wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 16, 2023 – 07:45 PM EDT

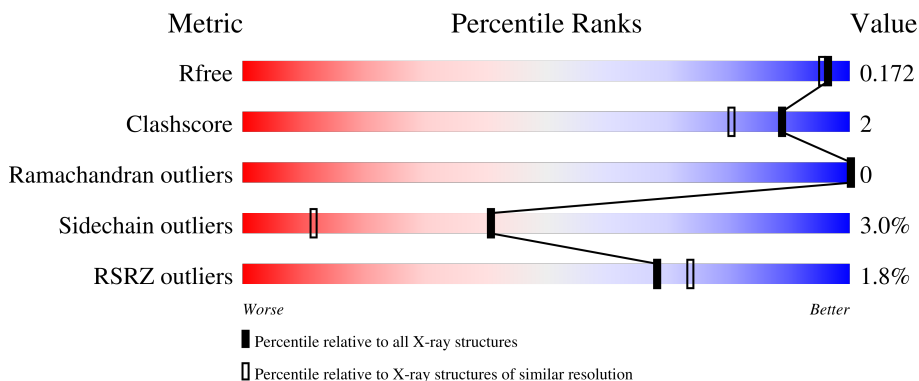|		
PDB ID	:	4WHP
Title	:	Resting Protocatechuate 3,4-dioxygenase (pseudomonas putida) at pH 6.5
Authors	:	Knoot, C.J.; Lipscomb, J.D.
Deposited on	:	2014-09-23
Resolution	:	1.54 Å(reported)

1 Overall quality at a glance ⓘ

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

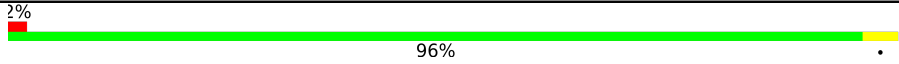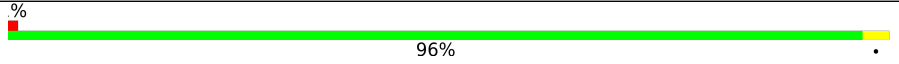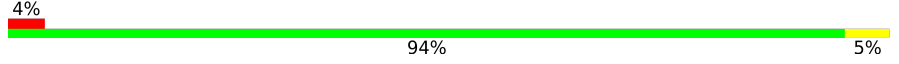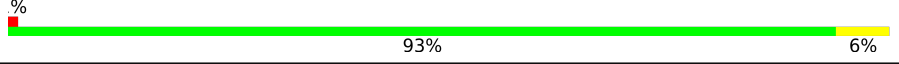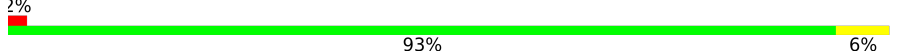The reported resolution of this entry is 1.54 Å.

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	Percentile Ranks	Value
Rfree		0.172
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		3.0%
RSRZ outliers		1.8%

Worse — *Better*

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution
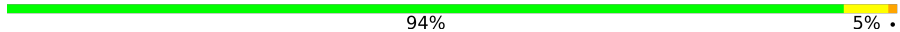
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	2556 (1.56-1.52)
Clashscore	141614	2634 (1.56-1.52)
Ramachandran outliers	138981	2580 (1.56-1.52)
Sidechain outliers	138945	2577 (1.56-1.52)
RSRZ outliers	127900	2524 (1.56-1.52)

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	200	2% / 96% ·
1	C	200	% / 96% ·
1	E	200	4% / 94% 5%
2	D	238	% / 93% 6%
2	F	238	2% / 93% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	B	238	94% 5%

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

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
5	SO4	D	602	-	-	X	-

2 Entry composition ⓘ

There are 8 unique types of molecules in this entry. The entry contains 11573 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 Protocatechuate 3,4-dioxygenase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	200	Total	C	N	O	S	0	2	0
			1593	1007	282	301	3			
1	E	200	Total	C	N	O	S	0	1	0
			1583	1002	278	300	3			
1	C	200	Total	C	N	O	S	0	0	0
			1575	996	277	299	3			

- Molecule 2 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	F	237	Total	C	N	O	S	0	1	0
			1884	1196	344	337	7			
2	D	238	Total	C	N	O	S	0	2	0
			1898	1205	346	339	8			

- Molecule 3 is a protein called Protocatechuate 3,4-dioxygenase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	B	238	Total	C	N	O	S	0	2	0
			1897	1202	346	341	8			

- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	F	1	Total	Fe	0	0
			1	1		
4	D	1	Total	Fe	0	0
			1	1		
4	B	1	Total	Fe	0	0
			1	1		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O_4S).

