wwPDB X-ray Structure Validation Summary Report (i)

May 25, 2020 – 06:24 pm BST

PDB ID	:	5FCR
Title	:	MOUSE COMPLEMENT FACTOR D
Authors	:	Mac Sweeney, A.
Deposited on	:	2015-12-15
Resolution	:	1.25 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.

We welcome your comments at *validation@mail.wwpdb.org*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

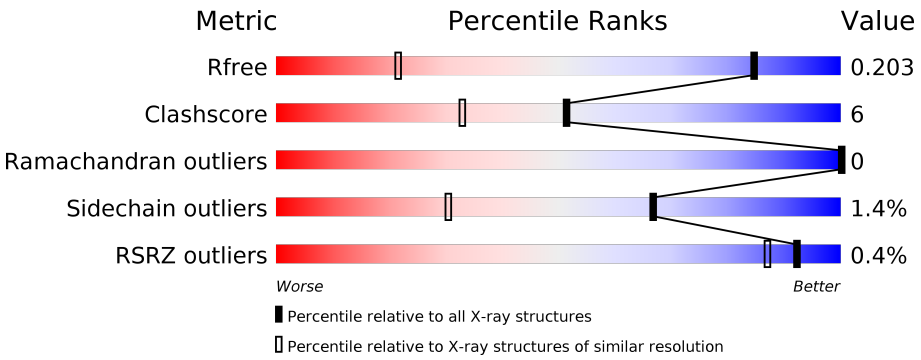
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 1.25 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}	130704	1023 (1.28-1.24)
Clashscore	141614	1060 (1.28-1.24)
Ramachandran outliers	138981	1029 (1.28-1.24)
Sidechain outliers	138945	1028 (1.28-1.24)
RSRZ outliers	127900	1004 (1.28-1.24)

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 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	234	79% 16% • •
1	B	234	78% 18% • •
1	C	234	79% 16% • •
1	D	234	% 80% 16% • •

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	GOL	B	302	-	X	X	-

2 Entry composition (i)

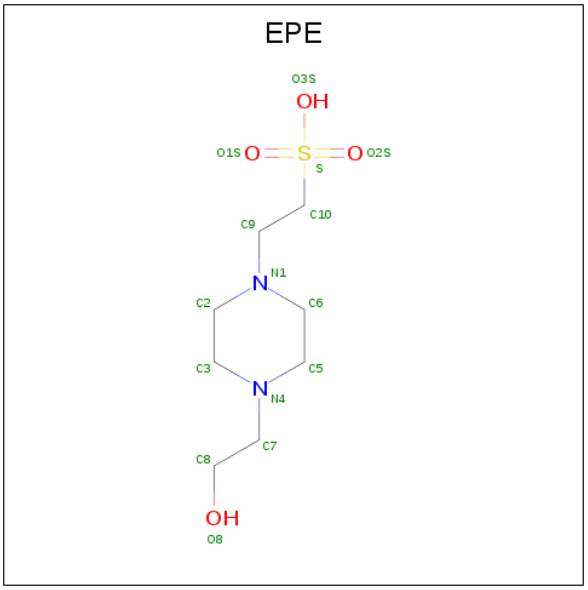
There are 6 unique types of molecules in this entry. The entry contains 8005 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Complement factor D.

