



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 07:19 AM GMT

PDB ID : 3ACZ  
Title : Crystal structure of Entamoeba histolytica methionine gamma-lyase 1  
Authors : Karaki, T.; Sato, D.; Shimizu, A.; Nozaki, T.; Harada, S.  
Deposited on : 2010-01-13  
Resolution : 1.97 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.  
We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)  
A user guide is available at  
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>  
with specific help available everywhere you see the ⓘ symbol.

---

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

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

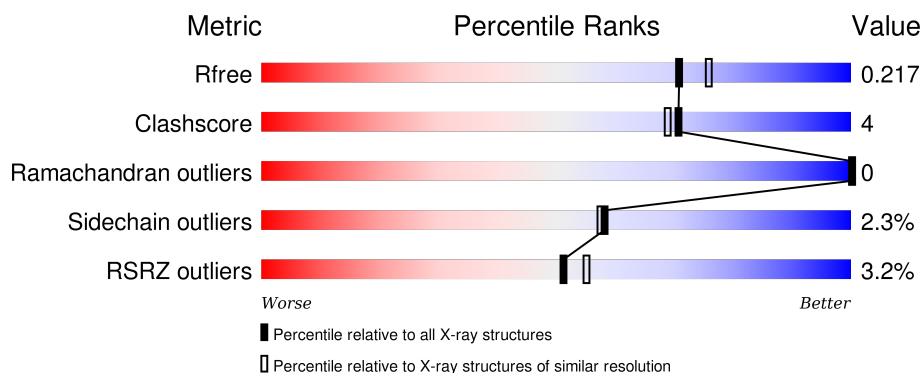
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.97 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	8664 (2.00-1.96)
Clashscore	102246	9905 (2.00-1.96)
Ramachandran outliers	100387	9792 (2.00-1.96)
Sidechain outliers	100360	9791 (2.00-1.96)
RSRZ outliers	91569	8679 (2.00-1.96)

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	389	<div> <div>2%</div> <div>89%</div> <div>10%</div> <div>..</div> </div>
1	B	389	<div> <div>3%</div> <div>89%</div> <div>9%</div> <div>.</div> </div>
1	C	389	<div> <div>3%</div> <div>88%</div> <div>11%</div> <div>..</div> </div>
1	D	389	<div> <div>4%</div> <div>86%</div> <div>12%</div> <div>.</div> </div>

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
2	SO4	A	2011	-	-	-	X
2	SO4	B	2005	-	-	-	X
2	SO4	C	2002	-	-	-	X
2	SO4	D	2008	-	-	-	X
3	GOL	C	2014	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 12951 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 Methionine gamma-lyase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	387	Total	C	N	O	P	S	0	3	0
			2975	1892	498	558	1	26			
1	B	386	Total	C	N	O	P	S	0	3	0
			2960	1881	496	556	1	26			
1	C	387	Total	C	N	O	P	S	0	5	0
			2987	1900	499	561	1	26			
1	D	384	Total	C	N	O	P	S	0	4	0
			2952	1877	494	554	1	26			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	308	LEU	SER	SEE REMARK 999	UNP Q86D28
B	808	LEU	SER	SEE REMARK 999	UNP Q86D28
C	1308	LEU	SER	SEE REMARK 999	UNP Q86D28
D	1808	LEU	SER	SEE REMARK 999	UNP Q86D28

- 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	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	A	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	B	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	C	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		
2	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 3 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
3	A	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		

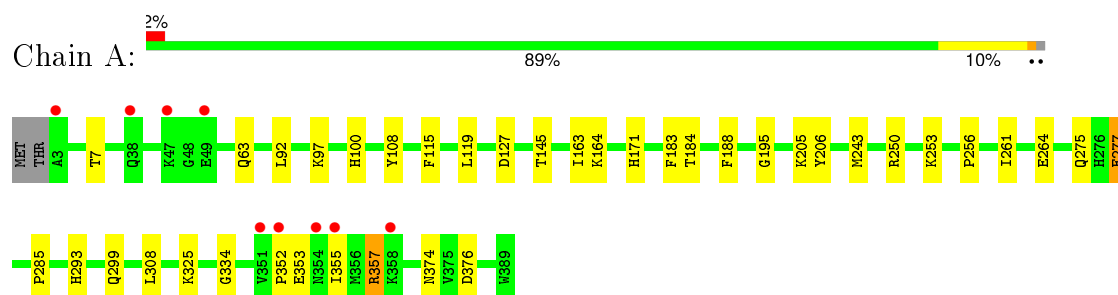
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	320	Total	O	0	0
			320	320		
4	B	197	Total	O	0	0
			197	197		
4	C	317	Total	O	0	0
			317	317		
4	D	171	Total	O	0	0
			171	171		

