WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report (i)

May 25, 2020 – 09:24 am BST

PDB ID	:	3MHS
Title	:	Structure of the SAGA Ubp8/Sgf11/Sus1/Sgf73 DUB module bound to ubiquitin aldehyde
Authors	:	Samara, N.L.; Datta, A.B.; Berndsen, C.E.; Zhang, X.; Yao, T.; Cohen, R.E.; Wolberger, C.
Deposited on	:	2010-04-08
Resolution	:	1.89 Å(reported)

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

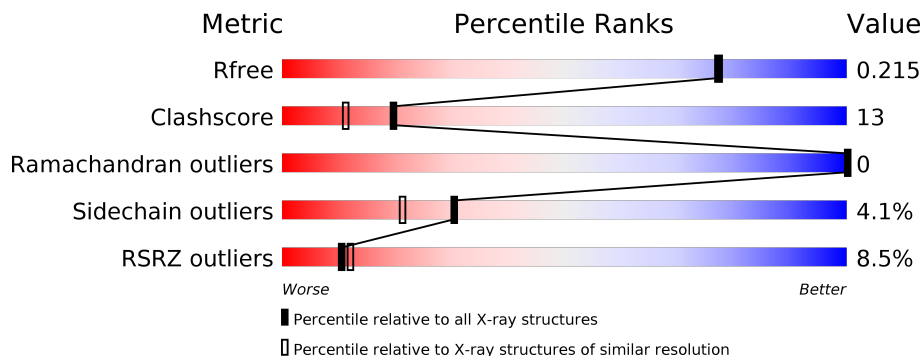|		
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (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 1.89 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		0.215
Clashscore		13
Ramachandran outliers		0
Sidechain outliers		4.1%
RSRZ outliers		8.5%

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	6207 (1.90-1.90)
Clashscore	141614	6847 (1.90-1.90)
Ramachandran outliers	138981	6760 (1.90-1.90)
Sidechain outliers	138945	6760 (1.90-1.90)
RSRZ outliers	127900	6082 (1.90-1.90)

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	476	6% 76% 18% • •
2	B	96	9% 77% 17% • 5%
3	C	99	11% 72% 16% 6% 6%
4	D	76	88% 12%
5	E	96	19% 68% 21% • 10%

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
7	GOL	A	1487	-	-	X	-

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 7460 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 Ubiquitin carboxyl-terminal hydrolase 8.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	455	Total 3758	C 2392	N 627	O 702	S 37	0	22	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-4	GLY	-	expression tag	UNP P50102
A	-3	ALA	-	expression tag	UNP P50102
A	-2	ALA	-	expression tag	UNP P50102
A	-1	ALA	-	expression tag	UNP P50102
A	0	ALA	-	expression tag	UNP P50102

- Molecule 2 is a protein called Protein SUS1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	91	Total 762	C 477	N 127	O 156	S 2	0	4	0

- Molecule 3 is a protein called SAGA-associated factor 11.

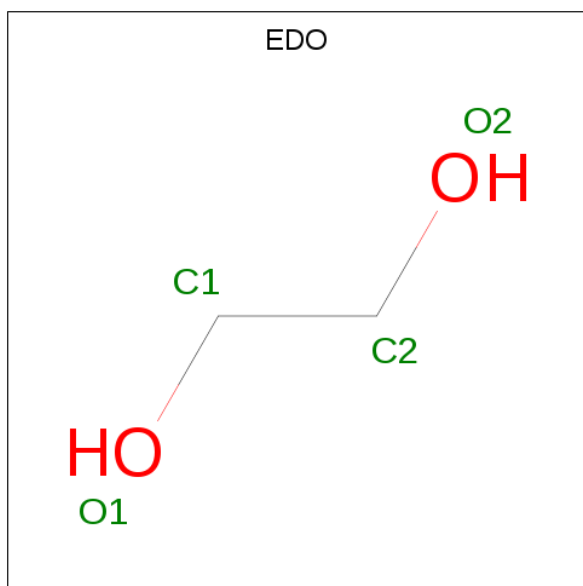| Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	93	Total 767	C 469	N 144	O 151	S 3	0	4	0

