



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:22 AM GMT

PDB ID : 3ED4
Title : Crystal structure of putative arylsulfatase from escherichia coli
Authors : Patskovsky, Y.; Ozyurt, S.; Gilmore, M.; Chang, S.; Bain, K.; Wasserman, S.; Koss, J.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on : 2008-09-02
Resolution : 1.70 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
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

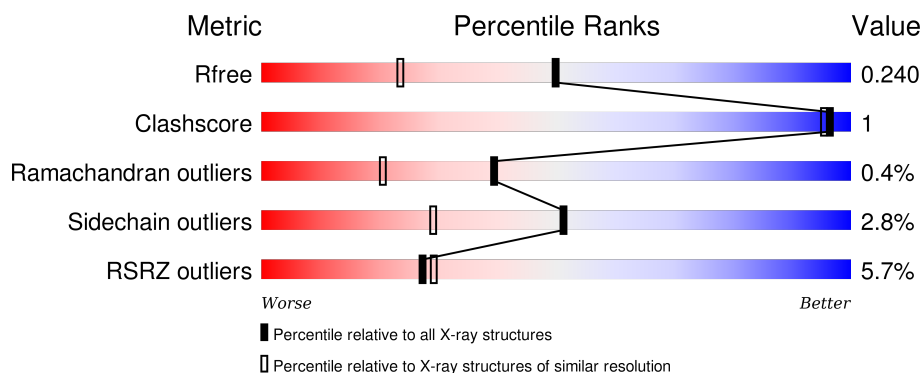
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.70 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	3190 (1.70-1.70)
Clashscore	102246	3585 (1.70-1.70)
Ramachandran outliers	100387	3527 (1.70-1.70)
Sidechain outliers	100360	3527 (1.70-1.70)
RSRZ outliers	91569	3200 (1.70-1.70)

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 ≥ 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	502	<div> <div>4%</div> <div>88%</div> <div>5%</div> <div>7%</div> </div>
1	B	502	<div> <div>3%</div> <div>90%</div> <div>•</div> <div>6%</div> </div>
1	C	502	<div> <div>8%</div> <div>88%</div> <div>•</div> <div>7%</div> </div>
1	D	502	<div> <div>6%</div> <div>87%</div> <div>7%</div> <div>6%</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	GOL	A	1	-	-	-	X
2	GOL	B	1	-	-	X	X
2	GOL	B	519	-	-	-	X
3	NA	A	518	-	-	-	X
3	NA	C	1	-	-	-	X
3	NA	C	518	-	-	-	X
3	NA	D	1	-	-	-	X
4	SO4	A	520	-	-	-	X
4	SO4	B	2	-	-	-	X
4	SO4	C	3	-	-	-	X
4	SO4	D	4	-	-	-	X
5	UNL	A	521	-	-	-	X
5	UNL	B	520	-	-	-	X
5	UNL	C	519	-	-	-	X
5	UNL	D	518	-	-	-	X
5	UNL	D	519	-	-	-	X
5	UNL	D	520	-	-	X	-

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16828 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 ARYLSULFATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	468	Total	C	N	O	S	0	12	0
			3800	2426	642	714	18			
1	B	474	Total	C	N	O	S	0	7	0
			3807	2428	643	718	18			
1	C	467	Total	C	N	O	S	0	4	0
			3747	2390	633	706	18			
1	D	474	Total	C	N	O	S	0	18	0
			3883	2481	658	725	19			

There are 44 discrepancies between the modelled and reference sequences:

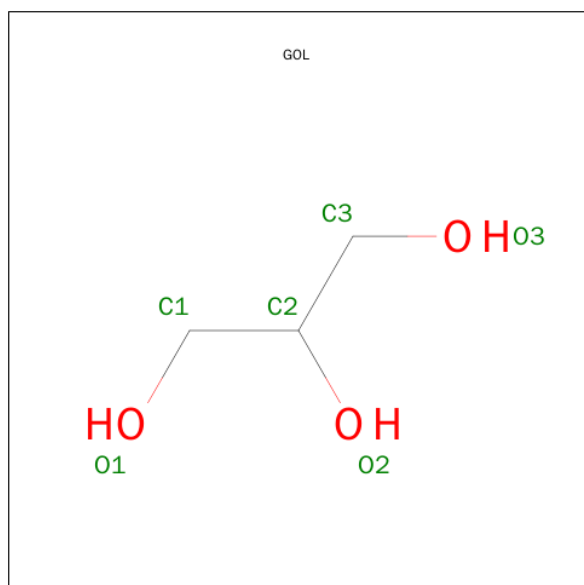
Chain	Residue	Modelled	Actual	Comment	Reference
A	16	MET	-	expression tag	UNP Q8FLC3
A	17	SER	-	expression tag	UNP Q8FLC3
A	18	LEU	-	expression tag	UNP Q8FLC3
A	510	GLU	-	expression tag	UNP Q8FLC3
A	511	GLY	-	expression tag	UNP Q8FLC3
A	512	HIS	-	expression tag	UNP Q8FLC3
A	513	HIS	-	expression tag	UNP Q8FLC3
A	514	HIS	-	expression tag	UNP Q8FLC3
A	515	HIS	-	expression tag	UNP Q8FLC3
A	516	HIS	-	expression tag	UNP Q8FLC3
A	517	HIS	-	expression tag	UNP Q8FLC3
B	16	MET	-	expression tag	UNP Q8FLC3
B	17	SER	-	expression tag	UNP Q8FLC3
B	18	LEU	-	expression tag	UNP Q8FLC3
B	510	GLU	-	expression tag	UNP Q8FLC3
B	511	GLY	-	expression tag	UNP Q8FLC3
B	512	HIS	-	expression tag	UNP Q8FLC3
B	513	HIS	-	expression tag	UNP Q8FLC3
B	514	HIS	-	expression tag	UNP Q8FLC3
B	515	HIS	-	expression tag	UNP Q8FLC3
B	516	HIS	-	expression tag	UNP Q8FLC3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	517	HIS	-	expression tag	UNP Q8FLC3
C	16	MET	-	expression tag	UNP Q8FLC3
C	17	SER	-	expression tag	UNP Q8FLC3
C	18	LEU	-	expression tag	UNP Q8FLC3
C	510	GLU	-	expression tag	UNP Q8FLC3
C	511	GLY	-	expression tag	UNP Q8FLC3
C	512	HIS	-	expression tag	UNP Q8FLC3
C	513	HIS	-	expression tag	UNP Q8FLC3
C	514	HIS	-	expression tag	UNP Q8FLC3
C	515	HIS	-	expression tag	UNP Q8FLC3
C	516	HIS	-	expression tag	UNP Q8FLC3
C	517	HIS	-	expression tag	UNP Q8FLC3
D	16	MET	-	expression tag	UNP Q8FLC3
D	17	SER	-	expression tag	UNP Q8FLC3
D	18	LEU	-	expression tag	UNP Q8FLC3
D	510	GLU	-	expression tag	UNP Q8FLC3
D	511	GLY	-	expression tag	UNP Q8FLC3
D	512	HIS	-	expression tag	UNP Q8FLC3
D	513	HIS	-	expression tag	UNP Q8FLC3
D	514	HIS	-	expression tag	UNP Q8FLC3
D	515	HIS	-	expression tag	UNP Q8FLC3
D	516	HIS	-	expression tag	UNP Q8FLC3
D	517	HIS	-	expression tag	UNP Q8FLC3

