Full wwPDB X-ray Structure Validation Report (i)

Oct 12, 2024 – 11:26 pm BST

PDB ID	:	4CYR
Title	:	G4 mutant of PAS, arylsulfatase from Pseudomonas Aeruginosa
Authors	:	Miton, C.M.; Jonas, S.; Mohammed, M.F.; Fischer, G.; Loo, B.v.; Kintses, B.; Hyvonen, M.; Tokuriki, N.; Hollfelder, F.
Deposited on	:	2014-04-14
Resolution	:	1.72 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

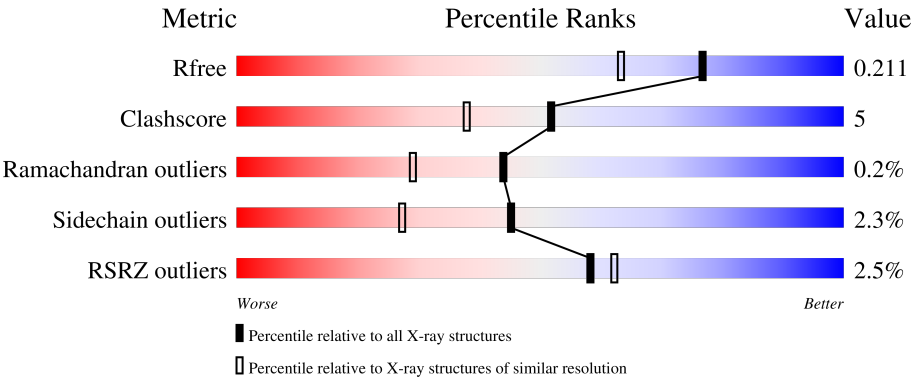
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.72 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}	164625	7106 (1.74-1.70)
Clashscore	180529	7746 (1.74-1.70)
Ramachandran outliers	177936	7654 (1.74-1.70)
Sidechain outliers	177891	7654 (1.74-1.70)
RSRZ outliers	164620	7104 (1.74-1.70)

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 of the lower bar indicate the fraction of residues that contain outliers for >=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 <=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	536	2% 87% 10%
1	B	536	3% 85% 11%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 9637 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.

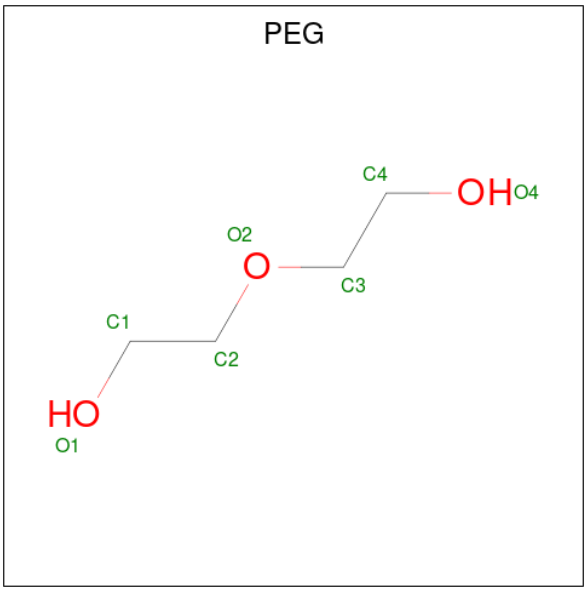
- Molecule 1 is a protein called ARYLSULFATASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	526	Total	C	N	O	S	0	1	0
			4162	2649	735	768	10			
1	B	526	Total	C	N	O	S	0	3	0
			4187	2662	743	772	10			

There are 12 discrepancies between the modelled and reference sequences:

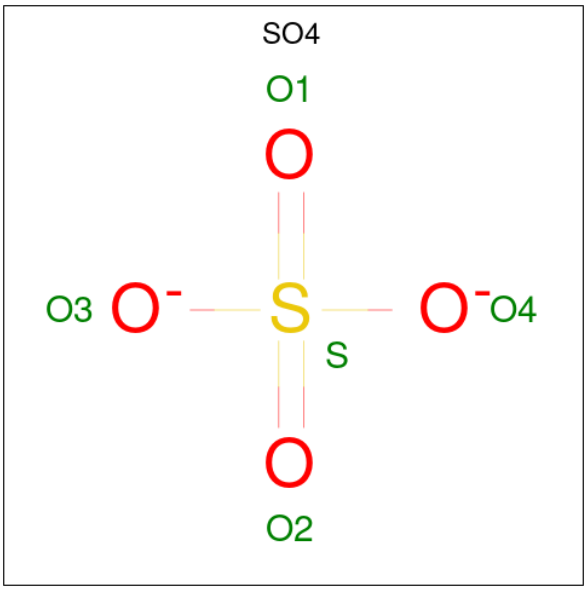
Chain	Residue	Modelled	Actual	Comment	Reference
A	22	THR	ALA	engineered mutation	UNP P51691
A	50	ALA	THR	engineered mutation	UNP P51691
A	134	ALA	SER	engineered mutation	UNP P51691
A	337	ASP	GLY	engineered mutation	UNP P51691
A	461	GLY	GLU	engineered mutation	UNP P51691
A	523	ASP	GLU	engineered mutation	UNP P51691
B	22	THR	ALA	engineered mutation	UNP P51691
B	50	ALA	THR	engineered mutation	UNP P51691
B	134	ALA	SER	engineered mutation	UNP P51691
B	337	ASP	GLY	engineered mutation	UNP P51691
B	461	GLY	GLU	engineered mutation	UNP P51691
B	523	ASP	GLU	engineered mutation	UNP P51691

- Molecule 2 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

Continued on next page...

Continued from previous page...

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

- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Ca 1 1	0	0
4	B	1	Total Ca 1 1	0	0

- Molecule 5 is water.

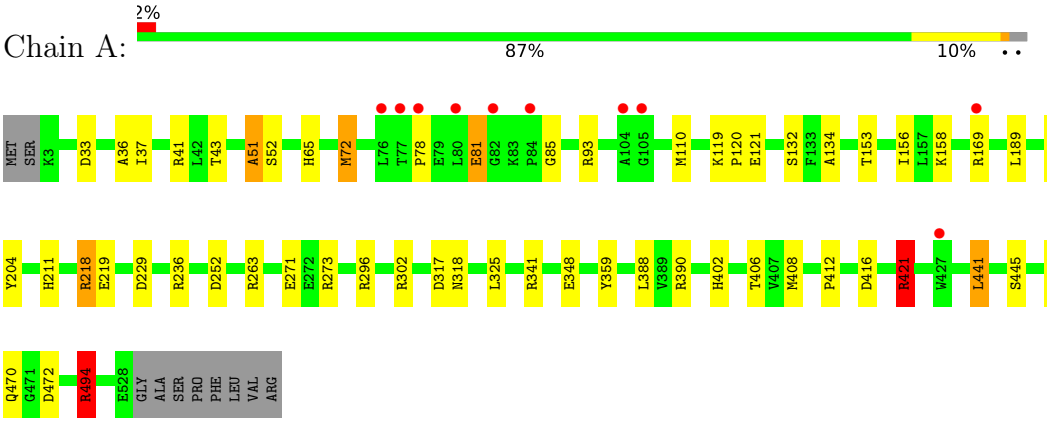
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	636	Total O 636 636	0	0
5	B	579	Total O 579 579	0	0

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.