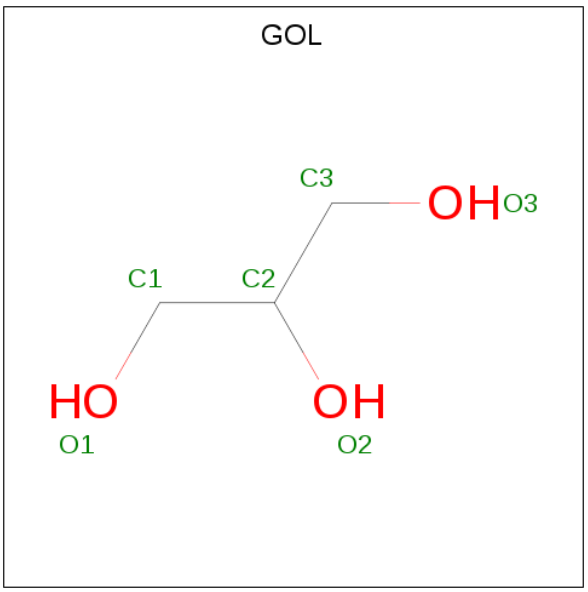
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	230	1799	1114	327	344	14	0	11	0
1	B	230	1762	1090	318	338	16	0	6	0
1	C	229	1784	1108	320	340	16	0	12	0
1	D	230	1771	1097	321	337	16	0	8	0

- Molecule 2 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



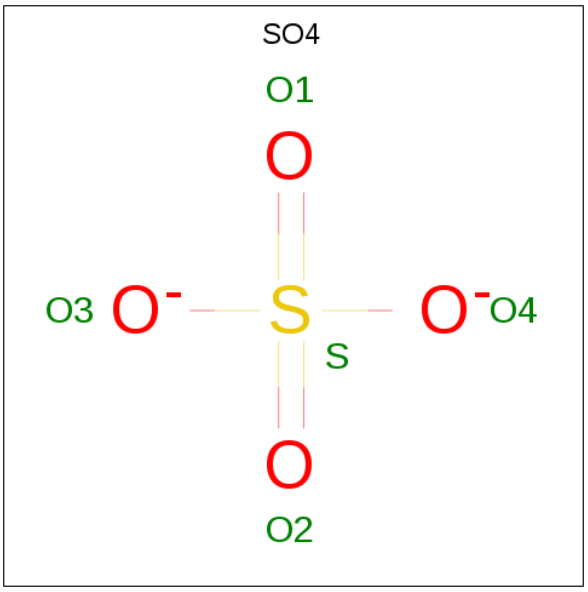
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	S		
2	B	1	15	8	2	4	1	0	0
2	C	1	15	8	2	4	1	0	0

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



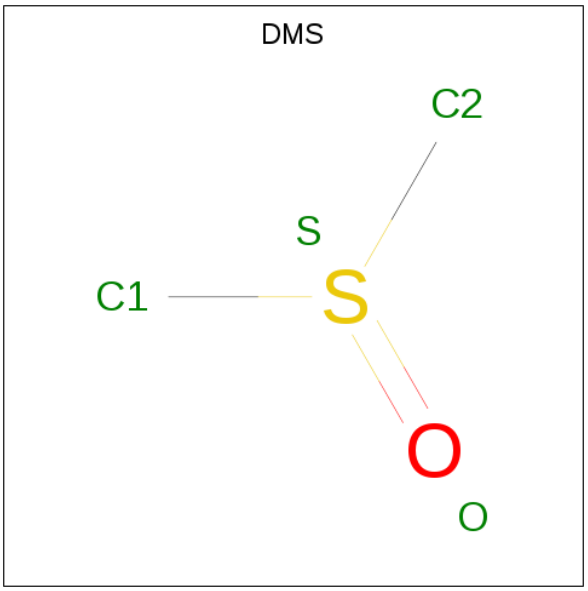
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	B	1	6	3	3	0	0

- Molecule 4 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 5 is DIMETHYL SULFOXIDE (three-letter code: DMS) (formula: C_2H_6OS).



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

- Molecule 6 is water.