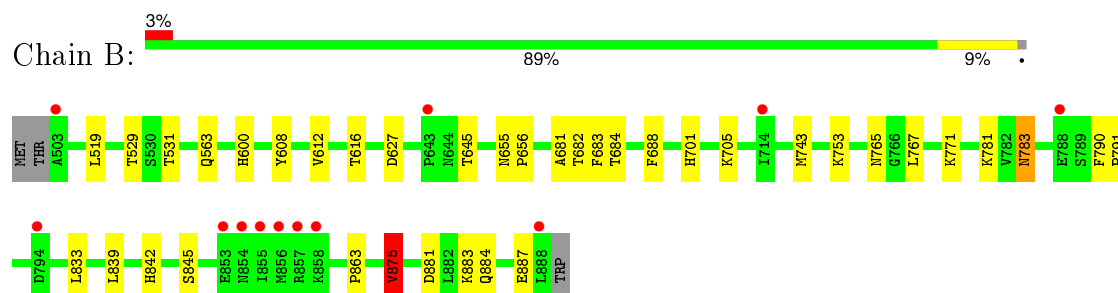
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

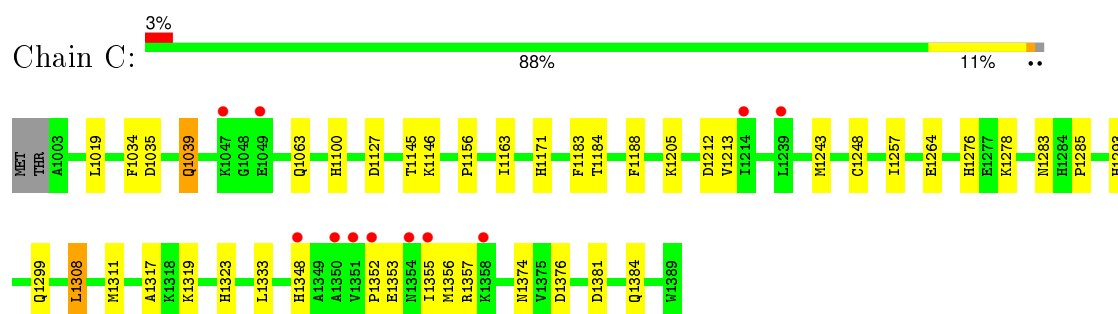
- Molecule 1: Methionine gamma-lyase



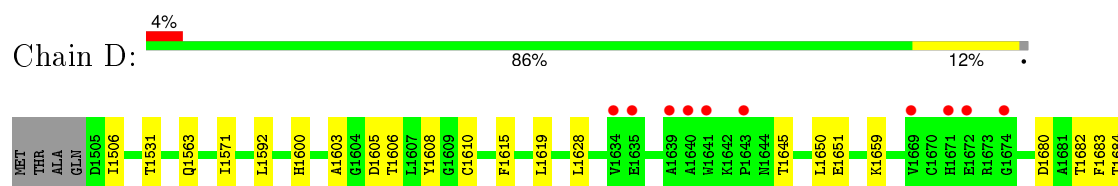
- Molecule 1: Methionine gamma-lyase

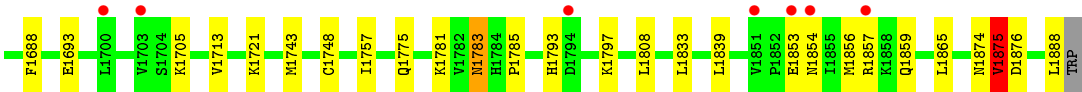


- Molecule 1: Methionine gamma-lyase



- Molecule 1: Methionine gamma-lyase







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.28 Å   85.34 Å   114.62 Å 90.00°   101.96°   90.00°	Depositor
Resolution (Å)	39.06 – 1.97 39.06 – 1.97	Depositor EDS
% Data completeness (in resolution range)	97.3 (39.06-1.97) 97.3 (39.06-1.97)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.38 (at 1.97 Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.156   ,   0.190 0.187   ,   0.217	Depositor DCC
$R_{free}$ test set	6491 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.9	Xtriage
Anisotropy	0.363	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 37.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 129283 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	12951	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.63% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, LLP, 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	0.78	0/3022	0.72	1/4084 (0.0%)
1	B	0.65	0/3005	0.67	1/4061 (0.0%)
1	C	0.79	0/3040	0.71	1/4108 (0.0%)
1	D	0.62	0/3000	0.67	1/4054 (0.0%)
All	All	0.72	0/12067	0.69	4/16307 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	875	VAL	CB-CA-C	-5.78	100.41	111.40
1	D	1875	VAL	CB-CA-C	-5.19	101.54	111.40
1	C	1308	LEU	CA-CB-CG	5.03	126.87	115.30
1	A	250	ARG	NE-CZ-NH2	5.03	122.81	120.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2975	0	2976	25	0
1	B	2960	0	2966	26	0
1	C	2987	0	2990	32	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	D	2952	0	2961	31	0
2	A	15	0	0	0	0
2	B	15	0	0	0	0
2	C	15	0	0	0	0
2	D	15	0	0	0	0
3	A	6	0	8	0	0
3	C	6	0	8	0	0
4	A	320	0	0	4	0
4	B	197	0	0	1	0
4	C	317	0	0	8	0
4	D	171	0	0	3	0
All	All	12951	0	11909	104	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 4.

