Full wwPDB X-ray Structure Validation Report ⓘ

Sep 26, 2023 – 05:38 PM EDT

PDB ID	:	6D65
Title	:	Crystal structure of the human dual specificity phosphatase 1 catalytic domain (C258S) as a maltose binding protein fusion in complex with the designed AR protein off7
Authors	:	Gumpena, R.; Lountos, G.T.; Waugh, D.S.
Deposited on	:	2018-04-20
Resolution	:	2.35 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 ⓘ 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 ⓘ) were used in the production of this report:

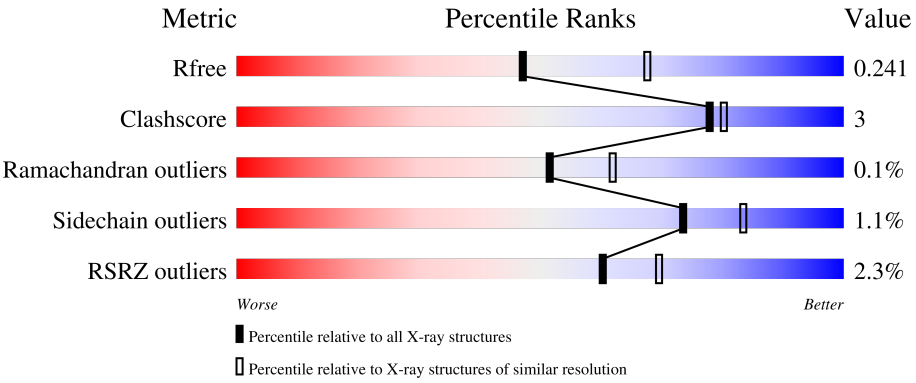
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 2.35 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	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	520	2% 92% 6% ·
1	C	520	2% 92% 6% ·
2	B	169	2% 84% 8% 8%
2	D	169	5% 86% 7% 8%

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	EOH	A	1412	-	-	X	-
5	EOH	C	1419	-	-	X	-

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11049 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 Maltose-binding periplasmic protein,Dual specificity protein phosphatase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	511	Total	C	N	O	S	0	2	0
			3948	2540	655	742	11			
1	C	511	Total	C	N	O	S	0	2	0
			3946	2538	654	743	11			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	initiating methionine	UNP P0AEX9
A	83	ALA	ASP	engineered mutation	UNP P0AEX9
A	84	ALA	LYS	engineered mutation	UNP P0AEX9
A	173	ALA	GLU	engineered mutation	UNP P0AEX9
A	174	ALA	ASN	engineered mutation	UNP P0AEX9
A	240	ALA	LYS	engineered mutation	UNP P0AEX9
A	360	ALA	GLU	engineered mutation	UNP P0AEX9
A	363	ALA	LYS	engineered mutation	UNP P0AEX9
A	364	ALA	ASP	engineered mutation	UNP P0AEX9
A	367	THR	-	linker	UNP P0AEX9
A	368	ASN	-	linker	UNP P0AEX9
A	369	ALA	-	linker	UNP P0AEX9
A	370	ALA	-	linker	UNP P0AEX9
A	371	ALA	-	linker	UNP P0AEX9
A	458	SER	CYS	engineered mutation	UNP P28562
A	515	HIS	-	expression tag	UNP P28562
A	516	HIS	-	expression tag	UNP P28562
A	517	HIS	-	expression tag	UNP P28562
A	518	HIS	-	expression tag	UNP P28562
A	519	HIS	-	expression tag	UNP P28562
A	520	HIS	-	expression tag	UNP P28562
C	1	MET	-	initiating methionine	UNP P0AEX9
C	83	ALA	ASP	engineered mutation	UNP P0AEX9
C	84	ALA	LYS	engineered mutation	UNP P0AEX9

Continued on next page...