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

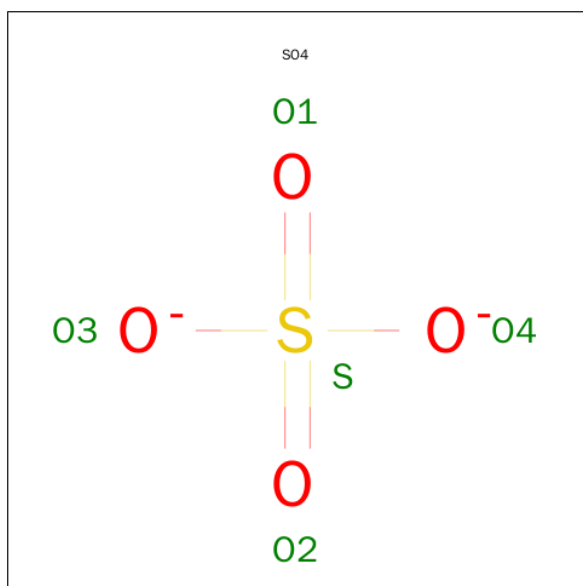


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		
2	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Na	0	0
			1	1		
3	A	2	Total	Na	0	0
			2	2		
3	D	1	Total	Na	0	0
			1	1		
3	C	2	Total	Na	0	0
			2	2		

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	S	0	0
			5	4	1		
4	B	1	Total	O	S	0	0
			5	4	1		
4	C	1	Total	O	S	0	0
			5	4	1		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	O	S	0	0
			5	4	1		

- Molecule 5 is UNKNOWN LIGAND (three-letter code: UNL) (formula:).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			4	1	3		
5	A	1	Total	C	O	0	0
			4	1	3		
5	D	3	Total	C	O	0	0
			12	3	9		
5	C	1	Total	C	O	0	0
			4	1	3		

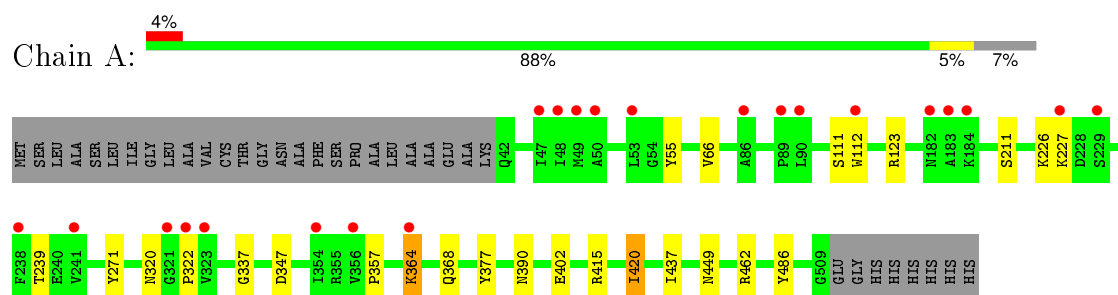
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	387	Total	O	0	0
			387	387		
6	B	438	Total	O	0	0
			438	438		
6	C	315	Total	O	0	0
			315	315		
6	D	382	Total	O	0	1
			383	383		

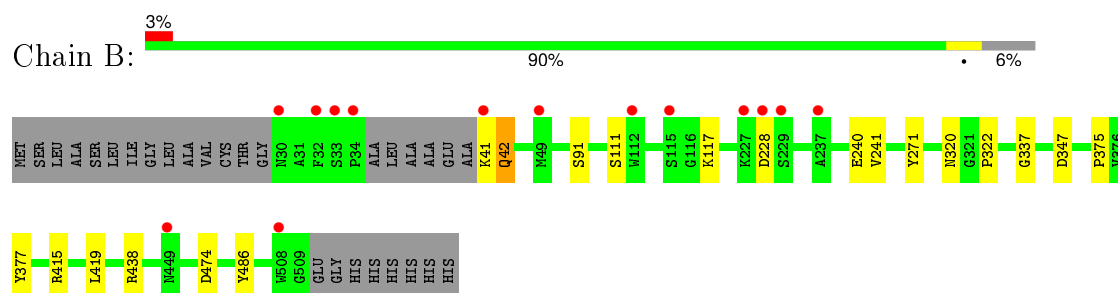
3 Residue-property plots [i](#)

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.