All (104) 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:A:352:PRO:HG2	1:A:355:ILE:HD12	1.57	0.87
1:C:1276:HIS:HD2	1:C:1278:LYS:H	1.30	0.80
1:A:261:ILE:HD11	1:C:1257:ILE:HG12	1.69	0.76
1:C:1352:PRO:HG2	1:C:1355:ILE:HD12	1.68	0.76
1:A:92:LEU:HD21	1:A:119:LEU:HD22	1.68	0.74
1:A:353:GLU:HB3	1:A:357:ARG:HH21	1.52	0.73
1:D:1781:LYS:HE3	1:D:1783:ASN:HD21	1.56	0.70
1:B:842:HIS:HE1	1:B:863:PRO:O	1.76	0.69
1:D:1600:HIS:HD2	1:D:1645:THR:OG1	1.77	0.68
1:B:600:HIS:HD2	1:B:645:THR:OG1	1.76	0.68
1:C:1035:ASP:H	1:C:1039:GLN:NE2	1.92	0.67
1:C:1243:MET:CE	1:D:1743:MET:HE1	2.26	0.66
1:A:285:PRO:O	1:A:293:HIS:HD2	1.80	0.65
1:B:781:LYS:HD3	1:B:783:ASN:HD21	1.61	0.64
1:D:1592:LEU:HD21	1:D:1619:LEU:HD22	1.81	0.63
1:C:1100:HIS:HD2	1:C:1145:THR:OG1	1.80	0.63
1:C:1063[A]:GLN:HG2	4:C:3564:HOH:O	1.99	0.61
1:B:600:HIS:HE1	1:B:627:ASP:OD2	1.82	0.61
1:D:1705:LLP:HD3	1:D:1833:LEU:HG	1.83	0.61
1:B:781:LYS:HD3	1:B:783:ASN:ND2	2.16	0.60
1:A:100:HIS:HD2	1:A:145:THR:OG1	1.84	0.60
1:A:264:GLU:HG3	4:A:3675:HOH:O	2.02	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:1682:THR:HB	1:D:1705:LLP:H2'2	1.84	0.59
1:D:1874:ASN:HD21	1:D:1876:ASP:HB2	1.67	0.59
1:C:1100:HIS:HE1	1:C:1127:ASP:OD2	1.85	0.59
1:C:1276:HIS:CD2	1:C:1278:LYS:H	2.15	0.58
1:D:1684:THR:HG22	1:D:1688:PHE:HB2	1.86	0.58
1:D:1608:TYR:CE1	1:D:1705:LLP:H4'1	2.39	0.57
1:C:1319:LYS:O	1:C:1323:HIS:HD2	1.88	0.56
1:A:243:MET:CE	1:B:743:MET:HE1	2.35	0.56
1:A:63:GLN:HG2	4:A:3343:HOH:O	2.04	0.56
1:B:765:ASN:ND2	1:B:875:VAL:HG22	2.21	0.55
1:B:682:THR:HB	1:B:705:LLP:H2'2	1.89	0.55
1:D:1875:VAL:HG22	4:D:3269:HOH:O	2.06	0.55
1:C:1285:PRO:O	1:C:1293:HIS:HD2	1.90	0.55
1:D:1781:LYS:CE	1:D:1783:ASN:HD21	2.20	0.53
1:D:1606:THR:HG22	1:D:1859:GLN:HG2	1.89	0.53
1:A:100:HIS:HE1	1:A:127:ASP:OD2	1.93	0.52
1:B:883:LYS:O	1:B:887:GLU:HG3	2.09	0.52
1:D:1651:GLU:HG3	1:D:1680:ASP:HB3	1.90	0.52
1:B:608:TYR:CE1	1:B:705:LLP:H4'1	2.44	0.52
1:C:1308:LEU:HB3	4:C:3348:HOH:O	2.10	0.52
1:C:1243:MET:CE	1:D:1743:MET:CE	2.88	0.52
1:C:1205:LLP:HD3	1:C:1333:LEU:HG	1.90	0.52
1:A:261:ILE:CD1	1:C:1257:ILE:HG12	2.40	0.51
1:C:1243:MET:HE1	1:D:1743:MET:CE	2.40	0.51
1:C:1264:GLU:HG3	4:C:3850:HOH:O	2.10	0.51
4:A:3182:HOH:O	1:D:1531:THR:HG22	2.11	0.50
1:C:1146:LYS:HE3	4:C:3317:HOH:O	2.11	0.50
1:D:1608:TYR:HE2	1:D:1610:CYS:HB2	1.78	0.49
1:A:184:THR:HG22	1:A:188:PHE:HB2	1.94	0.49
1:C:1034:PHE:HA	1:C:1039:GLN:HE22	1.77	0.49
1:C:1163:ILE:H	1:C:1299:GLN:HE22	1.60	0.49
1:C:1374:ASN:HD21	1:C:1376:ASP:HB2	1.78	0.49
1:D:1874:ASN:ND2	1:D:1876:ASP:HB2	2.28	0.48
1:A:206:TYR:CE2	1:A:334:GLY:HA2	2.49	0.48
1:C:1184:THR:HG22	1:C:1188:PHE:HB2	1.96	0.48