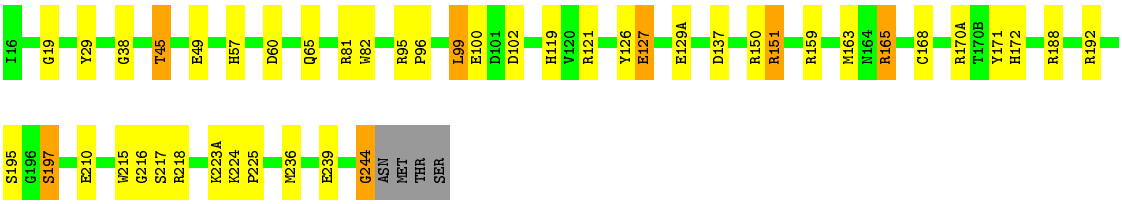
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	207	207	207	0	0
			Total	O		
6	B	234	234	234	0	0
			Total	O		
6	C	201	201	201	0	0
			Total	O		
6	D	198	198	198	0	0

3 Residue-property plots (i)

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

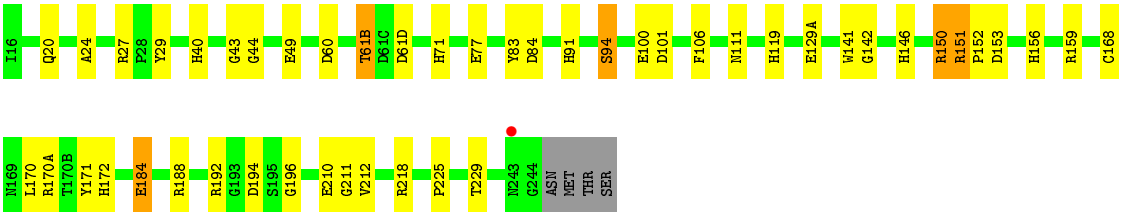
- Molecule 1: Complement factor D

Chain A:



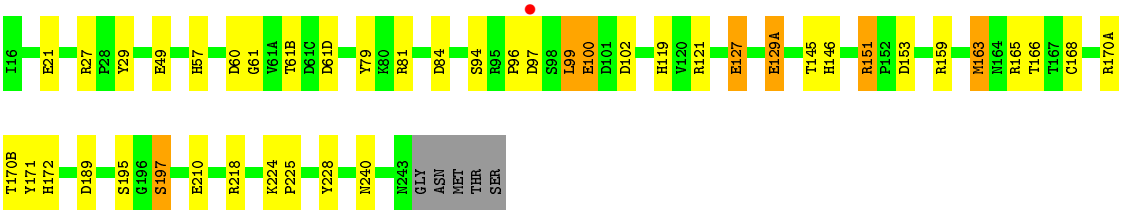
- Molecule 1: Complement factor D

Chain B:



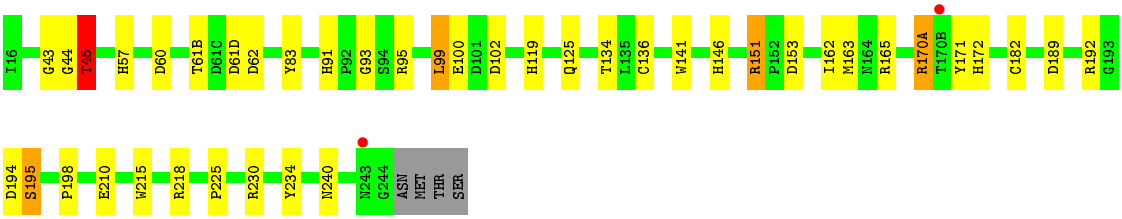
- Molecule 1: Complement factor D

Chain C:



- Molecule 1: Complement factor D

Chain D:


I16
G43
G44
T45
H57
D60
T61B
D61C
D61D
D62
Y83
H91
P92
G93
S94
R95
L99
E100
D101
D102
H119
Q125
T134
L135
C136
W141
H146
R151
P152
D153
I162
M163
N164
R165
R170A
T170B
Y171
H172
C182
D189
R192
G193
D194
S195
P198
E210
W215
R218
P225
R230
Y234
N240
N243
G244
ASN
MET
THR
SER