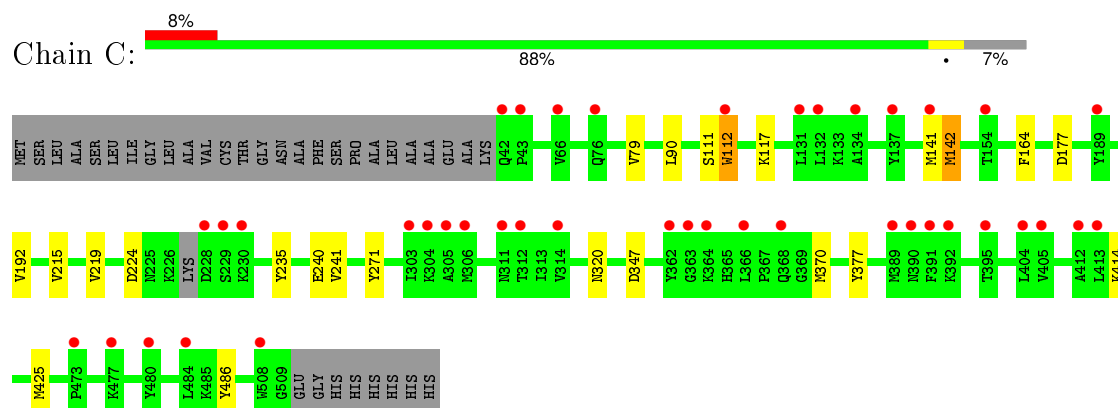
• Molecule 1: ARYLSULFATASE



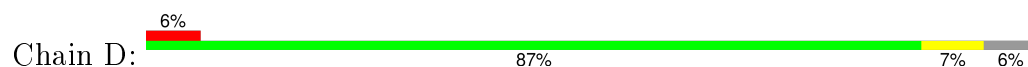
• Molecule 1: ARYLSULFATASE

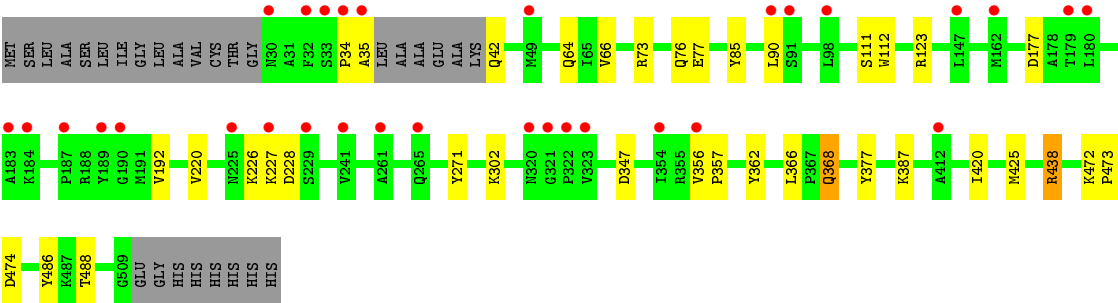


• Molecule 1: ARYLSULFATASE



• Molecule 1: ARYLSULFATASE





4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	71.93Å 178.86Å 76.28Å 90.00° 92.17° 90.00°	Depositor
Resolution (Å)	20.00 – 1.70 39.74 – 1.70	Depositor EDS
% Data completeness (in resolution range)	87.3 (20.00-1.70) 87.2 (39.74-1.70)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	0.06	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.40 (at 1.70Å)	Xtriage
Refinement program	REFMAC 5.3.0034	Depositor
R, R_{free}	0.196 , 0.242 0.194 , 0.240	Depositor DCC
R_{free} test set	5536 reflections (3.09%)	DCC
Wilson B-factor (Å ²)	29.8	Xtriage
Anisotropy	0.235	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 45.6	EDS
Estimated twinning fraction	0.027 for h,-k,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 184819 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	16828	wwPDB-VP
Average B, all atoms (Å ²)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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