SO4

O1

O3 O⁻ — S — O⁻ O4

S

O2

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

- Molecule 6 is BICARBONATE ION (three-letter code: BCT) (formula: CHO_3).

BCT

O1

O

OH O3

C

O⁻

O2
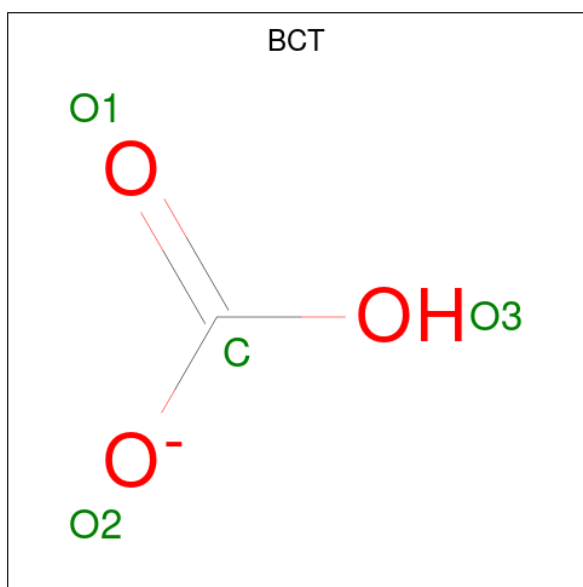
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	C	1	Total 4	C 1	O 3	0	0

- Molecule 7 is BETA-MERCAPTOETHANOL (three-letter code: BME) (formula: C_2H_6OS).

BME
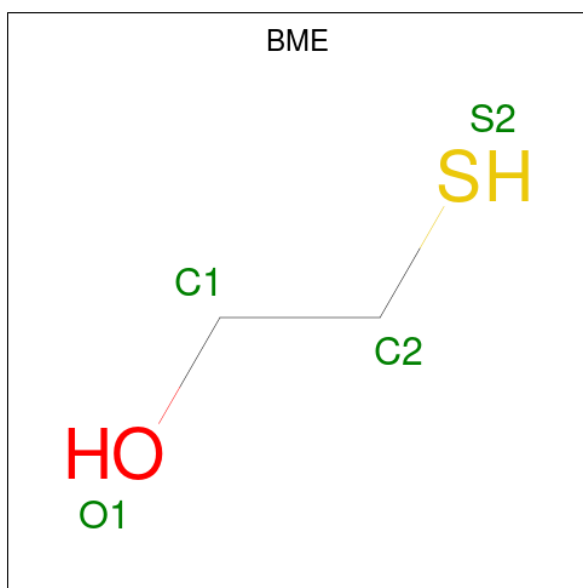
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	O	S		
7	D	1	4	2	1	1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	169	Total 169	O 169	0	0
8	E	131	Total 131	O 131	0	0
8	F	203	Total 203	O 203	0	0
8	C	166	Total 167	O 167	0	1
8	D	218	Total 218	O 218	0	0
8	B	229	Total 229	O 229	0	0

3 Residue-property plots ⓘ

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: Protocatechuate 3,4-dioxygenase alpha chain
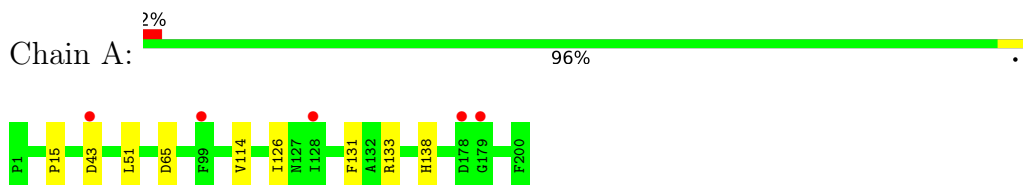
• Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain
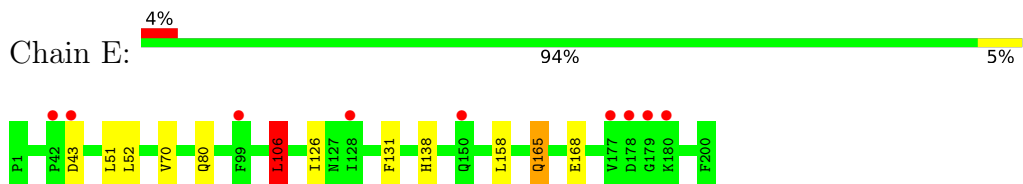
• Molecule 1: Protocatechuate 3,4-dioxygenase alpha chain

• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

• Molecule 2: Protocatechuate 3,4-dioxygenase beta chain

• Molecule 3: Protocatechuate 3,4-dioxygenase beta chain

Chain B:

94%

5%

P301 H361 N368 N369 T395 K411 L416 M422 L430 K473 L478 L485 M488 N497 P498 E499 H534 N537 C538
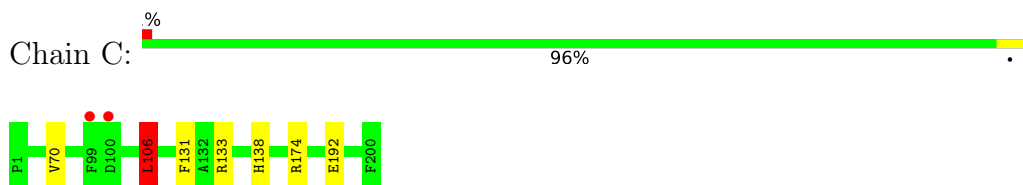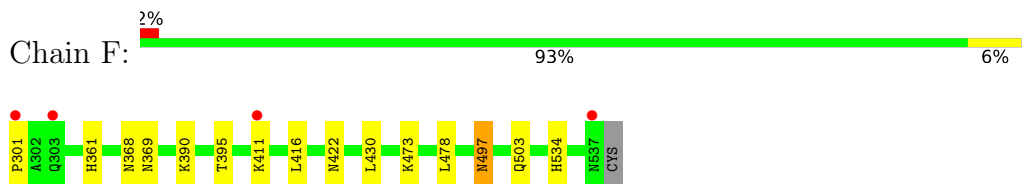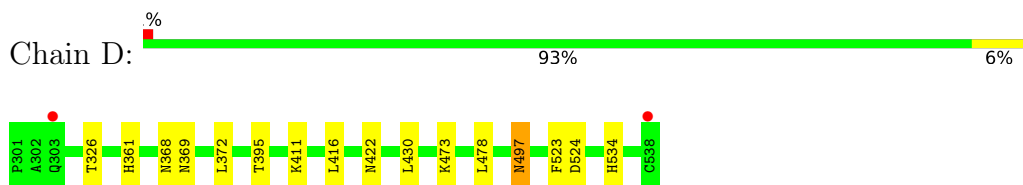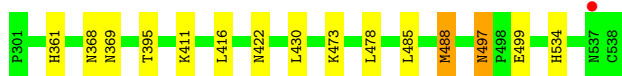
4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	127.92Å 140.77Å 168.19Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.02 – 1.54 32.46 – 1.54	Depositor EDS
% Data completeness (in resolution range)	93.0 (39.02-1.54) 93.0 (32.46-1.54)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.98 (at 1.54Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.140 , 0.172 0.140 , 0.172	Depositor DCC
R_{free} test set	10324 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	15.5	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.42 , 55.4	EDS
L-test for twinning[2]	$<|L|> = 0.48, <L^2> = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11573	wwPDB-VP
Average B, all atoms (Å^2)	21.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.88% 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: MHO, BCT, FE, SO4, BME

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	0.39	0/1633	0.64	0/2223
1	C	0.45	1/1615 (0.1%)	0.68	1/2199 (0.0%)
1	E	0.42	1/1623 (0.1%)	0.65	1/2210 (0.0%)
2	D	0.40	0/1954	0.67	0/2660
2	F	0.38	0/1940	0.67	0/2641
3	B	0.40	0/1943	0.67	0/2644
All	All	0.41	2/10708 (0.0%)	0.67	2/14577 (0.0%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	192	GLU	CD-OE2	-8.24	1.16	1.25
1	E	168	GLU	CD-OE2	-7.08	1.17	1.25

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	106	LEU	CA-CB-CG	5.51	127.97	115.30
1	E	106	LEU	CA-CB-CG	5.34	127.59	115.30

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	1593	0	1530	4	0
1	C	1575	0	1510	9	0
1	E	1583	0	1520	6	0
2	D	1898	0	1857	8	0
2	F	1884	0	1842	7	0
3	B	1897	0	1850	8	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
5	B	5	0	0	0	0
5	D	5	0	0	5	0
5	F	5	0	0	0	0
6	C	4	0	0	0	0
7	D	4	0	6	3	0
8	A	169	0	0	1	0
8	B	229	0	0	0	4
8	C	167	0	0	1	0
8	D	218	0	0	0	4
8	E	131	0	0	0	0
8	F	203	0	0	1	0
All	All	11573	0	10115	39	5

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 2.