1:B:705:LLP:HD3	1:B:833:LEU:HG	1.95	0.48
1:B:881:ASP:O	1:B:884:GLN:HB3	2.13	0.48
1:A:308:LEU:HB3	4:A:3314:HOH:O	2.13	0.48
1:D:1650:LEU:HA	4:D:3368:HOH:O	2.13	0.47
1:A:171:HIS:HE1	1:A:195:GLY:O	1.96	0.47
1:D:1615:PHE:HA	1:D:1619:LEU:HG	1.96	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:1757:ILE:HD12	4:D:3208:HOH:O	2.15	0.47
1:D:1605:ASP:OD2	1:D:1606:THR:HG23	2.14	0.47
1:A:7:THR:HG22	1:A:256:PRO:HB2	1.97	0.47
1:B:842:HIS:HD2	1:B:845:SER:OG	1.98	0.47
1:B:684:THR:HG22	1:B:688:PHE:HB2	1.97	0.46
1:D:1506:ILE:HG23	1:D:1571:ILE:HD12	1.97	0.46
1:C:1353:GLU:O	1:C:1357:ARG:HG3	2.16	0.46
1:D:1608:TYR:CE2	1:D:1610:CYS:HB2	2.51	0.45
1:B:681:ALA:HB3	1:B:701:HIS:CE1	2.51	0.45
1:C:1171:HIS:HD2	4:C:3041:HOH:O	1.98	0.45
1:B:705:LLP:NZ	1:B:705:LLP:O3	2.49	0.45
1:D:1785:PRO:O	1:D:1793:HIS:HD2	2.00	0.45
1:B:531:THR:HG22	4:C:3101:HOH:O	2.17	0.45
1:B:600:HIS:CE1	1:B:627:ASP:OD2	2.68	0.45
1:A:243:MET:HE1	1:B:743:MET:HE1	1.99	0.45
1:B:612:VAL:O	1:B:616:THR:HB	2.17	0.44
1:C:1311:MET:HG3	1:C:1317:ALA:HA	2.00	0.44
1:A:374:ASN:HD21	1:A:376:ASP:HB2	1.82	0.44
1:A:277:GLU:O	1:A:277:GLU:HG2	2.18	0.43
1:D:1713:VAL:HG11	1:D:1748:CYS:SG	2.59	0.43
1:A:115:PHE:HA	1:A:119:LEU:HG	2.01	0.43
1:B:790:PHE:HA	1:B:791:PRO:HD3	1.91	0.43
1:D:1603:ALA:HB3	1:D:1628:LEU:HD22	2.00	0.43
1:D:1808:LEU:HD21	1:D:1865:LEU:HD22	2.01	0.42
1:A:108:TYR:CE1	1:A:205:LLP:H4'1	2.53	0.42
1:B:753:LYS:NZ	4:B:3848:HOH:O	2.46	0.42
1:D:1853:GLU:HG3	1:D:1857:ARG:HH21	1.85	0.42
1:C:1348:HIS:CD2	4:C:3772:HOH:O	2.72	0.42
1:A:163:ILE:H	1:A:299:GLN:HE22	1.68	0.41
1:C:1213:VAL:HG11	1:C:1248:CYS:SG	2.60	0.41
1:B:529:THR:HA	1:C:1212:ASP:HA	2.02	0.41
1:C:1381:ASP:O	1:C:1384:GLN:HG3	2.20	0.41
1:A:243:MET:CE	1:B:743:MET:CE	2.98	0.41
1:B:767:LEU:O	1:B:771:LYS:HG3	2.20	0.41
1:C:1156:PRO:HD3	1:C:1348:HIS:NE2	2.35	0.41
1:B:655:ASN:HA	1:B:656:PRO:HA	1.83	0.41
1:C:1243:MET:HE1	1:D:1743:MET:HE1	1.98	0.41
1:A:253:LYS:HE3	4:C:3551:HOH:O	2.21	0.41
1:C:1353:GLU:HA	1:C:1356:MET:HB3	2.03	0.40
1:A:97:LYS:HE2	1:A:97:LYS:HB3	1.92	0.40
1:D:1705:LLP:O3	1:D:1705:LLP:NZ	2.50	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	387/389 (100%)	377 (97%)	10 (3%)	0	100	100
1	B	386/389 (99%)	376 (97%)	10 (3%)	0	100	100
1	C	389/389 (100%)	381 (98%)	8 (2%)	0	100	100
1	D	385/389 (99%)	380 (99%)	5 (1%)	0	100	100
All	All	1547/1556 (99%)	1514 (98%)	33 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	322/321 (100%)	316 (98%)	6 (2%)	65	65
1	B	321/321 (100%)	315 (98%)	6 (2%)	65	65
1	C	324/321 (101%)	320 (99%)	4 (1%)	78	80
1	D	321/321 (100%)	307 (96%)	14 (4%)	35	27
All	All	1288/1284 (100%)	1258 (98%)	30 (2%)	58	57