- Molecule 4 is a protein called Ubiquitin.
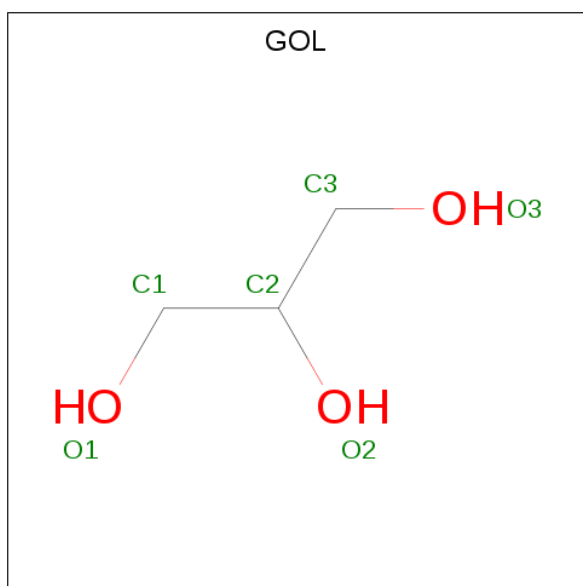
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	76	Total 611	C 384	N 105	O 121	S 1	0	2	0

- Molecule 5 is a protein called SAGA-associated factor 73.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	86	Total 697	C 442	N 116	O 135	S 4	0	2	0

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).

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

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).

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

- Molecule 8 is ZINC ION (three-letter code: ZN) (formula: Zn).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	6	Total 6	Zn 6	0	0
8	C	1	Total 1	Zn 1	0	0
8	E	1	Total 1	Zn 1	0	0

- Molecule 9 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	507	Total 509	O 509	0	2
9	B	57	Total 57	O 57	0	0
9	C	88	Total 88	O 88	0	0
9	D	91	Total 91	O 91	0	0
9	E	102	Total 102	O 102	0	0

3 Residue-property plots ⓘ

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.
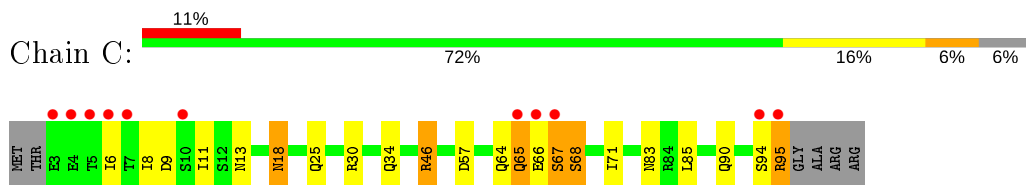
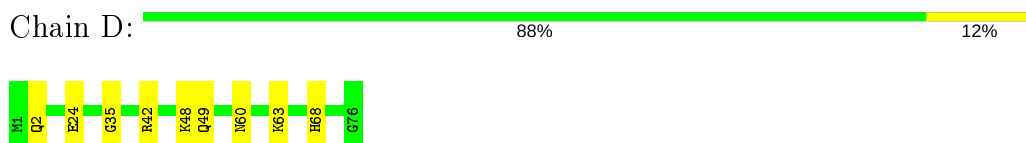• Molecule 1: Ubiquitin carboxyl-terminal hydrolase 8
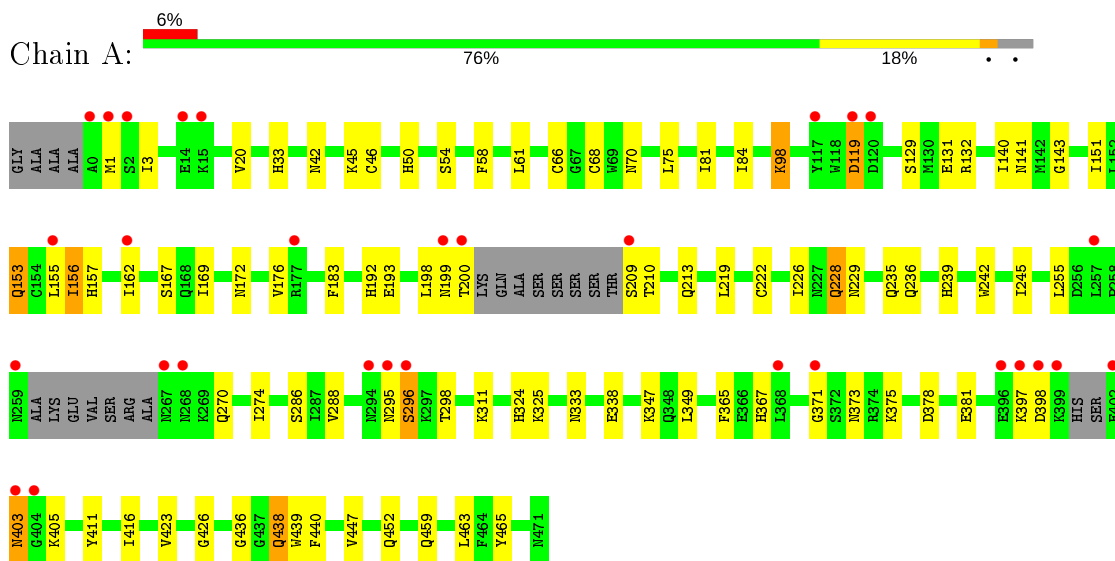
Chain A:

• Molecule 2: Protein SUS1
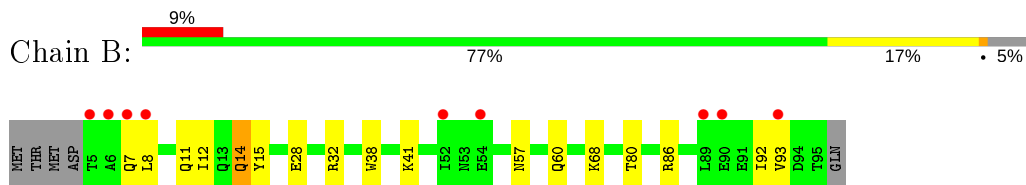
Chain B:
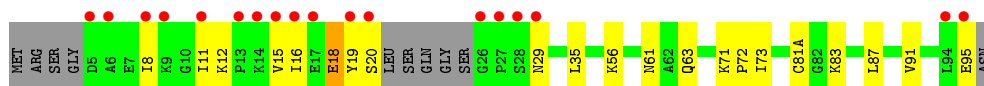
• Molecule 3: SAGA-associated factor 11

Chain C:

• Molecule 4: Ubiquitin

Chain D:

• Molecule 5: SAGA-associated factor 73

Chain E:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	78.60Å 102.08Å 120.90Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	35.10 – 1.89	Depositor
	35.10 – 1.89	EDS
% Data completeness	99.3 (35.10-1.89)	Depositor
(in resolution range)	99.3 (35.10-1.89)	EDS
R_{merge}	0.10	Depositor
R_{sym}	0.10	Depositor
$< I/\sigma(I) >$ [1]	2.02 (at 1.89Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, R_{free}	0.159 , 0.208	Depositor
	0.169 , 0.215	DCC
R_{free} test set	3934 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	24.3	Xtriage
Anisotropy	0.152	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 53.6	EDS
L-test for twinning[2]	$< |L| > = 0.50, < L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7460	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.35% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, ZN, GLZ, EDO

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	0.71	0/3891	0.70	1/5243 (0.0%)
2	B	0.54	0/780	0.57	0/1050
3	C	0.63	0/784	0.70	0/1059
4	D	0.83	1/619 (0.2%)	0.75	0/833
5	E	0.61	0/717	0.66	0/966
All	All	0.68	1/6791 (0.0%)	0.68	1/9151 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	24	GLU	CB-CG	-5.24	1.42	1.52

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	132	ARG	NE-CZ-NH1	7.56	124.08	120.30

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts (i)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	3758	0	3690	107	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	762	0	781	14	0
3	C	767	0	762	41	0
4	D	611	0	636	11	0
5	E	697	0	699	30	0
6	A	4	0	6	1	0
7	A	6	0	8	6	0
8	A	6	0	0	0	0
8	C	1	0	0	0	0
8	E	1	0	0	0	0
9	A	509	0	0	17	0
9	B	57	0	0	0	0
9	C	88	0	0	3	0
9	D	91	0	0	3	0
9	E	102	0	0	5	0
All	All	7460	0	6582	167	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 13.