5 Model quality [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, UNL, SO4, NA

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	0.49	0/3921	0.60	0/5308
1	B	0.48	0/3923	0.60	0/5312
1	C	0.43	0/3851	0.59	1/5216 (0.0%)
1	D	0.47	0/4032	0.64	1/5455 (0.0%)
All	All	0.47	0/15727	0.61	2/21291 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	90	LEU	CA-CB-CG	5.49	127.93	115.30
1	C	90	LEU	CA-CB-CG	5.06	126.94	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3800	0	3773	9	0
1	B	3807	0	3760	11	0
1	C	3747	0	3683	9	0
1	D	3883	0	3876	14	0
2	A	6	0	8	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	12	0	14	6	0
3	A	2	0	0	0	0
3	B	1	0	0	0	0
3	C	2	0	0	0	0
3	D	1	0	0	0	0
4	A	5	0	0	0	0
4	B	5	0	0	1	0
4	C	5	0	0	0	0
4	D	5	0	0	0	0
5	A	4	0	0	0	0
5	B	4	0	0	0	0
5	C	4	0	0	0	0
5	D	12	0	0	4	0
6	A	387	0	0	1	0
6	B	438	0	0	0	0
6	C	315	0	0	0	0
6	D	383	0	0	1	0
All	All	16828	0	15114	43	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 1.

All (43) 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:462:ARG:HH21	2:A:1:GOL:H32	1.34	0.88
1:A:239[B]:THR:HG23	6:A:802:HOH:O	1.93	0.67
1:B:375:PRO:HB2	2:B:1:GOL:H31	1.77	0.66
1:B:375:PRO:CB	2:B:1:GOL:H31	2.29	0.63
1:D:34:PRO:HA	1:D:35:ALA:HB3	1.82	0.62
1:D:473:PRO:HD2	5:D:520:UNL:O2	2.01	0.60
1:B:375:PRO:HG2	2:B:1:GOL:H31	1.83	0.58
1:A:420:ILE:HD12	1:A:437[A]:ILE:HG22	1.85	0.58
1:A:462:ARG:NH2	2:A:1:GOL:H32	2.13	0.56
1:B:438:ARG:HG3	2:B:1:GOL:O2	2.06	0.55
1:D:123[A]:ARG:HG3	1:D:123[A]:ARG:HH21	1.72	0.54
1:D:64:GLN:NE2	5:D:518:UNL:O2	2.40	0.54
1:B:375:PRO:CG	2:B:1:GOL:H31	2.39	0.53
1:C:112[A]:TRP:CZ2	1:C:425:MET:HG3	2.44	0.52
1:C:141:MET:HB2	1:C:164:PHE:CD1	2.44	0.52
1:D:474:ASP:HB2	5:D:520:UNL:O3	2.10	0.52