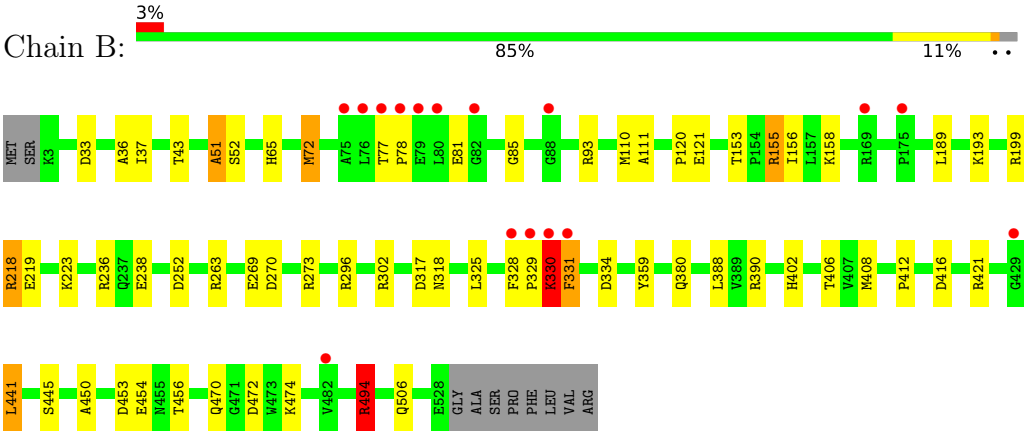
• Molecule 1: ARYLSULFATASE

Chain A:



• Molecule 1: ARYLSULFATASE

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	188.50Å 67.68Å 89.52Å 90.00° 93.93° 90.00°	Depositor
Resolution (Å)	94.03 – 1.72 94.03 – 1.72	Depositor EDS
% Data completeness (in resolution range)	99.3 (94.03-1.72) 99.3 (94.03-1.72)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.38 (at 1.72Å)	Xtriage
Refinement program	REFMAC 5.8.0069	Depositor
R, R_{free}	0.176 , 0.203 0.187 , 0.211	Depositor DCC
R_{free} test set	4041 reflections (3.40%)	wwPDB-VP
Wilson B-factor (Å^2)	15.1	Xtriage
Anisotropy	0.044	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 51.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	9637	wwPDB-VP
Average B, all atoms (Å^2)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.19% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: CA, PEG, SO4, DDZ

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	1.02	1/4270 (0.0%)	1.04	20/5807 (0.3%)
1	B	1.04	2/4295 (0.0%)	1.06	22/5839 (0.4%)
All	All	1.03	3/8565 (0.0%)	1.05	42/11646 (0.4%)

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	B	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	348	GLU	CG-CD	5.68	1.60	1.51
1	B	238	GLU	CD-OE1	-5.30	1.19	1.25
1	B	121	GLU	CD-OE1	5.04	1.31	1.25

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	494	ARG	NE-CZ-NH1	9.76	125.18	120.30
1	A	296	ARG	NE-CZ-NH2	-9.48	115.56	120.30
1	A	296	ARG	NE-CZ-NH1	9.19	124.89	120.30
1	A	494	ARG	NE-CZ-NH1	8.88	124.74	120.30
1	B	296	ARG	NE-CZ-NH1	8.80	124.70	120.30
1	B	296	ARG	NE-CZ-NH2	-8.15	116.23	120.30
1	B	390	ARG	NE-CZ-NH1	7.51	124.05	120.30
1	A	341	ARG	NE-CZ-NH1	7.30	123.95	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	390	ARG	NE-CZ-NH1	7.11	123.86	120.30
1	B	199	ARG	NE-CZ-NH2	-7.06	116.77	120.30
1	B	273	ARG	NE-CZ-NH1	6.86	123.73	120.30
1	B	390	ARG	NE-CZ-NH2	-6.70	116.95	120.30
1	B	334	ASP	CB-CG-OD2	6.61	124.25	118.30
1	A	390	ARG	NE-CZ-NH2	-6.60	117.00	120.30
1	B	236	ARG	NE-CZ-NH2	-6.30	117.15	120.30
1	A	33	ASP	CB-CG-OD2	-6.25	112.67	118.30
1	A	472	ASP	CB-CG-OD1	6.25	123.92	118.30
1	A	199	ARG	NE-CZ-NH2	-6.24	117.18	120.30
1	B	263	ARG	NE-CZ-NH2	-6.07	117.27	120.30
1	B	273	ARG	NE-CZ-NH2	-6.06	117.27	120.30
1	A	494	ARG	NE-CZ-NH2	-6.01	117.29	120.30
1	A	273	ARG	NE-CZ-NH2	-5.91	117.34	120.30
1	B	494	ARG	NE-CZ-NH2	-5.91	117.35	120.30
1	B	155	ARG	NE-CZ-NH1	5.85	123.22	120.30
1	B	472	ASP	CB-CG-OD1	5.72	123.45	118.30
1	A	33	ASP	CB-CG-OD1	5.64	123.38	118.30
1	A	236	ARG	NE-CZ-NH1	5.59	123.09	120.30
1	A	199	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	A	252	ASP	CB-CG-OD1	5.52	123.27	118.30
1	A	41	ARG	NE-CZ-NH1	5.51	123.06	120.30
1	B	236	ARG	NE-CZ-NH1	5.48	123.04	120.30
1	B	252	ASP	CB-CG-OD1	5.46	123.21	118.30
1	A	273	ARG	NE-CZ-NH1	5.24	122.92	120.30
1	A	229	ASP	CB-CG-OD2	-5.23	113.59	118.30
1	B	453	ASP	CB-CG-OD2	-5.16	113.66	118.30
1	B	199	ARG	NE-CZ-NH1	5.15	122.87	120.30
1	B	33	ASP	CB-CG-OD1	5.14	122.93	118.30
1	A	421	ARG	NE-CZ-NH1	5.10	122.85	120.30
1	B	33	ASP	CB-CG-OD2	-5.07	113.74	118.30
1	B	218[A]	ARG	NE-CZ-NH2	-5.04	117.78	120.30
1	B	218[B]	ARG	NE-CZ-NH2	-5.04	117.78	120.30
1	A	263	ARG	NE-CZ-NH2	-5.02	117.79	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	330	LYS	Peptide