All (167) 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:183[A]:PHE:CE2	1:A:226:ILE:HD13	1.67	1.28
1:A:416:ILE:HD13	1:A:463[A]:LEU:HD23	1.20	1.13
1:A:416:ILE:CD1	1:A:463[A]:LEU:CD2	2.31	1.08
1:A:270:GLN:NE2	5:E:61[B]:ASN:HD22	1.51	1.08
5:E:12:LYS:O	5:E:15:VAL:HG12	1.57	1.03
1:A:183[A]:PHE:CE2	1:A:226:ILE:CD1	2.41	1.03
1:A:167[B]:SER:OG	1:A:169:ILE:HG12	1.60	1.02
1:A:416:ILE:CD1	1:A:463[A]:LEU:HD23	1.87	1.01
1:A:381:GLU:HG3	9:A:838:HOH:O	1.61	1.00
1:A:183[A]:PHE:HE2	1:A:226:ILE:HD13	0.84	1.00
1:A:183[A]:PHE:HE2	1:A:226:ILE:CD1	1.75	1.00
1:A:416:ILE:HD13	1:A:463[A]:LEU:CD2	1.91	0.98
1:A:333:ASN:H	4:D:2:GLN:HE22	1.14	0.94
1:A:295:ASN:HB3	1:A:296:SER:HA	1.46	0.93
3:C:65:GLN:O	3:C:68:SER:HB3	1.69	0.93
1:A:270:GLN:HE22	5:E:61[B]:ASN:ND2	1.68	0.89
3:C:46[B]:ARG:HH11	3:C:46[B]:ARG:HB3	1.35	0.89
1:A:295:ASN:CB	1:A:296:SER:HA	2.06	0.86
1:A:235:GLN:HE22	1:A:426:GLY:H	1.25	0.84

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:270:GLN:HE22	5:E:61[B]:ASN:HD22	0.85	0.84
1:A:438[A]:GLN:HG2	1:A:440:PHE:CZ	2.17	0.80
1:A:162[A]:ILE:HG21	5:E:73:ILE:HD12	1.64	0.79
3:C:46[B]:ARG:HB3	3:C:46[B]:ARG:NH1	1.98	0.79
1:A:367:HIS:HD2	1:A:373:ASN:OD1	1.65	0.78
1:A:298[B]:THR:HG23	9:A:818:HOH:O	1.84	0.77
1:A:70:ASN:HD21	3:C:25:GLN:HE22	1.32	0.77
1:A:398:ASP:OD1	1:A:405:LYS:HD3	1.87	0.73
1:A:416:ILE:CD1	1:A:463[A]:LEU:HD22	2.19	0.73
1:A:375:LYS:HZ1	1:A:459:GLN:HE22	1.38	0.72
3:C:46[B]:ARG:HH11	3:C:46[B]:ARG:CB	2.03	0.72
1:A:270:GLN:NE2	5:E:61[B]:ASN:ND2	2.34	0.70
3:C:67:SER:HB3	3:C:83:ASN:ND2	2.05	0.70
1:A:45:LYS:NZ	2:B:14:GLN:HE22	1.89	0.70
3:C:6:ILE:CD1	5:E:8:ILE:HG23	2.22	0.70
1:A:228[A]:GLN:HE21	1:A:229:ASN:ND2	1.90	0.69
1:A:239:HIS:HD2	4:D:42:ARG:HH22	1.40	0.68
3:C:9:ASP:O	3:C:13:ASN:ND2	2.26	0.68
3:C:67:SER:HB3	3:C:83:ASN:HD21	1.57	0.67
1:A:192:HIS:ND1	9:A:526:HOH:O	2.26	0.67
4:D:42:ARG:NH1	4:D:49:GLN:OE1	2.28	0.66
1:A:209:SER:N	3:C:64:GLN:OE1	2.29	0.65
3:C:66:GLU:CD	3:C:67:SER:HA	2.16	0.65
1:A:129:SER:OG	1:A:131[B]:GLU:HG3	1.97	0.65
1:A:439:TRP:HE1	1:A:452:GLN:HE21	1.45	0.65
5:E:11:ILE:HG23	5:E:15:VAL:HG11	1.79	0.65
1:A:367:HIS:CD2	1:A:373:ASN:OD1	2.49	0.65
1:A:210:THR:HG21	3:C:71:ILE:CD1	2.28	0.64
3:C:6:ILE:O	5:E:8:ILE:HG22	1.97	0.64
1:A:416:ILE:HD11	1:A:463[A]:LEU:HD22	1.80	0.64
1:A:375:LYS:NZ	1:A:459:GLN:HE22	1.95	0.63
1:A:239:HIS:HD2	4:D:42:ARG:NH2	1.98	0.61
1:A:452:GLN:NE2	9:A:697:HOH:O	2.34	0.61
1:A:157:HIS:HE1	9:A:513:HOH:O	1.84	0.61
1:A:295:ASN:CB	1:A:296:SER:CA	2.79	0.61
1:A:222:CYS:O	1:A:226:ILE:HD12	2.00	0.61
1:A:324:HIS:HD2	1:A:349:LEU:O	1.84	0.61
2:B:86[A]:ARG:HG2	3:C:8:ILE:HG22	1.83	0.60
3:C:66:GLU:HA	3:C:66:GLU:OE2	2.00	0.60
2:B:15:TYR:CE2	2:B:92:ILE:HG12	2.37	0.59
1:A:3:ILE:HD11	9:A:853:HOH:O	2.02	0.59