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:91:SER:OG	4:B:2:SO4:O1	2.31	0.49
1:D:77:GLU:HA	1:D:368:GLN:HG3	1.95	0.48
1:A:55:TYR:HA	1:A:357:PRO:HD3	1.96	0.48
1:C:141:MET:HG3	1:C:235:TYR:HD2	1.79	0.47
1:B:419:LEU:HB2	1:B:438:ARG:HB3	1.95	0.47
1:A:402:GLU:HG3	1:A:415:ARG:HG3	1.97	0.47
1:C:215:VAL:O	1:C:219:VAL:HG23	2.15	0.47
1:A:364:LYS:HA	1:A:368:GLN:HE22	1.80	0.47
1:D:112:TRP:CZ2	1:D:425:MET:HG3	2.51	0.46
1:D:220:VAL:HG13	1:D:302:LYS:HD3	1.98	0.45
1:A:211:SER:OG	1:A:239[B]:THR:HG21	2.16	0.45
1:D:356:VAL:HB	1:D:357:PRO:HD2	1.98	0.45
1:D:76[A]:GLN:NE2	6:D:618:HOH:O	2.50	0.44
1:B:240:GLU:HA	1:B:241:VAL:HA	1.84	0.44
1:C:79:VAL:HG22	1:C:370:MET:HB3	2.00	0.43
1:C:142:MET:HB2	1:C:142:MET:HE2	1.73	0.43
1:D:85:TYR:OH	1:D:438[B]:ARG:NH1	2.51	0.43
1:D:177:ASP:HB2	1:D:192:VAL:HG13	1.99	0.43
1:A:322:PRO:HB2	1:A:337:GLY:HA3	2.01	0.42
1:C:240:GLU:HA	1:C:241:VAL:HA	1.82	0.41
1:C:414:LYS:HD2	1:C:414:LYS:HA	1.86	0.41
1:C:177:ASP:HB2	1:C:192:VAL:HG13	2.02	0.41
1:B:415:ARG:HG3	2:B:1:GOL:H11	2.01	0.41
1:D:472:LYS:HD3	5:D:520:UNL:O1	2.20	0.41
1:B:42:GLN:H	1:B:42:GLN:HG2	1.59	0.41
1:B:322:PRO:HB2	1:B:337:GLY:HA3	2.02	0.40
1:D:362:TYR:HB3	1:D:366:LEU:HG	2.04	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	478/502 (95%)	462 (97%)	14 (3%)	2 (0%)	39	20
1	B	477/502 (95%)	465 (98%)	10 (2%)	2 (0%)	39	20
1	C	467/502 (93%)	447 (96%)	18 (4%)	2 (0%)	39	20
1	D	489/502 (97%)	476 (97%)	11 (2%)	2 (0%)	39	20
All	All	1911/2008 (95%)	1850 (97%)	53 (3%)	8 (0%)	39	20

All (8) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	347	ASP
1	C	347	ASP
1	D	111	SER
1	D	347	ASP
1	A	111	SER
1	B	111	SER
1	B	347	ASP
1	C	111	SER

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	408/420 (97%)	394 (97%)	14 (3%)	44	21
1	B	408/420 (97%)	398 (98%)	10 (2%)	55	34
1	C	399/420 (95%)	389 (98%)	10 (2%)	55	34
1	D	419/420 (100%)	404 (96%)	15 (4%)	42	19
All	All	1634/1680 (97%)	1585 (97%)	49 (3%)	51	26

All (49) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	66	VAL
1	A	112	TRP
1	A	123	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	226[A]	LYS
1	A	226[B]	LYS
1	A	227	LYS
1	A	271	TYR
1	A	320	ASN
1	A	364	LYS
1	A	377	TYR
1	A	390	ASN
1	A	420	ILE
1	A	449	ASN
1	A	486	TYR
1	B	41	LYS
1	B	42	GLN
1	B	117	LYS
1	B	228	ASP
1	B	271	TYR
1	B	320	ASN
1	B	377	TYR
1	B	474[A]	ASP
1	B	474[B]	ASP
1	B	486	TYR
1	C	112[A]	TRP
1	C	112[B]	TRP
1	C	117[A]	LYS
1	C	117[B]	LYS
1	C	142	MET
1	C	224	ASP
1	C	271	TYR
1	C	320	ASN
1	C	377	TYR
1	C	486	TYR
1	D	42	GLN
1	D	66	VAL
1	D	73	ARG
1	D	226	LYS
1	D	227	LYS
1	D	228	ASP
1	D	271	TYR
1	D	368	GLN
1	D	377	TYR
1	D	387	LYS
1	D	420	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	438[A]	ARG
1	D	438[B]	ARG
1	D	486	TYR
1	D	488	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (14) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	295	GLN
1	A	368	GLN
1	B	30	ASN
1	B	42	GLN
1	B	64	GLN
1	B	225	ASN
1	B	449	ASN
1	C	64	GLN
1	C	135	GLN
1	C	295	GLN
1	D	42	GLN
1	D	64	GLN
1	D	478	GLN
1	D	492	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 19 ligands modelled in this entry, 6 are unknown and 6 are monoatomic - leaving 7 for Mogul analysis.