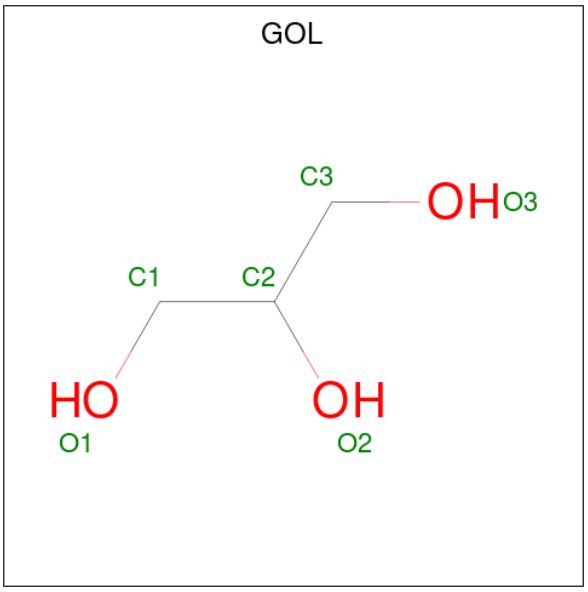
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	173	ALA	GLU	engineered mutation	UNP P0AEX9
C	174	ALA	ASN	engineered mutation	UNP P0AEX9
C	240	ALA	LYS	engineered mutation	UNP P0AEX9
C	360	ALA	GLU	engineered mutation	UNP P0AEX9
C	363	ALA	LYS	engineered mutation	UNP P0AEX9
C	364	ALA	ASP	engineered mutation	UNP P0AEX9
C	367	THR	-	linker	UNP P0AEX9
C	368	ASN	-	linker	UNP P0AEX9
C	369	ALA	-	linker	UNP P0AEX9
C	370	ALA	-	linker	UNP P0AEX9
C	371	ALA	-	linker	UNP P0AEX9
C	458	SER	CYS	engineered mutation	UNP P28562
C	515	HIS	-	expression tag	UNP P28562
C	516	HIS	-	expression tag	UNP P28562
C	517	HIS	-	expression tag	UNP P28562
C	518	HIS	-	expression tag	UNP P28562
C	519	HIS	-	expression tag	UNP P28562
C	520	HIS	-	expression tag	UNP P28562

- Molecule 2 is a protein called Designed AR protein off7.

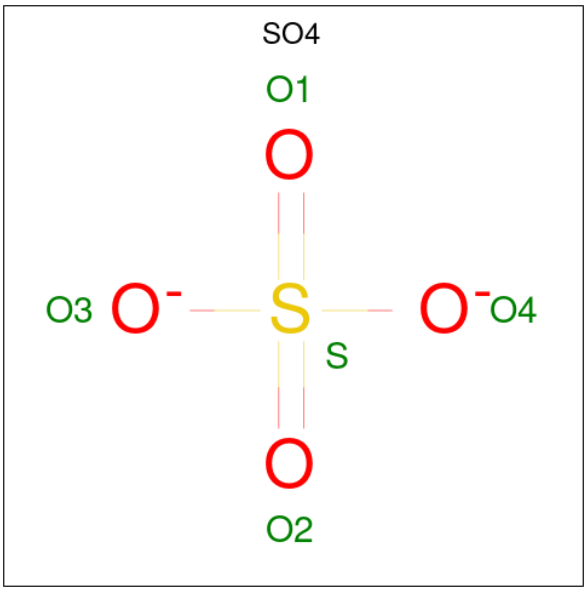
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	155	Total	C	N	O	S	0	3	0
			1175	742	200	230	3			
2	D	156	Total	C	N	O	S	0	0	0
			1160	733	199	225	3			

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



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

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



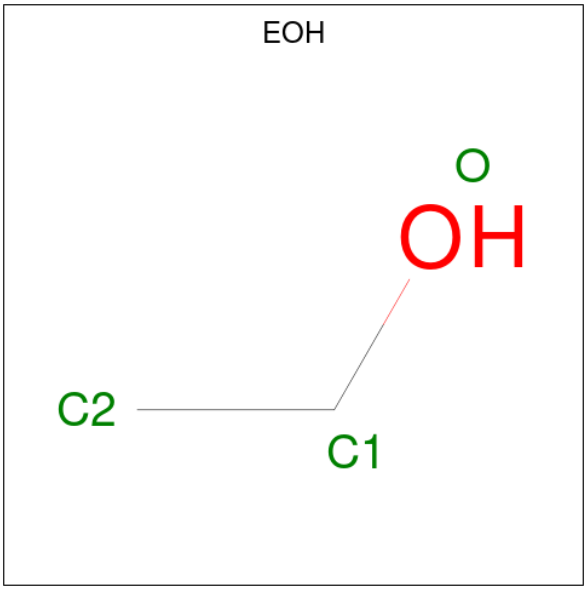
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	A	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0
4	B	1	Total O S 5 4 1	0	0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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is ETHANOL (three-letter code: EOH) (formula: C_2H_6O).