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:1487:GOL:H11	4:D:35:GLY:HA3	1.83	0.59
1:A:75:LEU:HD12	1:A:75:LEU:O	2.02	0.59
5:E:71:LYS:HB2	5:E:72:PRO:HD2	1.85	0.58
2:B:57:ASN:HD22	2:B:60:GLN:H	1.51	0.58
1:A:140:ILE:H	1:A:213:GLN:HE22	1.52	0.58
9:A:956:HOH:O	3:C:66:GLU:HG2	2.02	0.58
5:E:56[B]:LYS:HG3	9:E:348:HOH:O	2.04	0.58
1:A:119[A]:ASP:OD2	1:A:119[A]:ASP:N	2.37	0.57
1:A:167[B]:SER:HG	1:A:169:ILE:HG12	1.66	0.57
1:A:324:HIS:HE1	9:A:810:HOH:O	1.88	0.57
1:A:172:ASN:ND2	9:A:587:HOH:O	2.39	0.56
1:A:228[A]:GLN:HE21	1:A:229:ASN:HD22	1.54	0.56
4:D:63:LYS:NZ	9:D:835:HOH:O	2.29	0.56
2:B:57:ASN:HB3	2:B:60:GLN:HG2	1.88	0.56
3:C:66:GLU:HA	3:C:68:SER:H	1.71	0.55
1:A:416:ILE:HD12	1:A:463[A]:LEU:CD2	2.34	0.54
3:C:6:ILE:CD1	5:E:8:ILE:CG2	2.86	0.54
1:A:193:GLU:HG3	1:A:198:LEU:HG	1.91	0.53
1:A:236:GLN:OE1	6:A:1295:EDO:H12	2.08	0.53
5:E:11:ILE:CG2	5:E:15:VAL:CG1	2.86	0.53
3:C:94[B]:SER:O	3:C:95:ARG:CB	2.57	0.53
1:A:274:ILE:HD12	5:E:63:GLN:NE2	2.24	0.53
3:C:6:ILE:HD12	5:E:8:ILE:CG2	2.38	0.53
5:E:11:ILE:CG2	5:E:15:VAL:HG11	2.39	0.53
1:A:239:HIS:CD2	4:D:42:ARG:HH22	2.26	0.53
5:E:83:LYS:NZ	9:E:803:HOH:O	2.42	0.53
1:A:228[B]:GLN:HG2	3:C:90[B]:GLN:NE2	2.24	0.52
1:A:438[A]:GLN:HG2	1:A:440:PHE:CE2	2.44	0.52
4:D:68:HIS:HE1	9:D:299:HOH:O	1.92	0.52
1:A:45:LYS:CE	2:B:14:GLN:HE22	2.23	0.52
1:A:141:ASN:ND2	1:A:143:GLY:H	2.08	0.52
3:C:94[B]:SER:OG	9:C:491:HOH:O	2.19	0.52
2:B:8:LEU:HD23	3:C:11:ILE:HD11	1.91	0.52
1:A:228[B]:GLN:HG2	3:C:90[B]:GLN:HE22	1.76	0.51
1:A:20:VAL:HG13	1:A:61:LEU:HD21	1.92	0.50
5:E:91:VAL:HG23	9:E:875:HOH:O	2.11	0.50
1:A:325:LYS:HE3	9:A:594:HOH:O	2.11	0.50
5:E:73:ILE:CD1	9:E:393:HOH:O	2.60	0.50
2:B:93:VAL:HG21	3:C:8:ILE:HD11	1.93	0.50
1:A:405:LYS:HE2	9:A:707:HOH:O	2.13	0.49
2:B:12:ILE:HD12	3:C:11:ILE:HD13	1.93	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:295:ASN:ND2	1:A:338:GLU:OE2	2.46	0.49
1:A:167[B]:SER:OG	1:A:169:ILE:CG1	2.46	0.49
1:A:140:ILE:H	1:A:213:GLN:NE2	2.09	0.49
5:E:15:VAL:HG13	5:E:16:ILE:N	2.27	0.49
1:A:54:SER:HB3	1:A:68:CYS:HB2	1.95	0.48
1:A:311:LYS:NZ	9:A:820:HOH:O	2.46	0.48
1:A:81:ILE:HD12	1:A:81:ILE:N	2.28	0.48
5:E:81(A):CYS:HB3	5:E:95:GLU:OE1	2.14	0.47
3:C:6:ILE:HG13	3:C:6:ILE:O	2.13	0.47
1:A:157:HIS:HD2	9:A:815:HOH:O	1.97	0.47
1:A:200:THR:HG23	1:A:200:THR:O	2.14	0.47
3:C:94[A]:SER:O	3:C:95:ARG:CB	2.62	0.47
1:A:367:HIS:HE1	7:A:1487:GOL:C3	2.27	0.47
3:C:6:ILE:HD12	5:E:8:ILE:HG21	1.96	0.47
1:A:371:GLY:HA2	7:A:1487:GOL:C3	2.45	0.47
1:A:288:VAL:HG22	1:A:296:SER:HB2	1.96	0.47
1:A:367:HIS:CE1	7:A:1487:GOL:H31	2.50	0.46
1:A:58:PHE:O	1:A:66:CYS:HA	2.15	0.46