The worst 5 of 39 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:C:133:ARG:HG3	5:D:602:SO4:O1	1.57	1.03
1:C:133:ARG:HD2	5:D:602:SO4:O2	1.86	0.76
2:D:524:ASP:H	7:D:603:BME:H22	1.58	0.67
1:C:70:VAL:HG11	1:C:106:LEU:HD21	1.76	0.66
1:E:165:GLN:H	1:E:165:GLN:HE21	1.45	0.65

All (5) 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)
8:D:755:HOH:O	8:B:746:HOH:O[2_555]	0.84	1.36
8:D:749:HOH:O	8:B:758:HOH:O[2_555]	1.27	0.93
8:D:701:HOH:O	8:B:754:HOH:O[2_555]	1.43	0.77

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
8:D:719:HOH:O	8:D:756:HOH:O[4_555]	1.58	0.62
8:B:752:HOH:O	8:B:752:HOH:O[2_555]	2.05	0.15

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	200/200 (100%)	194 (97%)	6 (3%)	0	100	100
1	C	198/200 (99%)	195 (98%)	3 (2%)	0	100	100
1	E	199/200 (100%)	195 (98%)	4 (2%)	0	100	100
2	D	238/238 (100%)	234 (98%)	4 (2%)	0	100	100
2	F	236/238 (99%)	232 (98%)	4 (2%)	0	100	100
3	B	237/238 (100%)	233 (98%)	4 (2%)	0	100	100
All	All	1308/1314 (100%)	1283 (98%)	25 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	165/163 (101%)	162 (98%)	3 (2%)	59	29
1	C	163/163 (100%)	162 (99%)	1 (1%)	86	72
1	E	164/163 (101%)	158 (96%)	6 (4%)	34	7

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	D	204/202 (101%)	194 (95%)	10 (5%)	25	3
2	F	202/202 (100%)	192 (95%)	10 (5%)	24	3
3	B	203/201 (101%)	195 (96%)	8 (4%)	32	6
All	All	1101/1094 (101%)	1063 (96%)	38 (4%)	41	8

5 of 38 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	D	497	ASN
3	B	473	LYS
2	D	534	HIS
3	B	416	LEU
3	B	534	HIS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 19 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	361	HIS
3	B	422	ASN
3	B	497	ASN
3	B	412	ASN
2	F	503	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\#\|Z\| > 2$
3	MHO	B	488	3	7,8,9	1.00	0	4,9,11	1.50	1 (25%)

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
3	MHO	B	488	3	-	0/6/7/9	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	488	MHO	OD1-SD-CE	2.59	111.50	106.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	488	MHO	1	0

5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 8 ligands modelled in this entry, 3 are monoatomic - leaving 5 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).

WORLDWIDE
wwPDB
PROTEIN DATA BANK

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#|Z|>2$	Counts	RMSZ	$\#|Z|>2$
7	BME	D	603	-	3,3,3	0.39	0	1,2,2	0.57	0
5	SO4	B	601	-	4,4,4	0.27	0	6,6,6	0.36	0
5	SO4	F	602	-	4,4,4	0.40	0	6,6,6	0.39	0
6	BCT	C	301	-	2,3,3	0.68	0	2,3,3	0.92	0
5	SO4	D	602	-	4,4,4	0.28	0	6,6,6	0.78	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
7	BME	D	603	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	D	603	BME	O1-C1-C2-S2

There are no ring outliers.

2 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	D	603	BME	3	0
5	D	602	SO4	5	0

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	200/200 (100%)	-0.27	5 (2%)	57	63	14, 19, 32, 52	0
1	C	200/200 (100%)	-0.38	2 (1%)	82	85	14, 19, 31, 48	0
1	E	200/200 (100%)	0.05	9 (4%)	33	38	15, 23, 38, 59	0
2	D	238/238 (100%)	-0.44	2 (0%)	86	88	15, 18, 27, 42	0
2	F	237/238 (99%)	-0.34	4 (1%)	70	74	15, 19, 29, 44	0
3	B	237/238 (99%)	-0.42	1 (0%)	92	94	15, 17, 27, 37	0
All	All	1312/1314 (99%)	-0.31	23 (1%)	68	74	14, 19, 32, 59	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	537	ASN	4.0
1	E	178	ASP	3.9
2	F	301	PRO	3.7
1	E	43	ASP	3.6
1	E	128	ILE	3.5

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. 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	B-factors(Å^2)	Q<0.9
3	MHO	B	488	9/10	0.97	0.08	17,20,29,46	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, 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	B-factors($Å^2$)	Q<0.9
7	BME	D	603	4/4	0.66	0.19	22,24,31,35	0
6	BCT	C	301	4/4	0.94	0.14	32,32,44,45	0
5	SO4	D	602	5/5	0.97	0.10	28,30,33,34	0
5	SO4	F	602	5/5	0.98	0.07	27,30,34,34	0
5	SO4	B	601	5/5	0.99	0.07	24,27,31,32	0
4	FE	B	602	1/1	1.00	0.06	17,17,17,17	1
4	FE	F	601	1/1	1.00	0.06	17,17,17,17	1
4	FE	D	601	1/1	1.00	0.05	17,17,17,17	1

6.5 Other polymers (i)

There are no such residues in this entry.