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

Continued on next page...

Continued from previous page...

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

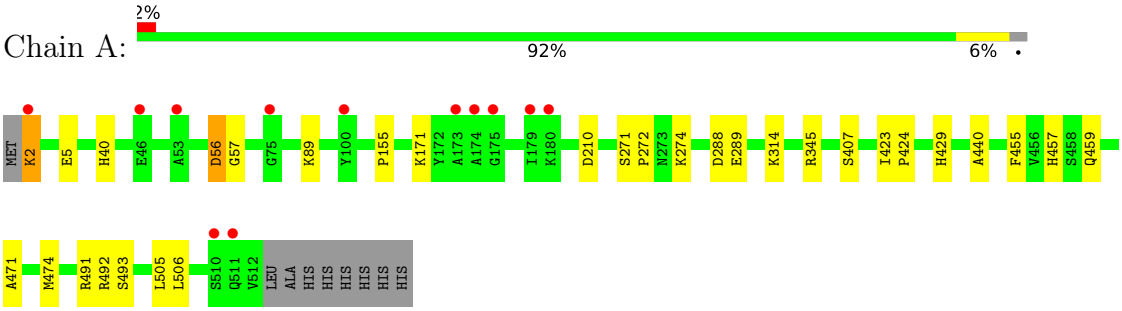
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	268	Total O 268 268	0	0
6	B	91	Total O 91 91	0	0
6	C	189	Total O 189 189	0	0
6	D	58	Total O 58 58	0	0

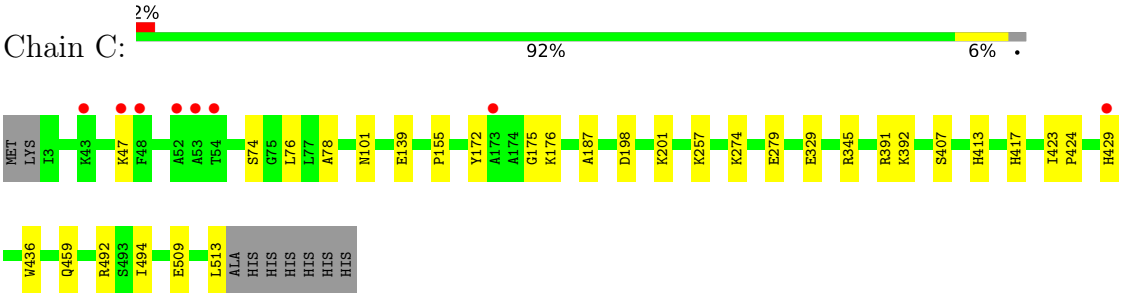
3 Residue-property plots (i)

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.

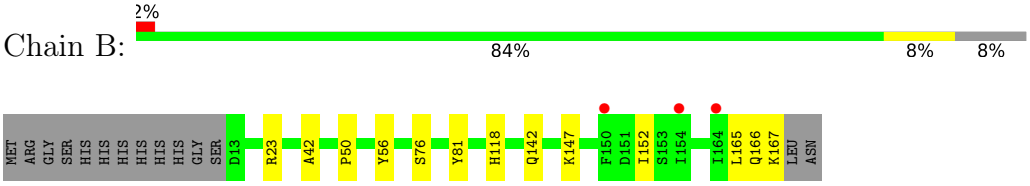
• Molecule 1: Maltose-binding periplasmic protein,Dual specificity protein phosphatase 1