1:A:199:ASN:HA	1:A:200:THR:HA	1.73	0.46
3:C:6:ILE:HG13	5:E:8:ILE:CG2	2.46	0.46
2:B:93:VAL:CG2	3:C:8:ILE:HD11	2.45	0.46
3:C:6:ILE:HG13	5:E:8:ILE:HG23	1.98	0.46
1:A:403:ASN:OD1	1:A:403:ASN:N	2.43	0.46
1:A:438[A]:GLN:CG	1:A:440:PHE:CE2	2.99	0.46
1:A:210:THR:HG21	3:C:71:ILE:HD11	1.98	0.45
1:A:367:HIS:HE1	7:A:1487:GOL:H31	1.81	0.45
9:A:478:HOH:O	4:D:68:HIS:HD2	1.99	0.45
3:C:18:ASN:ND2	9:C:138:HOH:O	2.49	0.45
7:A:1487:GOL:H12	4:D:35:GLY:O	2.17	0.45
1:A:438[A]:GLN:CG	1:A:440:PHE:CZ	2.95	0.45
1:A:192:HIS:CD2	1:A:198:LEU:HD23	2.52	0.45
1:A:311:LYS:HB2	9:A:936:HOH:O	2.17	0.44
1:A:156:ILE:O	1:A:162[A]:ILE:HD11	2.18	0.44
5:E:19:TYR:O	5:E:20:SER:CB	2.65	0.43
2:B:32:ARG:CZ	2:B:80:THR:HG21	2.47	0.43
1:A:436:GLY:HA2	9:C:102:HOH:O	2.19	0.43
2:B:38:TRP:HA	2:B:41:LYS:HE2	2.00	0.43
1:A:274:ILE:CD1	5:E:63:GLN:HE22	2.32	0.43
1:A:141:ASN:HD21	1:A:143:GLY:C	2.23	0.42
1:A:153:GLN:HA	1:A:153:GLN:HE21	1.84	0.42
1:A:398:ASP:CG	1:A:405:LYS:HD3	2.40	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:C:66:GLU:HA	3:C:68:SER:N	2.34	0.42
1:A:298[B]:THR:HG22	9:A:827:HOH:O	2.19	0.42
2:B:8:LEU:HD22	3:C:6:ILE:CD1	2.49	0.42
4:D:48:LYS:NZ	9:D:700:HOH:O	2.30	0.42
1:A:70:ASN:ND2	3:C:25:GLN:HE22	2.07	0.42
1:A:151:ILE:HG23	1:A:242:TRP:HE3	1.84	0.42
3:C:94[A]:SER:O	3:C:95:ARG:HB2	2.20	0.42
5:E:18:GLU:HG2	5:E:18:GLU:O	2.20	0.42
1:A:447:VAL:O	3:C:57:ASP:HB2	2.20	0.42
1:A:84:ILE:HG21	1:A:98:LYS:CD	2.50	0.41
1:A:155[A]:LEU:HA	1:A:155[A]:LEU:HD23	1.87	0.41
1:A:255[B]:LEU:HD12	9:A:902:HOH:O	2.19	0.41
1:A:33:HIS:HE1	9:E:839:HOH:O	2.02	0.41
3:C:6:ILE:CG1	5:E:8:ILE:HG23	2.50	0.41
1:A:42:ASN:HB2	5:E:35:LEU:HD11	2.02	0.41
2:B:14:GLN:HB2	2:B:14:GLN:HE21	1.64	0.41
3:C:30:ARG:O	3:C:34:GLN:HG3	2.20	0.41
1:A:46:CYS:O	1:A:50:HIS:HA	2.22	0.40
5:E:71:LYS:HB2	5:E:72:PRO:CD	2.49	0.40
1:A:228[A]:GLN:NE2	1:A:229:ASN:HD22	2.19	0.40
1:A:286:SER:O	1:A:347:LYS:HA	2.22	0.40
1:A:54:SER:HB3	1:A:68:CYS:CB	2.51	0.40
1:A:411:TYR:HB3	1:A:465:TYR:HB3	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	469/476 (98%)	460 (98%)	9 (2%)	0	100	100
2	B	93/96 (97%)	93 (100%)	0	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	95/99 (96%)	94 (99%)	1 (1%)	0	100	100
4	D	76/76 (100%)	75 (99%)	1 (1%)	0	100	100
5	E	84/96 (88%)	82 (98%)	2 (2%)	0	100	100
All	All	817/843 (97%)	804 (98%)	13 (2%)	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	435/429 (101%)	418 (96%)	17 (4%)	32	23
2	B	90/91 (99%)	84 (93%)	6 (7%)	16	7
3	C	88/89 (99%)	80 (91%)	8 (9%)	9	3
4	D	70/68 (103%)	69 (99%)	1 (1%)	67	65
5	E	80/86 (93%)	77 (96%)	3 (4%)	33	24
All	All	763/763 (100%)	728 (95%)	35 (5%)	30	17