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	GOL	A	1	-	5,5,5	0.22	0	5,5,5	0.70	0
4	SO4	A	520	-	4,4,4	0.36	0	6,6,6	0.06	0
2	GOL	B	1	-	5,5,5	1.65	1 (20%)	5,5,5	0.86	0
4	SO4	B	2	-	4,4,4	0.47	0	6,6,6	0.18	0
2	GOL	B	519	-	5,5,5	0.30	0	5,5,5	0.36	0
4	SO4	C	3	-	4,4,4	0.31	0	6,6,6	0.07	0
4	SO4	D	4	-	4,4,4	1.43	0	6,6,6	0.37	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	GOL	A	1	-	-	0/4/4/4	0/0/0/0
4	SO4	A	520	-	-	0/0/0/0	0/0/0/0
2	GOL	B	1	-	-	0/4/4/4	0/0/0/0
4	SO4	B	2	-	-	0/0/0/0	0/0/0/0
2	GOL	B	519	-	-	0/4/4/4	0/0/0/0
4	SO4	C	3	-	-	0/0/0/0	0/0/0/0
4	SO4	D	4	-	-	0/0/0/0	0/0/0/0

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1	GOL	O2-C2	-2.59	1.35	1.43

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1	GOL	2	0
2	B	1	GOL	6	0
4	B	2	SO4	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, 95th 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 ²)	Q<0.9
1	A	468/502 (93%)	0.23	22 (4%) 35 39	18, 30, 51, 94	0
1	B	474/502 (94%)	0.07	14 (2%) 54 58	17, 28, 51, 88	0
1	C	467/502 (93%)	0.50	41 (8%) 12 13	21, 37, 66, 87	0
1	D	474/502 (94%)	0.34	31 (6%) 22 24	19, 30, 52, 84	0
All	All	1883/2008 (93%)	0.28	108 (5%) 27 29	17, 31, 56, 94	0