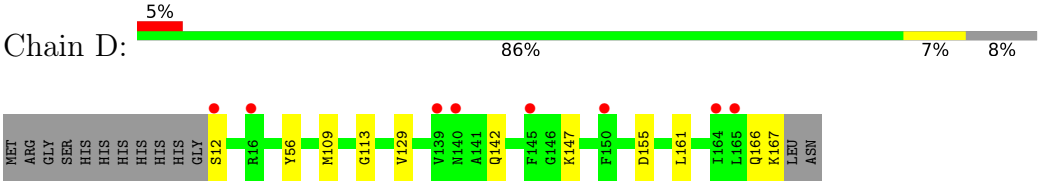
• Molecule 1: Maltose-binding periplasmic protein,Dual specificity protein phosphatase 1



• Molecule 2: Designed AR protein off7



• Molecule 2: Designed AR protein off7



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.33Å 109.50Å 218.56Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.67 – 2.35 38.68 – 2.35	Depositor EDS
% Data completeness (in resolution range)	99.3 (38.67-2.35) 99.3 (38.68-2.35)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.57 (at 2.34Å)	Xtriage
Refinement program	PHENIX (1.11.1_2575: ???)	Depositor
R, R_{free}	0.185 , 0.241 0.185 , 0.241	Depositor DCC
R_{free} test set	3728 reflections (4.92%)	wwPDB-VP
Wilson B-factor (Å^2)	34.6	Xtriage
Anisotropy	0.106	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 37.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	11049	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.55% 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: SO4, GOL, EOH

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	0.43	0/4051	0.57	0/5502
1	C	0.42	0/4046	0.54	0/5499
2	B	0.41	0/1201	0.58	0/1633
2	D	0.41	0/1180	0.58	0/1605
All	All	0.42	0/10478	0.56	0/14239

There are no bond length outliers.

There are no bond angle outliers.

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	3948	0	3888	25	0
1	C	3946	0	3874	22	0
2	B	1175	0	1143	6	0
2	D	1160	0	1128	6	0
3	A	24	0	32	4	0
3	C	12	0	16	2	0
4	A	40	0	0	1	0
4	B	15	0	0	0	0
4	C	45	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	D	15	0	0	0	0
5	A	21	0	42	7	0
5	B	6	0	12	1	0
5	C	30	0	60	3	0
5	D	6	0	12	0	0
6	A	268	0	0	1	0
6	B	91	0	0	0	0
6	C	189	0	0	2	0
6	D	58	0	0	0	0
All	All	11049	0	10207	59	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 3.