All (35) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1	MET
1	A	98	LYS
1	A	119[A]	ASP
1	A	119[B]	ASP
1	A	153	GLN
1	A	156	ILE
1	A	176	VAL
1	A	219	LEU
1	A	228[A]	GLN
1	A	228[B]	GLN
1	A	296	SER
1	A	365	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	397	LYS
1	A	403	ASN
1	A	423	VAL
1	A	438[A]	GLN
1	A	438[B]	GLN
2	B	7	GLN
2	B	11	GLN
2	B	14	GLN
2	B	28[A]	GLU
2	B	28[B]	GLU
2	B	68	LYS
3	C	18	ASN
3	C	46[A]	ARG
3	C	46[B]	ARG
3	C	65	GLN
3	C	67	SER
3	C	68	SER
3	C	85	LEU
3	C	95	ARG
4	D	60	ASN
5	E	18	GLU
5	E	29	ASN
5	E	87	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (43) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	8	GLN
1	A	9	GLN
1	A	25	ASN
1	A	70	ASN
1	A	90	ASN
1	A	110	ASN
1	A	141	ASN
1	A	153	GLN
1	A	157	HIS
1	A	168	GLN
1	A	172	ASN
1	A	199	ASN
1	A	213	GLN
1	A	229	ASN
1	A	235	GLN