All (108) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	35	ALA	9.0
1	C	412	ALA	6.6
1	B	34	PRO	5.5
1	C	303	ILE	5.5
1	A	227	LYS	5.2
1	D	34	PRO	5.2
1	C	405	VAL	5.1
1	C	311	ASN	4.7
1	B	229	SER	4.7
1	D	229	SER	4.6
1	B	228	ASP	4.6
1	B	30	ASN	4.5
1	C	391	PHE	4.1
1	B	33[A]	SER	4.0
1	C	229	SER	3.9
1	C	228	ASP	3.7
1	C	132	LEU	3.7
1	D	323	VAL	3.7
1	C	304	LYS	3.5
1	B	41	LYS	3.4
1	D	261	ALA	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	134	ALA	3.3
1	D	184	LYS	3.2
1	A	354	ILE	3.2
1	A	322	PRO	3.1
1	C	477	LYS	3.1
1	D	30	ASN	3.1
1	D	183	ALA	3.1
1	B	112	TRP	3.1
1	D	90	LEU	3.0
1	B	227	LYS	3.0
1	A	183	ALA	3.0
1	C	404	LEU	3.0
1	C	484	LEU	3.0
1	A	49	MET	3.0
1	D	49	MET	2.9
1	A	112	TRP	2.8
1	C	364	LYS	2.8
1	D	322	PRO	2.8
1	B	32	PHE	2.7
1	A	90	LEU	2.7
1	C	413	LEU	2.7
1	D	265	GLN	2.7
1	C	42	GLN	2.7
1	C	366	LEU	2.7
1	A	323	VAL	2.6
1	A	48	ILE	2.6
1	D	225	ASN	2.6
1	C	43	PRO	2.6
1	C	392	LYS	2.6
1	C	480	TYR	2.6
1	C	305	ALA	2.5
1	D	227	LYS	2.5
1	A	229	SER	2.5
1	D	187	PRO	2.5
1	D	321	GLY	2.5
1	D	33	SER	2.5
1	D	189	TYR	2.4
1	C	141	MET	2.4
1	C	306	MET	2.4
1	A	50	ALA	2.4
1	A	364	LYS	2.4
1	A	356	VAL	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	363	GLY	2.4
1	C	137	TYR	2.4
1	C	508	TRP	2.4
1	C	66	VAL	2.4
1	D	412	ALA	2.4
1	C	473	PRO	2.3
1	A	321	GLY	2.3
1	B	115	SER	2.3
1	C	395	THR	2.3
1	D	179	THR	2.3
1	A	241	VAL	2.3
1	B	49	MET	2.3
1	D	180	LEU	2.3
1	C	312	THR	2.3
1	A	184	LYS	2.3
1	C	390	ASN	2.3
1	D	190	GLY	2.3
1	D	162[A]	MET	2.3
1	C	230	LYS	2.2
1	B	508	TRP	2.2
1	D	354	ILE	2.2
1	D	241	VAL	2.2
1	C	112[A]	TRP	2.2
1	C	362	TYR	2.2
1	C	76	GLN	2.2
1	A	89	PRO	2.1
1	A	47	ILE	2.1
1	A	182	ASN	2.1
1	D	32	PHE	2.1
1	A	86	ALA	2.1
1	C	131	LEU	2.1
1	D	98	LEU	2.1
1	C	368	GLN	2.0
1	C	314	VAL	2.0
1	D	356	VAL	2.0
1	D	91	SER	2.0
1	C	189	TYR	2.0
1	B	237	ALA	2.0
1	A	53	LEU	2.0
1	B	449	ASN	2.0
1	D	320	ASN	2.0
1	A	238	PHE	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	154	THR	2.0
1	C	389	MET	2.0
1	D	147	LEU	2.0

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

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

6.3 Carbohydrates [i](#)

There are no carbohydrates in this entry.

6.4 Ligands [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, 95th 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 ²)	Q<0.9
5	UNL	C	519	4/-	0.80	0.33	12.44	34,52,57,58	0
4	SO4	D	4	5/5	0.95	0.52	10.11	25,34,47,55	5
4	SO4	A	520	5/5	0.96	0.42	7.86	23,26,37,46	5
4	SO4	B	2	5/5	0.98	0.31	7.59	24,26,40,41	5
3	NA	C	518	1/1	0.96	0.20	5.43	30,30,30,30	0
2	GOL	B	1	6/6	0.86	0.19	5.37	24,36,59,59	0
5	UNL	D	519	4/-	0.87	0.21	4.72	52,56,58,59	0
5	UNL	B	520	4/-	0.88	0.19	4.40	25,37,44,55	0
2	GOL	B	519	6/6	0.85	0.22	4.33	48,57,62,70	0
4	SO4	C	3	5/5	0.98	0.20	4.19	23,31,33,50	5
3	NA	D	1	1/1	0.97	0.25	3.77	27,27,27,27	0
5	UNL	D	518	4/-	0.87	0.14	3.14	31,42,50,56	0
3	NA	C	1	1/1	0.98	0.20	3.11	19,19,19,19	0
3	NA	A	518	1/1	0.99	0.18	2.84	19,19,19,19	0
5	UNL	A	521	4/-	0.78	0.15	2.14	30,36,44,45	0
2	GOL	A	1	6/6	0.89	0.15	2.01	24,46,52,52	0
3	NA	B	518	1/1	0.98	0.13	-0.18	21,21,21,21	0
3	NA	A	519	1/1	0.99	0.11	-4.01	24,24,24,24	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
5	UNL	D	520	4/-	0.93	0.22	-	54,57,61,62	0

6.5 Other polymers [i](#)

There are no such residues in this entry.