All (59) 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:429:HIS:HB2	3:A:1403:GOL:H32	1.68	0.76
1:C:47:LYS:NZ	3:C:1421:GOL:O1	2.22	0.72
1:C:509:GLU:HG3	1:C:513:LEU:HD12	1.72	0.70
2:D:147:LYS:NZ	2:D:155:ASP:OD2	2.29	0.65
1:C:198:ASP:HA	1:C:201:LYS:HG2	1.79	0.65
1:C:187:ALA:H	3:C:1413:GOL:H2	1.62	0.64
1:A:274:LYS:H	5:A:1411:EOH:H23	1.67	0.58
1:C:417[A]:HIS:ND1	4:C:1404:SO4:O4	2.30	0.57
1:A:288:ASP:HB2	5:A:1412:EOH:H11	1.87	0.56
1:C:155:PRO:HD3	1:C:345:ARG:HB2	1.86	0.56
1:A:289:GLU:HG3	5:A:1412:EOH:H22	1.88	0.54
1:C:492:ARG:NH2	1:C:494:ILE:HD11	2.23	0.54
1:A:474:MET:HE1	1:A:505:LEU:HD22	1.91	0.53
1:C:391:ARG:HH12	5:C:1416:EOH:H11	1.72	0.53
2:B:165:LEU:O	2:B:167:LYS:N	2.42	0.52
1:C:101:ASN:HA	1:C:176:LYS:NZ	2.26	0.51
1:A:2:LYS:HG3	1:A:56:ASP:OD1	2.10	0.51
1:C:429:HIS:CE1	5:C:1419:EOH:H12	2.46	0.50
1:C:423:ILE:HD12	1:C:436:TRP:HB3	1.94	0.50
2:B:23:ARG:HH11	5:B:204:EOH:H21	1.78	0.49
2:D:142:GLN:HA	2:D:147:LYS:O	2.12	0.49
1:A:429:HIS:CG	3:A:1403:GOL:H11	2.49	0.48
1:A:210:ASP:HB2	4:A:1419:SO4:O4	2.13	0.48
1:C:257:LYS:HE2	1:C:329:GLU:HG2	1.96	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:459:GLN:HG2	6:C:1615:HOH:O	2.13	0.48
2:D:109:MET:HE3	2:D:113:GLY:HA2	1.96	0.48
1:C:407:SER:O	1:C:424:PRO:HA	2.14	0.47
1:A:2:LYS:HE3	1:A:57:GLY:H	1.80	0.47
1:A:5:GLU:HG2	1:A:272:PRO:HB2	1.96	0.47
1:C:413:HIS:HE1	6:C:1505:HOH:O	1.97	0.46
1:C:429:HIS:HE1	5:C:1419:EOH:H12	1.79	0.46
1:A:271:SER:O	1:A:274:LYS:NZ	2.48	0.46
2:B:142:GLN:HA	2:B:147:LYS:O	2.16	0.46
1:C:345:ARG:HH11	1:C:345:ARG:HG2	1.81	0.45
1:A:314:LYS:NZ	5:A:1413:EOH:H11	2.32	0.45
1:A:155:PRO:HD3	1:A:345:ARG:HB2	1.99	0.45
1:A:289:GLU:HG3	5:A:1412:EOH:O	2.17	0.44
1:A:493:SER:H	3:A:1415:GOL:H12	1.83	0.44
1:A:89:LYS:HD3	6:A:1622:HOH:O	2.18	0.43
1:A:40:HIS:O	5:A:1410:EOH:H21	2.18	0.43
1:C:172:TYR:OH	1:C:175:GLY:HA2	2.19	0.43
2:D:166:GLN:O	2:D:167:LYS:HB2	2.17	0.43
1:A:407:SER:O	1:A:424:PRO:HA	2.18	0.42
1:C:78:ALA:HB2	1:C:274:LYS:HE3	2.01	0.42
2:B:42:ALA:O	2:B:50:PRO:HD3	2.19	0.42
2:D:129:VAL:HG11	2:D:161:LEU:HD11	2.02	0.41
2:B:118:HIS:HD2	2:B:152:ILE:HD12	1.85	0.41
1:A:289:GLU:HG3	5:A:1412:EOH:C2	2.50	0.41
1:A:455:PHE:CE2	1:A:457:HIS:HB3	2.55	0.41
1:A:474:MET:CE	1:A:505:LEU:HD22	2.50	0.41
1:C:392:LYS:HB2	1:C:392:LYS:HE3	1.98	0.41
1:C:74:SER:HB2	1:C:76:LEU:HG	2.02	0.41
1:A:429:HIS:H	3:A:1403:GOL:C1	2.34	0.41
1:A:506:LEU:HD23	1:A:506:LEU:HA	1.94	0.41
1:A:423:ILE:HD11	1:A:440:ALA:HB2	2.02	0.41
1:A:471:ALA:HA	1:A:474:MET:HE3	2.03	0.40
1:C:155:PRO:CD	1:C:345:ARG:HB2	2.51	0.40
2:B:76:SER:HA	2:B:81:TYR:O	2.21	0.40
2:D:129:VAL:HG21	2:D:161:LEU:HD11	2.03	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	511/520 (98%)	493 (96%)	18 (4%)	0	100 100
1	C	511/520 (98%)	495 (97%)	16 (3%)	0	100 100
2	B	156/169 (92%)	153 (98%)	2 (1%)	1 (1%)	25 26
2	D	154/169 (91%)	151 (98%)	3 (2%)	0	100 100
All	All	1332/1378 (97%)	1292 (97%)	39 (3%)	1 (0%)	51 62

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	166	GLN

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	404/414 (98%)	397 (98%)	7 (2%)	60 72
1	C	403/414 (97%)	401 (100%)	2 (0%)	88 93
2	B	119/132 (90%)	118 (99%)	1 (1%)	81 89
2	D	116/132 (88%)	114 (98%)	2 (2%)	60 72
All	All	1042/1092 (95%)	1030 (99%)	12 (1%)	73 82