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	164	LYS
1	A	183	PHE
1	A	275	GLN
1	A	277	GLU
1	A	325	LYS
1	A	357	ARG
1	B	519	LEU
1	B	563	GLN
1	B	683	PHE
1	B	783	ASN
1	B	839	LEU
1	B	875	VAL
1	C	1019	LEU
1	C	1039	GLN
1	C	1183	PHE
1	C	1283	ASN
1	D	1563[A]	GLN
1	D	1563[B]	GLN
1	D	1659	LYS
1	D	1683	PHE
1	D	1693	GLU
1	D	1721	LYS
1	D	1775	GLN
1	D	1783	ASN
1	D	1797	LYS
1	D	1839	LEU
1	D	1854	ASN
1	D	1856	MET
1	D	1875	VAL
1	D	1888	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (37) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	28	GLN
1	A	100	HIS
1	A	265	ASN
1	A	275	GLN
1	A	293	HIS
1	A	299	GLN
1	A	354	ASN
1	A	374	ASN
1	B	528	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	563	GLN
1	B	596	GLN
1	B	600	HIS
1	B	765	ASN
1	B	783	ASN
1	B	793	HIS
1	B	799	GLN
1	B	842	HIS
1	B	884	GLN
1	C	1028	GLN
1	C	1039	GLN
1	C	1100	HIS
1	C	1117	HIS
1	C	1265	ASN
1	C	1276	HIS
1	C	1283	ASN
1	C	1293	HIS
1	C	1299	GLN
1	C	1323	HIS
1	C	1374	ASN
1	D	1600	HIS
1	D	1617	HIS
1	D	1765	ASN
1	D	1775	GLN
1	D	1783	ASN
1	D	1793	HIS
1	D	1854	ASN
1	D	1874	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 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	23,24,25	1.97	8 (34%)	28,32,34	1.81	7 (25%)
1	LLP	B	705	1	23,24,25	2.21	6 (26%)	28,32,34	2.10	8 (28%)
1	LLP	C	1205	1	23,24,25	1.89	5 (21%)	28,32,34	1.85	5 (17%)
1	LLP	D	1705	1	23,24,25	2.05	5 (21%)	28,32,34	1.81	6 (21%)