Continued from previous page...

Mol	Chain	Res	Type
1	A	239	HIS
1	A	251	GLN
1	A	270	GLN
1	A	324	HIS
1	A	348	GLN
1	A	361	GLN
1	A	367	HIS
1	A	424	ASN
1	A	452	GLN
1	A	459	GLN
1	A	469	GLN
2	B	7	GLN
2	B	11	GLN
2	B	14	GLN
2	B	27	ASN
2	B	57	ASN
2	B	60	GLN
3	C	13	ASN
3	C	18	ASN
3	C	34	GLN
3	C	59	ASN
3	C	64	GLN
4	D	2	GLN
4	D	25	ASN
4	D	60	ASN
4	D	62	GLN
4	D	68	HIS
5	E	59	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#|Z| > 2$
4	GLZ	D	76	4	3,3,3	0.69	0	0,2,2	0.00	-

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
4	GLZ	D	76	4	-	0/0/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 10 ligands modelled in this entry, 8 are monoatomic - leaving 2 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	$\#|Z| > 2$
7	GOL	A	1487	-	5,5,5	0.47	0	5,5,5	0.63	0
6	EDO	A	1295	-	3,3,3	0.73	0	2,2,2	0.13	0

Full wwPDB X-ray Structure Validation Report

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	GOL	A	1487	-	-	4/4/4/4	-
6	EDO	A	1295	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	A	1487	GOL	O1-C1-C2-C3
7	A	1487	GOL	C1-C2-C3-O3
7	A	1487	GOL	O1-C1-C2-O2
6	A	1295	EDO	O1-C1-C2-O2
7	A	1487	GOL	O2-C2-C3-O3

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	1487	GOL	6	0
6	A	1295	EDO	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	455/476 (95%)	0.29	30 (6%)	18	20	18, 32, 60, 77	3 (0%)
2	B	91/96 (94%)	0.32	9 (9%)	7	8	29, 49, 77, 97	1 (1%)
3	C	93/99 (93%)	0.47	11 (11%)	4	5	25, 39, 65, 88	1 (1%)
4	D	75/76 (98%)	-0.30	0	100	100	21, 28, 36, 45	0
5	E	86/96 (89%)	0.73	18 (20%)	1	1	21, 38, 96, 103	0
All	All	800/843 (94%)	0.30	68 (8%)	10	12	18, 35, 71, 103	5 (0%)

All (68) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	E	15	VAL	7.7
1	A	268	ASN	7.7
5	E	29	ASN	6.4
1	A	398	ASP	5.9
5	E	19	TYR	5.7
1	A	0	ALA	5.6
3	C	6	ILE	5.5
5	E	27	PRO	5.3
1	A	267	ASN	5.1
1	A	403	ASN	5.1
3	C	4	GLU	4.9
2	B	93	VAL	4.9
5	E	26	GLY	4.8
3	C	65	GLN	4.7
3	C	95	ARG	4.6
5	E	94	LEU	4.6
1	A	199	ASN	4.4
3	C	5	THR	4.1
5	E	16	ILE	4.1
5	E	11	ILE	4.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	259	ASN	4.0
2	B	6	ALA	3.9
1	A	397	LYS	3.8
1	A	402	GLU	3.7
5	E	5	ASP	3.7
2	B	52	ILE	3.6
5	E	8	ILE	3.6
3	C	67	SER	3.4
3	C	7	THR	3.4
1	A	200	THR	3.3
5	E	6	ALA	3.3
1	A	295	ASN	3.3
3	C	3	GLU	3.3
1	A	1	MET	3.2
1	A	296	SER	3.2
5	E	20	SER	3.2
2	B	5	THR	3.1
2	B	90	GLU	3.1
1	A	404	GLY	3.0
5	E	13	PRO	3.0
1	A	399	LYS	3.0
3	C	66	GLU	3.0
5	E	28	SER	2.9
3	C	10	SER	2.9
5	E	14	LYS	2.9
1	A	294	ASN	2.8
1	A	368	LEU	2.7
5	E	9	LYS	2.6
1	A	209	SER	2.6
1	A	120[A]	ASP	2.6
1	A	177	ARG	2.5
2	B	8	LEU	2.5
1	A	371	GLY	2.4
2	B	54	GLU	2.4
1	A	257	LEU	2.4
1	A	396	GLU	2.3
1	A	119[A]	ASP	2.3
2	B	7	GLN	2.3
5	E	17	GLU	2.2
5	E	95	GLU	2.2
1	A	155[A]	LEU	2.2
1	A	14	GLU	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	162[A]	ILE	2.1
1	A	117	TYR	2.1
1	A	15	LYS	2.1
2	B	89	LEU	2.1
1	A	2	SER	2.0
3	C	94[A]	SER	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, 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
4	GLZ	D	76	4/4	0.97	0.09	23,26,27,30	0

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
6	EDO	A	1295	4/4	0.77	0.23	43,46,55,55	0
7	GOL	A	1487	6/6	0.86	0.38	51,58,60,63	0
8	ZN	A	473	1/1	0.93	0.06	44,44,44,44	0
8	ZN	A	474	1/1	0.94	0.06	48,48,48,48	0
8	ZN	A	472	1/1	0.95	0.05	44,44,44,44	0
8	ZN	A	477	1/1	0.97	0.04	34,34,34,34	0
8	ZN	E	97	1/1	0.98	0.06	44,44,44,44	0
8	ZN	C	100	1/1	0.98	0.05	32,32,32,32	0
8	ZN	A	475	1/1	0.98	0.05	32,32,32,32	0
8	ZN	A	476	1/1	1.00	0.07	23,23,23,23	0

6.5 Other polymers (i)

There are no such residues in this entry.