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	4162	0	4028	39	0
1	B	4187	0	4054	47	1
2	A	21	0	30	2	0
3	A	30	0	0	1	0
3	B	20	0	0	1	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	636	0	0	6	2
5	B	579	0	0	6	1
All	All	9637	0	8112	82	2

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 5.

All (82) 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:169:ARG:NH2	5:A:2251:HOH:O	1.96	0.96
1:B:329:PRO:O	1:B:330:LYS:O	1.92	0.86
1:A:421:ARG:O	5:A:2519:HOH:O	2.10	0.70
1:B:218[B]:ARG:HH11	1:B:218[B]:ARG:HG2	1.57	0.70
1:B:52:SER:H	1:B:72:MET:HE1	1.58	0.69
1:A:421:ARG:HH11	1:A:421:ARG:HB3	1.58	0.69
1:B:223:LYS:NZ	5:B:2277:HOH:O	2.28	0.66
1:A:110[B]:MET:HG3	1:A:204:TYR:CD2	2.31	0.64
1:A:454:GLU:HG2	5:A:2559:HOH:O	1.97	0.64
1:B:155:ARG:HD2	1:B:331:PHE:HD2	1.64	0.63
1:A:406:THR:CG2	1:A:408:MET:H	2.12	0.62
1:B:155:ARG:HD2	1:B:331:PHE:CD2	2.34	0.62
1:B:406:THR:CG2	1:B:408:MET:H	2.14	0.61
1:A:52:SER:H	1:A:72:MET:HE1	1.67	0.59
1:A:121:GLU:CB	2:A:1532:PEG:H31	2.34	0.58
1:B:51:DDZ:HB	1:B:317:ASP:OD2	2.02	0.58
1:A:110[B]:MET:HG3	1:A:204:TYR:HD2	1.67	0.57
1:A:406:THR:HG22	1:A:408:MET:H	1.68	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:494:ARG:HG2	1:A:494:ARG:HH11	1.70	0.56
1:B:380:GLN:HE21	1:B:474:LYS:NZ	2.02	0.56
1:B:330:LYS:HD3	5:B:2410:HOH:O	2.04	0.56
1:A:51:DDZ:HB	1:A:317:ASP:OD2	2.03	0.56
1:B:454:GLU:HG2	5:B:2515:HOH:O	2.04	0.56
1:B:189:LEU:HG	1:B:193:LYS:HE3	1.88	0.54
1:A:156:ILE:CD1	1:A:325:LEU:HD13	2.38	0.54
1:B:421:ARG:HG3	1:B:421:ARG:HH11	1.73	0.54
1:B:218[B]:ARG:HG2	1:B:218[B]:ARG:NH1	2.24	0.53
1:B:380:GLN:HE21	1:B:474:LYS:HZ1	1.55	0.53
1:B:406:THR:HG22	1:B:408:MET:H	1.73	0.53
1:A:406:THR:HG22	1:A:408:MET:N	2.24	0.53
1:B:155:ARG:CD	1:B:331:PHE:HD2	2.22	0.53
1:A:52:SER:H	1:A:72:MET:CE	2.22	0.52
1:A:153:THR:O	1:A:158:LYS:HE2	2.10	0.52
1:B:52:SER:H	1:B:72:MET:CE	2.23	0.52
1:B:329:PRO:O	1:B:330:LYS:C	2.48	0.52
1:B:156:ILE:CD1	1:B:325:LEU:HD13	2.39	0.51
1:A:189:LEU:HG	1:A:193:LYS:HE3	1.93	0.51