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	LYS
1	A	56	ASP
1	A	171	LYS
1	A	459	GLN
1	A	491	ARG
1	A	492[A]	ARG
1	A	492[B]	ARG
2	B	56	TYR
1	C	139	GLU
1	C	279	GLU
2	D	12	SER
2	D	56	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	429	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

50 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	EOH	A	1410	-	2,2,2	0.47	0	1,1,1	0.17	0
3	GOL	C	1413	-	5,5,5	0.39	0	5,5,5	0.28	0
4	SO4	B	203	-	4,4,4	0.18	0	6,6,6	0.28	0
4	SO4	A	1408	-	4,4,4	0.17	0	6,6,6	0.58	0
5	EOH	C	1415	-	2,2,2	0.45	0	1,1,1	0.17	0
4	SO4	C	1404	-	4,4,4	0.17	0	6,6,6	0.23	0
5	EOH	A	1418	-	2,2,2	0.45	0	1,1,1	0.20	0
4	SO4	D	201	-	4,4,4	0.13	0	6,6,6	0.14	0
4	SO4	C	1407	-	4,4,4	0.17	0	6,6,6	0.39	0
4	SO4	D	203	-	4,4,4	0.12	0	6,6,6	0.27	0
5	EOH	C	1417	-	2,2,2	0.47	0	1,1,1	0.09	0
4	SO4	B	202	-	4,4,4	0.16	0	6,6,6	0.24	0
5	EOH	C	1416	-	2,2,2	0.48	0	1,1,1	0.09	0
5	EOH	C	1410	-	2,2,2	0.50	0	1,1,1	0.14	0
4	SO4	C	1402	-	4,4,4	0.18	0	6,6,6	0.28	0
5	EOH	D	204	-	2,2,2	0.47	0	1,1,1	0.14	0
4	SO4	A	1406	-	4,4,4	0.13	0	6,6,6	0.16	0
4	SO4	B	201	-	4,4,4	0.14	0	6,6,6	0.15	0
5	EOH	C	1412	-	2,2,2	0.46	0	1,1,1	0.33	0
4	SO4	C	1408	-	4,4,4	0.13	0	6,6,6	0.25	0
4	SO4	D	202	-	4,4,4	0.14	0	6,6,6	0.18	0
5	EOH	A	1413	-	2,2,2	0.44	0	1,1,1	0.02	0
3	GOL	A	1402	-	5,5,5	0.26	0	5,5,5	0.39	0
4	SO4	C	1409	-	4,4,4	0.14	0	6,6,6	0.29	0
5	EOH	A	1411	-	2,2,2	0.47	0	1,1,1	0.20	0
5	EOH	B	205	-	2,2,2	0.42	0	1,1,1	0.19	0
5	EOH	C	1418	-	2,2,2	0.48	0	1,1,1	0.09	0
5	EOH	C	1420	-	2,2,2	0.49	0	1,1,1	0.13	0
4	SO4	A	1404	-	4,4,4	0.14	0	6,6,6	0.20	0
5	EOH	D	205	-	2,2,2	0.45	0	1,1,1	0.18	0
4	SO4	A	1409	-	4,4,4	0.23	0	6,6,6	0.16	0
4	SO4	A	1414	-	4,4,4	0.14	0	6,6,6	0.16	0
5	EOH	A	1416	-	2,2,2	0.47	0	1,1,1	0.21	0
5	EOH	B	204	-	2,2,2	0.42	0	1,1,1	0.22	0
4	SO4	C	1406	-	4,4,4	0.15	0	6,6,6	0.34	0
3	GOL	C	1421	-	5,5,5	0.30	0	5,5,5	0.58	0
5	EOH	C	1419	-	2,2,2	0.48	0	1,1,1	0.15	0
4	SO4	A	1419	-	4,4,4	0.18	0	6,6,6	0.30	0
4	SO4	C	1405	-	4,4,4	0.19	0	6,6,6	0.19	0
4	SO4	C	1401	-	4,4,4	0.14	0	6,6,6	0.18	0
5	EOH	C	1411	-	2,2,2	0.48	0	1,1,1	0.14	0
3	GOL	A	1401	-	5,5,5	0.25	0	5,5,5	0.45	0
3	GOL	A	1415	-	5,5,5	0.34	0	5,5,5	0.67	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
4	SO4	A	1407	-	4,4,4	0.15	0	6,6,6	0.45	0
5	EOH	A	1417	-	2,2,2	0.46	0	1,1,1	0.30	0
5	EOH	A	1412	-	2,2,2	0.48	0	1,1,1	0.01	0
4	SO4	C	1403	-	4,4,4	0.18	0	6,6,6	0.15	0
4	SO4	A	1405	-	4,4,4	0.15	0	6,6,6	0.18	0
5	EOH	C	1414	-	2,2,2	0.50	0	1,1,1	0.08	0
3	GOL	A	1403	-	5,5,5	0.39	0	5,5,5	0.23	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
3	GOL	C	1413	-	-	3/4/4/4	-
3	GOL	A	1401	-	-	0/4/4/4	-
3	GOL	A	1415	-	-	1/4/4/4	-
3	GOL	A	1402	-	-	4/4/4/4	-
3	GOL	C	1421	-	-	4/4/4/4	-
3	GOL	A	1403	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1402	GOL	C1-C2-C3-O3
3	C	1421	GOL	O1-C1-C2-C3
3	A	1402	GOL	O2-C2-C3-O3
3	A	1402	GOL	O1-C1-C2-C3
3	A	1403	GOL	C1-C2-C3-O3
3	C	1413	GOL	O1-C1-C2-C3
3	C	1421	GOL	C1-C2-C3-O3
3	C	1421	GOL	O1-C1-C2-O2
3	C	1421	GOL	O2-C2-C3-O3
3	A	1403	GOL	O2-C2-C3-O3
3	A	1402	GOL	O1-C1-C2-O2
3	A	1415	GOL	O1-C1-C2-C3
3	C	1413	GOL	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
3	C	1413	GOL	O2-C2-C3-O3