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	705	1	-	0/15/17/19	0/1/1/1
1	LLP	C	1205	1	-	0/15/17/19	0/1/1/1
1	LLP	D	1705	1	-	0/15/17/19	0/1/1/1

All (24) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	705	LLP	O3-C3	-6.06	1.22	1.37
1	A	205	LLP	O3-C3	-5.87	1.23	1.37
1	D	1705	LLP	O3-C3	-5.45	1.24	1.37
1	C	1205	LLP	O3-C3	-4.76	1.25	1.37
1	A	205	LLP	C2-N1	2.11	1.38	1.34
1	A	205	LLP	CB-CA	2.32	1.56	1.53
1	A	205	LLP	CD-CE	2.33	1.58	1.51
1	B	705	LLP	CD-CE	2.33	1.58	1.51
1	A	205	LLP	C4-C4'	2.37	1.50	1.46
1	C	1205	LLP	C2-N1	2.54	1.39	1.34
1	A	205	LLP	CE-NZ	2.56	1.52	1.46
1	A	205	LLP	C4'-NZ	2.63	1.35	1.27
1	B	705	LLP	C6-N1	2.63	1.40	1.34
1	A	205	LLP	C6-N1	2.84	1.40	1.34
1	C	1205	LLP	C4'-NZ	2.95	1.36	1.27
1	C	1205	LLP	C4-C4'	3.06	1.52	1.46
1	D	1705	LLP	CE-NZ	3.15	1.53	1.46
1	D	1705	LLP	C4'-NZ	3.26	1.37	1.27
1	B	705	LLP	C4-C4'	3.27	1.52	1.46
1	D	1705	LLP	C4-C4'	3.36	1.52	1.46

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	705	LLP	C4'-NZ	3.38	1.37	1.27
1	C	1205	LLP	CE-NZ	3.59	1.54	1.46
1	D	1705	LLP	C6-N1	3.62	1.42	1.34
1	B	705	LLP	CE-NZ	4.07	1.55	1.46