1:B:110[B]:MET:HG2	1:B:111:ALA:N	2.25	0.51
1:B:93:ARG:HD2	5:B:2082:HOH:O	2.12	0.50
1:A:156:ILE:HD13	1:A:325:LEU:HD13	1.93	0.50
1:A:302:ARG:NH1	1:B:37:ILE:HD12	2.27	0.50
1:B:406:THR:HG22	1:B:408:MET:N	2.26	0.49
1:A:37:ILE:HD12	1:B:302:ARG:NH1	2.28	0.49
1:B:153:THR:O	1:B:158:LYS:HE2	2.12	0.48
1:A:494:ARG:HH11	1:A:494:ARG:CG	2.26	0.48
1:B:65:HIS:O	1:B:85:GLY:HA3	2.13	0.48
1:B:156:ILE:HD13	1:B:325:LEU:HD13	1.95	0.47
1:A:65:HIS:O	1:A:85:GLY:HA3	2.15	0.47
1:A:211:HIS:NE2	3:A:1539:SO4:O3	2.33	0.47
1:B:494:ARG:HH11	1:B:494:ARG:HG3	1.79	0.47
1:A:78:PRO:O	1:A:81:GLU:HB2	2.15	0.47
1:B:77:THR:O	1:B:81:GLU:HB2	2.16	0.46
1:A:37:ILE:HD12	1:B:302:ARG:CZ	2.46	0.46
1:B:211:HIS:NE2	3:B:1530:SO4:O2	2.33	0.46
1:A:456:THR:O	1:A:470:GLN:HA	2.17	0.45
1:B:36:ALA:HB2	1:B:388:LEU:HD22	1.99	0.45
1:B:78:PRO:HA	1:B:81:GLU:OE1	2.16	0.44
1:B:328:PHE:O	1:B:329:PRO:C	2.56	0.44
1:B:441:LEU:HD22	1:B:441:LEU:HA	1.79	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:302:ARG:CZ	1:B:37:ILE:HD12	2.48	0.44
1:B:450:ALA:HB1	5:B:2462:HOH:O	2.16	0.44
1:A:110[B]:MET:HE1	1:A:132:SER:HB2	1.99	0.44
1:B:456:THR:O	1:B:470:GLN:HA	2.18	0.43
1:B:43:THR:OG1	1:B:402:HIS:CD2	2.72	0.43
1:B:218[A]:ARG:HG2	1:B:219:GLU:N	2.34	0.43
1:B:93:ARG:HB3	5:B:2122:HOH:O	2.18	0.43
1:B:406:THR:HG23	1:B:408:MET:H	1.83	0.43
1:A:271:GLU:HB2	5:A:2397:HOH:O	2.18	0.42
1:B:110[B]:MET:HB2	1:B:204:TYR:HB3	2.02	0.42
1:A:93:ARG:HB3	5:A:2145:HOH:O	2.19	0.42
1:A:36:ALA:HB2	1:A:388:LEU:HD22	2.02	0.42
1:A:119:LYS:NZ	5:A:2185:HOH:O	2.18	0.42
1:A:441:LEU:HD22	1:A:441:LEU:HA	1.78	0.41
1:B:441:LEU:O	1:B:445:SER:HB3	2.19	0.41
1:A:43:THR:OG1	1:A:402:HIS:CD2	2.73	0.41
1:A:110[B]:MET:CE	1:A:134:ALA:HB2	2.50	0.41
1:A:121:GLU:HB3	2:A:1532:PEG:H31	2.02	0.41
1:A:412:PRO:O	1:A:416:ASP:HB2	2.21	0.41
1:B:218[B]:ARG:NH1	1:B:218[B]:ARG:CG	2.84	0.41
1:A:218:ARG:HG2	1:A:219:GLU:N	2.35	0.40
1:A:441:LEU:O	1:A:445:SER:HB3	2.22	0.40
1:B:412:PRO:O	1:B:416:ASP:HB2	2.21	0.40