There are no ring outliers.

13 monomers are involved in 19 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1410	EOH	1	0
3	C	1413	GOL	1	0
4	C	1404	SO4	1	0
5	C	1416	EOH	1	0
5	A	1413	EOH	1	0
5	A	1411	EOH	1	0
5	B	204	EOH	1	0
3	C	1421	GOL	1	0
5	C	1419	EOH	2	0
4	A	1419	SO4	1	0
3	A	1415	GOL	1	0
5	A	1412	EOH	4	0
3	A	1403	GOL	3	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, 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	511/520 (98%)	-0.13	12 (2%) 60 69	22, 33, 55, 74	0
1	C	511/520 (98%)	-0.07	8 (1%) 72 80	22, 34, 57, 67	0
2	B	155/169 (91%)	-0.08	3 (1%) 66 75	24, 34, 61, 72	0
2	D	156/169 (92%)	0.03	8 (5%) 28 39	24, 37, 65, 71	0
All	All	1333/1378 (96%)	-0.08	31 (2%) 60 69	22, 34, 58, 74	0

All (31) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	174	ALA	5.2
1	A	173	ALA	4.4
2	D	164	ILE	4.0
1	A	75	GLY	3.6
2	D	150	PHE	3.4
2	D	140	ASN	3.2
2	D	139	VAL	3.0
1	A	100	TYR	2.9
1	A	53	ALA	2.8
1	C	43	LYS	2.7
1	A	510	SER	2.6
1	A	175	GLY	2.5
2	D	165	LEU	2.5
2	B	154	ILE	2.4
1	A	46	GLU	2.4
1	C	429	HIS	2.3
1	C	47	LYS	2.3
1	C	52	ALA	2.3
2	D	12	SER	2.3
2	D	145	PHE	2.2
1	A	511	GLN	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	54	THR	2.2
2	B	164	ILE	2.2
1	A	2	LYS	2.2
1	C	48	PHE	2.2
2	B	150	PHE	2.2
1	A	180	LYS	2.1
1	C	53	ALA	2.1
2	D	16	ARG	2.0
1	A	179	ILE	2.0
1	C	173	ALA	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 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
5	EOH	C	1420	3/3	0.70	0.19	57,57,60,63	0
5	EOH	C	1418	3/3	0.79	0.17	57,57,60,68	0
4	SO4	C	1403	5/5	0.84	0.34	64,76,95,99	0
5	EOH	C	1411	3/3	0.84	0.14	41,41,44,48	0
5	EOH	C	1414	3/3	0.85	0.16	46,46,54,54	0
5	EOH	B	205	3/3	0.85	0.19	34,34,40,55	0
3	GOL	A	1401	6/6	0.85	0.33	51,57,59,63	0
3	GOL	A	1402	6/6	0.86	0.16	41,50,57,61	0
5	EOH	A	1418	3/3	0.88	0.10	54,54,55,61	0