All (26) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	1205	LLP	O-C-CA	-2.83	118.12	125.49
1	B	705	LLP	O-C-CA	-2.78	118.25	125.49
1	C	1205	LLP	OP4-P-OP1	-2.74	100.16	107.14
1	A	205	LLP	O-C-CA	-2.67	118.53	125.49
1	D	1705	LLP	C4-C4'-NZ	-2.63	110.44	125.06
1	B	705	LLP	OP4-P-OP1	-2.60	100.52	107.14
1	D	1705	LLP	O-C-CA	-2.33	119.42	125.49
1	D	1705	LLP	OP4-P-OP1	-2.31	101.27	107.14
1	A	205	LLP	C4-C4'-NZ	-2.27	112.44	125.06
1	A	205	LLP	OP4-P-OP1	-2.15	101.68	107.14
1	D	1705	LLP	CE-NZ-C4'	-2.14	112.79	118.97
1	B	705	LLP	CE-NZ-C4'	-2.08	112.97	118.97
1	C	1205	LLP	C5-C6-N1	-2.06	120.28	123.86
1	B	705	LLP	C5-C6-N1	-2.04	120.32	123.86
1	B	705	LLP	C4-C4'-NZ	-2.01	113.86	125.06
1	A	205	LLP	CE-NZ-C4'	-2.01	113.18	118.97
1	A	205	LLP	OP3-P-OP1	2.01	117.04	110.58
1	B	705	LLP	OP3-P-OP1	2.42	118.36	110.58
1	D	1705	LLP	CD-CE-NZ	3.70	117.04	110.98
1	C	1205	LLP	CD-CE-NZ	3.99	117.50	110.98
1	A	205	LLP	OP4-C5'-C5	4.91	117.11	108.99
1	A	205	LLP	CD-CE-NZ	5.79	120.45	110.98
1	B	705	LLP	CD-CE-NZ	5.99	120.78	110.98
1	D	1705	LLP	OP4-C5'-C5	6.29	119.39	108.99
1	C	1205	LLP	OP4-C5'-C5	6.32	119.44	108.99
1	B	705	LLP	OP4-C5'-C5	6.44	119.63	108.99

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	205	LLP	1	0
1	B	705	LLP	4	0
1	C	1205	LLP	1	0
1	D	1705	LLP	4	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

14 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	A	2001	-	4,4,4	0.84	0	6,6,6	0.81	0
2	SO4	A	2003	-	4,4,4	0.30	0	6,6,6	0.19	0
2	SO4	A	2011	-	4,4,4	0.21	0	6,6,6	0.11	0
3	GOL	A	2013	-	5,5,5	0.41	0	5,5,5	0.31	0
2	SO4	B	2004	-	4,4,4	0.46	0	6,6,6	0.36	0
2	SO4	B	2005	-	4,4,4	0.25	0	6,6,6	0.14	0
2	SO4	B	2006	-	4,4,4	0.24	0	6,6,6	0.14	0
2	SO4	C	2002	-	4,4,4	0.21	0	6,6,6	0.22	0
2	SO4	C	2007	-	4,4,4	0.93	0	6,6,6	0.62	0
2	SO4	C	2009	-	4,4,4	0.26	0	6,6,6	0.29	0
3	GOL	C	2014	-	5,5,5	0.38	0	5,5,5	0.38	0
2	SO4	D	2008	-	4,4,4	0.19	0	6,6,6	0.07	0
2	SO4	D	2010	-	4,4,4	0.40	0	6,6,6	0.54	0
2	SO4	D	2012	-	4,4,4	0.28	0	6,6,6	0.24	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	SO4	A	2001	-	-	0/0/0/0	0/0/0/0
2	SO4	A	2003	-	-	0/0/0/0	0/0/0/0
2	SO4	A	2011	-	-	0/0/0/0	0/0/0/0
3	GOL	A	2013	-	-	0/4/4/4	0/0/0/0
2	SO4	B	2004	-	-	0/0/0/0	0/0/0/0
2	SO4	B	2005	-	-	0/0/0/0	0/0/0/0
2	SO4	B	2006	-	-	0/0/0/0	0/0/0/0
2	SO4	C	2002	-	-	0/0/0/0	0/0/0/0
2	SO4	C	2007	-	-	0/0/0/0	0/0/0/0
2	SO4	C	2009	-	-	0/0/0/0	0/0/0/0
3	GOL	C	2014	-	-	0/4/4/4	0/0/0/0
2	SO4	D	2008	-	-	0/0/0/0	0/0/0/0
2	SO4	D	2010	-	-	0/0/0/0	0/0/0/0
2	SO4	D	2012	-	-	0/0/0/0	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.