All (2) 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)
5:A:2136:HOH:O	5:B:2110:HOH:O[4_555]	1.82	0.38
1:B:438[A]:ARG:NH2	5:A:2558:HOH:O[4_546]	2.05	0.15

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	524/536 (98%)	508 (97%)	16 (3%)	0	100 100
1	B	526/536 (98%)	509 (97%)	15 (3%)	2 (0%)	30 17
All	All	1050/1072 (98%)	1017 (97%)	31 (3%)	2 (0%)	44 29

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	330	LYS
1	B	331	PHE

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	418/428 (98%)	409 (98%)	9 (2%)	47 29
1	B	421/428 (98%)	411 (98%)	10 (2%)	44 25
All	All	839/856 (98%)	820 (98%)	19 (2%)	45 27

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	72	MET
1	A	81	GLU
1	A	120	PRO
1	A	218	ARG
1	A	318	ASN
1	A	359	TYR
1	A	421	ARG
1	A	441	LEU
1	A	494	ARG
1	B	72	MET
1	B	120	PRO
1	B	269	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	270	ASP
1	B	318	ASN
1	B	359	TYR
1	B	435	PRO
1	B	441	LEU
1	B	494	ARG
1	B	506	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	402	HIS
1	A	506	GLN
1	B	89	HIS
1	B	380	GLN
1	B	402	HIS
1	B	506	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

2 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	DDZ	B	51	1	4,6,7	0.92	0	3,7,9	3.13	2 (66%)
1	DDZ	A	51	1	4,6,7	0.84	0	3,7,9	3.16	2 (66%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the

Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	DDZ	B	51	1	-	2/2/6/8	-
1	DDZ	A	51	1	-	2/2/6/8	-

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	51	DDZ	OG2-CB-OG1	4.83	119.78	111.27
1	B	51	DDZ	OG2-CB-OG1	4.67	119.50	111.27
1	B	51	DDZ	O-C-CA	-2.71	117.68	124.78
1	A	51	DDZ	O-C-CA	-2.50	118.23	124.78

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	51	DDZ	N-CA-CB-OG1
1	A	51	DDZ	N-CA-CB-OG2
1	B	51	DDZ	N-CA-CB-OG1
1	B	51	DDZ	N-CA-CB-OG2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	B	51	DDZ	1	0
1	A	51	DDZ	1	0

5.5 Carbohydrates (i)

There are no oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 15 ligands modelled in this entry, 2 are monoatomic - leaving 13 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	$\#\|Z\| > 2$
3	SO4	A	1536	-	4,4,4	0.55	0	6,6,6	0.89	0
3	SO4	A	1535	-	4,4,4	0.41	0	6,6,6	0.50	0
2	PEG	A	1531	-	6,6,6	0.72	0	5,5,5	0.85	0
3	SO4	A	1538	-	4,4,4	0.45	0	6,6,6	0.42	0
3	SO4	B	1534	-	4,4,4	0.47	0	6,6,6	0.26	0
2	PEG	A	1530	-	6,6,6	0.51	0	5,5,5	0.31	0
3	SO4	A	1533	-	4,4,4	0.42	0	6,6,6	0.99	0
3	SO4	B	1530	-	4,4,4	0.51	0	6,6,6	1.10	0
3	SO4	B	1531	-	4,4,4	0.64	0	6,6,6	1.48	1 (16%)
3	SO4	B	1532	-	4,4,4	0.67	0	6,6,6	1.06	0
3	SO4	A	1534	-	4,4,4	0.29	0	6,6,6	1.07	0
2	PEG	A	1532	-	6,6,6	0.78	0	5,5,5	1.36	1 (20%)
3	SO4	A	1539	-	4,4,4	0.82	0	6,6,6	0.61	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	PEG	A	1532	-	-	2/4/4/4	-
2	PEG	A	1531	-	-	2/4/4/4	-
2	PEG	A	1530	-	-	2/4/4/4	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1531	SO4	O4-S-O3	3.07	122.16	109.06
2	A	1532	PEG	O2-C2-C1	2.04	119.02	110.07