3	GOL	C	1421	6/6	0.88	0.24	56,65,65,66	0
5	EOH	C	1419	3/3	0.89	0.12	43,43,45,46	0
5	EOH	C	1410	3/3	0.89	0.20	42,42,50,53	0
5	EOH	C	1415	3/3	0.90	0.11	48,48,52,53	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
5	EOH	A	1416	3/3	0.90	0.08	56,56,58,58	0
5	EOH	A	1412	3/3	0.90	0.34	45,45,45,53	0
5	EOH	A	1413	3/3	0.90	0.18	41,41,43,46	0
5	EOH	A	1417	3/3	0.91	0.13	55,55,57,57	0
5	EOH	A	1411	3/3	0.91	0.29	38,38,44,46	0
5	EOH	C	1416	3/3	0.91	0.17	55,55,58,59	0
4	SO4	C	1401	5/5	0.91	0.23	61,69,81,83	0
4	SO4	C	1402	5/5	0.91	0.23	61,67,73,85	0
4	SO4	B	203	5/5	0.91	0.21	53,61,70,74	0
3	GOL	C	1413	6/6	0.92	0.19	53,55,67,67	0
3	GOL	A	1415	6/6	0.92	0.14	45,50,52,59	0
5	EOH	C	1417	3/3	0.92	0.15	53,53,55,56	0
4	SO4	A	1406	5/5	0.94	0.24	52,64,71,76	0
4	SO4	D	203	5/5	0.94	0.23	51,55,69,72	0
3	GOL	A	1403	6/6	0.94	0.12	49,52,53,58	0
4	SO4	A	1404	5/5	0.95	0.23	72,76,76,80	0
5	EOH	A	1410	3/3	0.95	0.12	44,44,48,51	0
5	EOH	D	204	3/3	0.95	0.10	41,41,42,47	0
5	EOH	D	205	3/3	0.95	0.11	40,40,42,43	0
4	SO4	A	1405	5/5	0.96	0.14	51,57,65,66	0
4	SO4	A	1409	5/5	0.96	0.20	41,53,57,59	0
5	EOH	C	1412	3/3	0.96	0.14	34,34,36,43	0
4	SO4	C	1409	5/5	0.96	0.18	54,58,68,69	0
4	SO4	D	201	5/5	0.97	0.15	42,43,50,60	0
4	SO4	B	201	5/5	0.97	0.08	44,44,55,64	0
5	EOH	B	204	3/3	0.97	0.16	29,29,33,42	0
4	SO4	A	1408	5/5	0.97	0.10	40,44,50,55	0
4	SO4	C	1405	5/5	0.97	0.16	50,50,62,63	0
4	SO4	C	1406	5/5	0.97	0.09	45,52,57,63	0
4	SO4	C	1408	5/5	0.97	0.22	47,56,60,70	0
4	SO4	A	1419	5/5	0.97	0.17	34,39,44,52	5
4	SO4	A	1414	5/5	0.98	0.10	33,44,46,51	5
4	SO4	C	1407	5/5	0.99	0.11	32,34,38,39	0
4	SO4	D	202	5/5	0.99	0.10	37,40,45,48	0
4	SO4	B	202	5/5	0.99	0.08	39,42,47,50	0
4	SO4	C	1404	5/5	0.99	0.20	43,47,53,56	0
4	SO4	A	1407	5/5	1.00	0.12	30,32,40,41	0

6.5 Other polymers (i)

There are no such residues in this entry.