP	1	0
1	H	205	LLP	1	0
1	I	205	LLP	1	0
1	J	205	LLP	1	0
1	K	205	LLP	1	0
1	L	205	LLP	2	0
1	M	205	LLP	1	0
1	N	205	LLP	5	0
1	O	205	LLP	1	0
1	P	205	LLP	1	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

7 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	IMD	A	501	-	3,5,5	0.44	0	4,5,5	0.51	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	GOL	A	502	-	5,5,5	0.56	0	5,5,5	0.67	0
3	GOL	B	501[A]	-	5,5,5	0.34	0	5,5,5	0.38	0
3	GOL	B	501[B]	-	5,5,5	0.44	0	5,5,5	0.29	0
4	PO4	C	501	-	4,4,4	0.75	0	6,6,6	0.58	0
2	IMD	D	501	-	3,5,5	0.41	0	4,5,5	0.46	0
3	GOL	O	501	-	5,5,5	0.44	0	5,5,5	0.30	0

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	IMD	A	501	-	-	0/0/0/0	0/1/1/1
3	GOL	A	502	-	-	0/4/4/4	0/0/0/0
3	GOL	B	501[A]	-	-	0/4/4/4	0/0/0/0
3	GOL	B	501[B]	-	-	0/4/4/4	0/0/0/0
4	PO4	C	501	-	-	0/0/0/0	0/0/0/0
2	IMD	D	501	-	-	0/0/0/0	0/1/1/1
3	GOL	O	501	-	-	0/4/4/4	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	502	GOL	1	0
3	B	501[A]	GOL	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 ⓘ

### 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	419/424 (98%)	0.08	20 (4%)	31	43	8, 18, 49, 77	0
1	B	419/424 (98%)	0.15	16 (3%)	41	53	8, 20, 50, 79	0
1	C	419/424 (98%)	0.10	15 (3%)	43	55	6, 18, 45, 57	0
1	D	420/424 (99%)	0.08	15 (3%)	43	55	6, 18, 46, 85	0
1	E	419/424 (98%)	0.15	18 (4%)	36	48	9, 19, 50, 86	0
1	F	420/424 (99%)	0.10	15 (3%)	43	55	7, 19, 45, 76	0
1	G	420/424 (99%)	0.17	17 (4%)	39	51	9, 22, 52, 76	0
1	H	420/424 (99%)	0.15	17 (4%)	39	51	7, 20, 50, 76	0
1	I	401/424 (94%)	0.05	12 (2%)	51	62	8, 20, 45, 69	0
1	J	405/424 (95%)	0.18	20 (4%)	30	42	8, 22, 48, 71	0
1	K	404/424 (95%)	0.26	25 (6%)	21	30	9, 22, 51, 84	0
1	L	405/424 (95%)	0.20	18 (4%)	35	47	8, 24, 54, 83	0
1	M	405/424 (95%)	0.06	15 (3%)	42	54	8, 21, 47, 80	0
1	N	403/424 (95%)	0.17	21 (5%)	28	40	8, 23, 48, 69	0
1	O	403/424 (95%)	0.12	19 (4%)	32	44	8, 20, 47, 86	0
1	P	410/424 (96%)	0.15	22 (5%)	26	37	8, 20, 52, 98	0
All	All	6592/6784 (97%)	0.14	285 (4%)	36	48	6, 20, 49, 98	0

The worst 5 of 285 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	M	423	SER	6.0
1	P	423	SER	5.5
1	G	15	ASN	5.5
1	E	394	ALA	5.3
1	B	423	SER	5.3

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
1	LLP	F	205	24/25	0.96	0.14	-	13,16,19,21	0
1	LLP	I	205	24/25	0.95	0.15	-	13,18,24,27	0
1	LLP	K	205	24/25	0.95	0.17	-	14,21,25,26	0
1	LLP	M	205	24/25	0.96	0.13	-	12,14,18,18	0
1	LLP	O	205	24/25	0.95	0.15	-	13,23,29,37	0
1	LLP	A	205	24/25	0.95	0.17	-	12,15,18,24	0
1	LLP	C	205	24/25	0.96	0.15	-	11,14,17,23	0
1	LLP	L	205	24/25	0.96	0.15	-	14,23,26,29	0
1	LLP	E	205	24/25	0.97	0.14	-	12,15,18,19	0
1	LLP	G	205	24/25	0.96	0.17	-	13,18,21,24	0
1	LLP	P	205	24/25	0.95	0.15	-	12,19,24,25	0
1	LLP	J	205	24/25	0.95	0.15	-	16,19,22,24	0
1	LLP	D	205	24/25	0.95	0.15	-	13,16,18,20	0
1	LLP	N	205	24/25	0.94	0.17	-	15,23,30,32	0
1	LLP	H	205	24/25	0.96	0.14	-	13,17,20,23	0
1	LLP	B	205	24/25	0.95	0.15	-	12,16,17,20	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
3	GOL	O	501	6/6	0.93	0.18	3.59	26,27,28,33	0
3	GOL	A	502	6/6	0.92	0.18	0.88	26,27,31,31	0
4	PO4	C	501	5/5	0.96	0.16	0.19	45,47,49,51	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
2	IMD	D	501	5/5	0.90	0.15	-0.13	38,38,43,45	0
3	GOL	B	501[B]	6/6	0.93	0.15	-0.43	19,20,21,21	6
3	GOL	B	501[A]	6/6	0.93	0.15	-	11,11,12,13	6
2	IMD	A	501	5/5	0.91	0.20	-	26,28,30,31	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.