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1532	PEG	C1-C2-O2-C3
2	A	1530	PEG	O1-C1-C2-O2
2	A	1531	PEG	O2-C3-C4-O4
2	A	1532	PEG	O1-C1-C2-O2
2	A	1530	PEG	O2-C3-C4-O4
2	A	1531	PEG	C1-C2-O2-C3

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	1530	SO4	1	0
2	A	1532	PEG	2	0
3	A	1539	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, 95^{th} 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	525/536 (97%)	0.00	10 (1%)	66	70	7, 16, 36, 71	5 (0%)
1	B	525/536 (97%)	0.03	16 (3%)	52	57	7, 16, 38, 83	6 (1%)
All	All	1050/1072 (97%)	0.02	26 (2%)	58	62	7, 16, 37, 83	11 (1%)

All (26) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	331	PHE	4.9
1	B	77	THR	4.8
1	B	78	PRO	4.7
1	B	80	LEU	4.3
1	B	76	LEU	4.0
1	A	80	LEU	3.5
1	B	79	GLU	3.5
1	B	329	PRO	3.4
1	B	328	PHE	3.1
1	A	77	THR	2.9
1	A	78	PRO	2.9
1	B	82	GLY	2.8
1	B	75	ALA	2.7
1	B	429	GLY	2.5
1	B	330	LYS	2.5
1	A	104	ALA	2.4
1	A	169	ARG	2.4
1	B	482	VAL	2.4
1	A	82	GLY	2.4
1	A	84	PRO	2.3
1	B	88	GLY	2.3
1	B	169	ARG	2.2
1	A	105	GLY	2.1
1	A	76	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	427	TRP	2.1
1	B	175	PRO	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. The B-factors column lists the minimum, median, 95^{th} 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^2)	Q<0.9
1	DDZ	A	51	7/8	0.95	0.06	10,12,13,14	0
1	DDZ	B	51	7/8	0.96	0.05	11,12,13,13	0

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^{th} 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^2)	Q<0.9
3	SO4	A	1536	5/5	0.71	0.15	70,72,79,81	0
2	PEG	A	1532	7/7	0.75	0.20	44,48,53,61	0
2	PEG	A	1530	7/7	0.76	0.21	49,50,54,56	0
3	SO4	B	1534	5/5	0.78	0.14	67,69,72,81	0
3	SO4	A	1538	5/5	0.79	0.14	70,75,79,90	0
3	SO4	A	1535	5/5	0.83	0.13	52,55,67,68	0
2	PEG	A	1531	7/7	0.85	0.13	37,39,40,42	0
3	SO4	B	1530	5/5	0.86	0.17	42,44,53,60	0
3	SO4	A	1539	5/5	0.86	0.17	47,48,55,64	0
3	SO4	A	1534	5/5	0.95	0.14	24,28,37,39	0
3	SO4	B	1532	5/5	0.96	0.11	26,34,38,40	0
3	SO4	A	1533	5/5	0.97	0.06	21,22,23,28	0
3	SO4	B	1531	5/5	0.97	0.07	23,24,25,27	0
4	CA	A	1537	1/1	0.99	0.12	21,21,21,21	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
4	CA	B	1533	1/1	0.99	0.11	23,23,23,23	0

6.5 Other polymers (i)

There are no such residues in this entry.