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	51.27Å 52.87Å 84.28Å 90.46° 98.04° 90.14°	Depositor
Resolution (Å)	83.45 – 1.25 83.45 – 1.25	Depositor EDS
% Data completeness (in resolution range)	88.3 (83.45-1.25) 88.3 (83.45-1.25)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.23 (at 1.25Å)	Xtriage
Refinement program	REFMAC 5.8.0131	Depositor
R, R_{free}	0.150 , 0.196 0.160 , 0.203	Depositor DCC
R_{free} test set	10776 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	8.7	Xtriage
Anisotropy	0.867	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 44.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.137 for -h,k,-l	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	8005	wwPDB-VP
Average B, all atoms (Å^2)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 42.16 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 2.1252e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, EPE, DMS, 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	1.56	16/1866 (0.9%)	1.55	38/2544 (1.5%)
1	B	1.66	24/1816 (1.3%)	1.49	25/2475 (1.0%)
1	C	1.58	15/1851 (0.8%)	1.46	21/2525 (0.8%)
1	D	1.59	16/1832 (0.9%)	1.43	20/2498 (0.8%)
All	All	1.60	71/7365 (1.0%)	1.48	104/10042 (1.0%)

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	A	0	1

The worst 5 of 71 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	168	CYS	CB-SG	-13.11	1.59	1.82
1	A	100	GLU	CD-OE2	-12.81	1.11	1.25
1	D	100	GLU	CD-OE2	-10.46	1.14	1.25
1	B	184	GLU	CD-OE2	-9.93	1.14	1.25
1	B	168[A]	CYS	CB-SG	-9.29	1.66	1.82

The worst 5 of 104 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	C	81	ARG	NE-CZ-NH2	-12.02	114.29	120.30
1	B	192	ARG	NE-CZ-NH2	-11.44	114.58	120.30
1	A	236	MET	CG-SD-CE	10.79	117.47	100.20
1	A	218	ARG	NE-CZ-NH2	-10.78	114.91	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	151	ARG	NE-CZ-NH2	-10.68	114.96	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	95	ARG	Mainchain

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	1799	0	1751	14	4
1	B	1762	0	1698	24	0
1	C	1784	0	1735	16	3
1	D	1771	0	1709	28	0
2	B	15	0	18	1	0
2	C	15	0	18	1	0
3	B	6	0	6	6	0
4	C	5	0	0	0	0
5	D	8	0	12	2	0
6	A	207	0	0	9	4
6	B	234	0	0	9	1
6	C	201	0	0	9	3
6	D	198	0	0	8	3
All	All	8005	0	6947	79	9

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 6.

The worst 5 of 79 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:D:136:CYS:SG	1:D:162:ILE:HD11	2.01	1.00
1:D:45[B]:THR:HG22	1:D:198:PRO:HG3	1.53	0.90

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:134:THR:O	1:D:162:ILE:HD12	1.76	0.84
1:D:91:HIS:HD1	1:D:93:GLY:H	1.25	0.80
1:B:49:GLU:HG2	6:B:585:HOH:O	1.84	0.76

The worst 5 of 9 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)
1:A:127:GLU:OE2	6:C:459:HOH:O[1_645]	0.33	1.87
1:C:127:GLU:OE2	6:A:404:HOH:O[1_455]	0.35	1.85
1:C:127:GLU:CD	6:A:404:HOH:O[1_455]	1.27	0.93
6:B:401:HOH:O	6:D:510:HOH:O[1_445]	1.31	0.89
1:A:127:GLU:CD	6:C:459:HOH:O[1_645]	1.38	0.82

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	239/234 (102%)	232 (97%)	7 (3%)	0	100	100
1	B	234/234 (100%)	223 (95%)	11 (5%)	0	100	100
1	C	238/234 (102%)	229 (96%)	9 (4%)	0	100	100
1	D	236/234 (101%)	225 (95%)	11 (5%)	0	100	100
All	All	947/936 (101%)	909 (96%)	38 (4%)	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	202/197 (102%)	198 (98%)	4 (2%)	55	17
1	B	196/197 (100%)	196 (100%)	0	100	100
1	C	201/197 (102%)	195 (97%)	6 (3%)	41	6
1	D	197/197 (100%)	191 (97%)	6 (3%)	41	6
All	All	796/788 (101%)	780 (98%)	16 (2%)	67	17