No monomer is involved in short contacts.

## 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	386/389 (99%)	0.08	9 (2%) 64 67	14, 19, 38, 65	1 (0%)
1	B	385/389 (98%)	0.26	12 (3%) 52 56	15, 30, 52, 76	1 (0%)
1	C	386/389 (99%)	0.08	11 (2%) 56 60	14, 19, 38, 67	0
1	D	383/389 (98%)	0.32	17 (4%) 38 42	16, 32, 56, 75	0
All	All	1540/1556 (98%)	0.18	49 (3%) 51 55	14, 23, 52, 76	2 (0%)

All (49) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	1351	VAL	6.8
1	B	503	ALA	6.7
1	A	351	VAL	5.4
1	C	1354	ASN	5.0
1	B	854	ASN	4.7
1	A	355	ILE	4.7
1	C	1355	ILE	4.6
1	A	354	ASN	4.0
1	C	1352	PRO	4.0
1	B	853	GLU	3.9
1	D	1634	VAL	3.8
1	D	1854	ASN	3.7
1	B	857	ARG	3.7
1	A	352	PRO	3.7
1	D	1794	ASP	3.6
1	B	858	LYS	3.2
1	A	358	LYS	3.2
1	D	1669	VAL	3.2
1	C	1350	ALA	3.1
1	D	1671	HIS	3.0
1	D	1857	ARG	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	1047	LYS	3.0
1	D	1643	PRO	2.9
1	B	788	GLU	2.8
1	B	855	ILE	2.8
1	D	1853	GLU	2.7
1	D	1641	TRP	2.7
1	B	888	LEU	2.6
1	D	1639	ALA	2.6
1	B	856	MET	2.5
1	C	1214	ILE	2.5
1	A	38	GLN	2.5
1	D	1700	LEU	2.5
1	D	1674	GLY	2.5
1	D	1851	VAL	2.4
1	D	1640	ALA	2.3
1	A	49	GLU	2.3
1	D	1635	GLU	2.2
1	A	47	LYS	2.2
1	B	794	ASP	2.2
1	D	1703	VAL	2.2
1	A	3	ALA	2.1
1	C	1358	LYS	2.1
1	B	714	ILE	2.1
1	B	643	PRO	2.1
1	D	1672	GLU	2.1
1	C	1239	LEU	2.0
1	C	1348	HIS	2.0
1	C	1049	GLU	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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. 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	B	705	24/25	0.95	0.27	-	15,21,26,29	10
1	LLP	C	1205	24/25	0.97	0.14	-	14,18,20,21	6
1	LLP	A	205	24/25	0.96	0.16	-	14,17,20,20	6

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
1	LLP	D	1705	24/25	0.95	0.26	-	17,23,28,29	10

### 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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( $\text{\AA}^2$ )	Q<0.9
2	SO4	D	2008	5/5	0.76	0.29	11.24	106,106,106,106	0
2	SO4	A	2011	5/5	0.64	0.29	7.17	110,110,111,111	0
2	SO4	B	2005	5/5	0.83	0.21	3.96	104,104,104,104	0
2	SO4	C	2002	5/5	0.72	0.22	3.85	90,90,91,91	0
3	GOL	C	2014	6/6	0.93	0.12	2.02	40,45,46,48	0
3	GOL	A	2013	6/6	0.87	0.11	1.05	45,47,48,49	0
2	SO4	B	2004	5/5	0.94	0.12	-0.27	46,47,49,52	0
2	SO4	C	2007	5/5	0.97	0.11	-0.46	30,32,35,40	0
2	SO4	D	2010	5/5	0.96	0.11	-0.69	44,46,47,50	0
2	SO4	A	2001	5/5	0.98	0.09	-1.41	28,31,33,38	0
2	SO4	B	2006	5/5	0.94	0.26	-	73,73,74,74	0
2	SO4	D	2012	5/5	0.95	0.26	-	73,73,73,74	0
2	SO4	C	2009	5/5	0.94	0.15	-	61,61,62,63	0
2	SO4	A	2003	5/5	0.94	0.12	-	62,63,63,64	0

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

There are no such residues in this entry.