5 of 16 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	197[B]	SER
1	C	240[A]	ASN
1	D	99	LEU
1	C	197[A]	SER
1	D	170(A)	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	172	HIS
1	C	86	GLN
1	D	156	HIS
1	B	240	ASN
1	C	50	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)

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)

6 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	$\#\|Z\| > 2$
5	DMS	D	301	-	3,3,3	1.92	1 (33%)	3,3,3	1.97	1 (33%)
5	DMS	D	302	-	3,3,3	1.18	0	3,3,3	3.26	1 (33%)
2	EPE	C	301	-	15,15,15	1.81	2 (13%)	18,20,20	2.97	6 (33%)
3	GOL	B	302	-	5,5,5	2.21	3 (60%)	5,5,5	2.46	2 (40%)
4	SO4	C	302	-	4,4,4	0.71	0	6,6,6	0.91	0
2	EPE	B	301	-	15,15,15	1.53	3 (20%)	18,20,20	2.58	9 (50%)

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	EPE	C	301	-	-	2/9/19/19	0/1/1/1
2	EPE	B	301	-	-	3/9/19/19	0/1/1/1
3	GOL	B	302	-	-	2/4/4/4	-

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	301	EPE	O2S-S	4.58	1.58	1.45
2	B	301	EPE	C10-S	-3.95	1.71	1.77
2	C	301	EPE	C10-S	-3.81	1.72	1.77
5	D	301	DMS	C1-S	-3.31	1.51	1.75
3	B	302	GOL	C3-C2	-2.98	1.39	1.51

The worst 5 of 19 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	C	301	EPE	O3S-S-O1S	7.31	129.13	111.27
2	C	301	EPE	O3S-S-O2S	-7.11	93.90	111.27
2	B	301	EPE	O3S-S-C10	-6.17	95.79	105.77
5	D	302	DMS	C2-S-C1	5.56	127.03	98.44
2	B	301	EPE	O1S-S-C10	5.40	113.42	106.92

There are no chirality outliers.

5 of 7 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	301	EPE	S-C10-C9-N1
2	B	301	EPE	S-C10-C9-N1
3	B	302	GOL	O1-C1-C2-O2
2	B	301	EPE	N4-C7-C8-O8
2	B	301	EPE	C8-C7-N4-C3

There are no ring outliers.

5 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	D	301	DMS	1	0
5	D	302	DMS	1	0
2	C	301	EPE	1	0
3	B	302	GOL	6	0
2	B	301	EPE	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 (i)

6.1 Protein, DNA and RNA chains (i)

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	230/234 (98%)	-0.39	0	100	100	9, 15, 36, 71	0
1	B	230/234 (98%)	-0.36	1 (0%)	92	87	7, 13, 33, 64	0
1	C	229/234 (97%)	-0.33	1 (0%)	92	87	9, 15, 34, 66	1 (0%)
1	D	230/234 (98%)	-0.27	2 (0%)	84	76	8, 14, 37, 66	0
All	All	919/936 (98%)	-0.34	4 (0%)	92	87	7, 14, 36, 71	1 (0%)

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	243	ASN	4.1
1	D	243	ASN	2.8
1	C	97	ASP	2.6
1	D	170(B)	THR	2.5

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. 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	GOL	B	302	6/6	0.93	0.16	18,21,29,34	0
2	EPE	C	301	15/15	0.97	0.12	19,29,42,44	0
4	SO4	C	302	5/5	0.97	0.16	28,40,47,77	0
5	DMS	D	302	4/4	0.98	0.10	19,23,26,27	0
5	DMS	D	301	4/4	0.98	0.06	13,20,24,29	0
2	EPE	B	301	15/15	0.98	0.12	18,24,32,37	0

6.5 Other polymers (i)

There are no such residues in this entry.