Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 03:19 PM BST

PDB ID	:	4LSU
Title	:	Crystal structure of broadly and potently neutralizing antibody VRC-PG20 in complex with HIV-1 clade A/E 93TH057 gp120
Authors	:	Zhou, T.; Moquin, S.; Kwong, P.D.
Deposited on	:	2013-07-23
Resolution	:	2.30 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 following versions of software and data (see references ⓘ) were used in the production of this report:

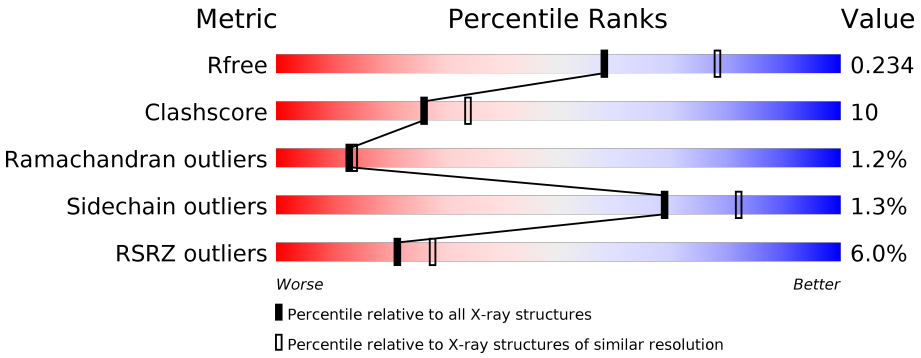
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.13.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.13.1

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 2.30 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	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (2.30-2.30)

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	G	353	7% 80% 18%
2	H	227	4% 79% 17%
3	L	204	6% 83% 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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
6	EDO	G	513	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 6365 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 HIV-1 CLADE A/E STRAIN 93TH057 GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	G	346	2705	1696	470	516	23	0	0	0

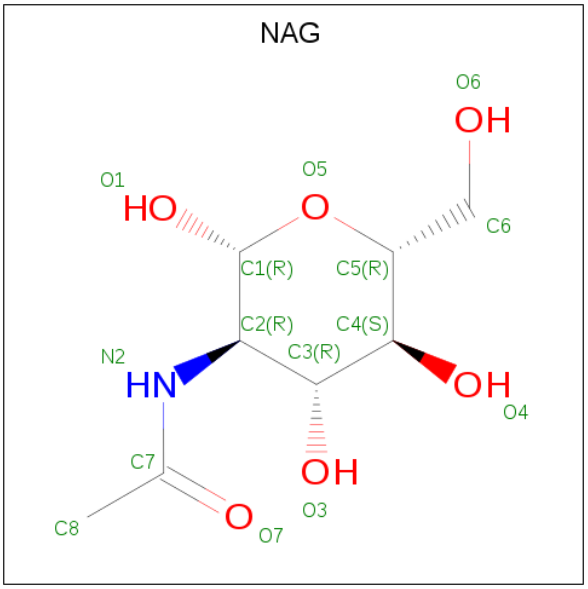
- Molecule 2 is a protein called HEAVY CHAIN OF ANTIBODY VRC-PG20.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	H	222	1742	1103	305	323	11	0	1	0

- Molecule 3 is a protein called LIGHT CHAIN OF ANTIBODY VRC-PG20.

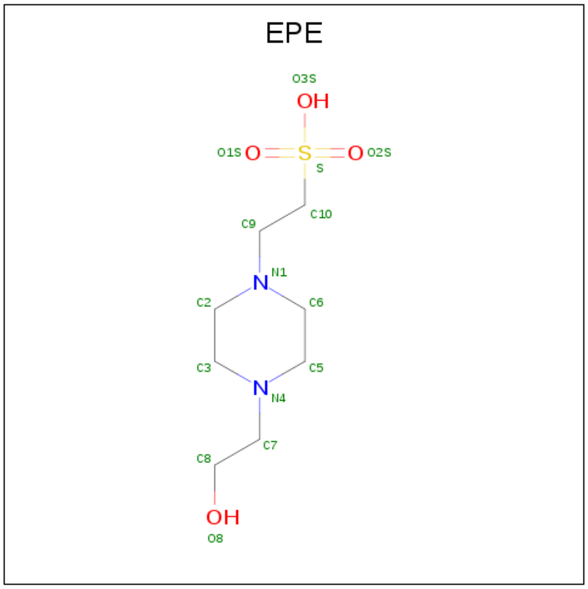
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	L	203	1508	943	249	311	5	0	0	0

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



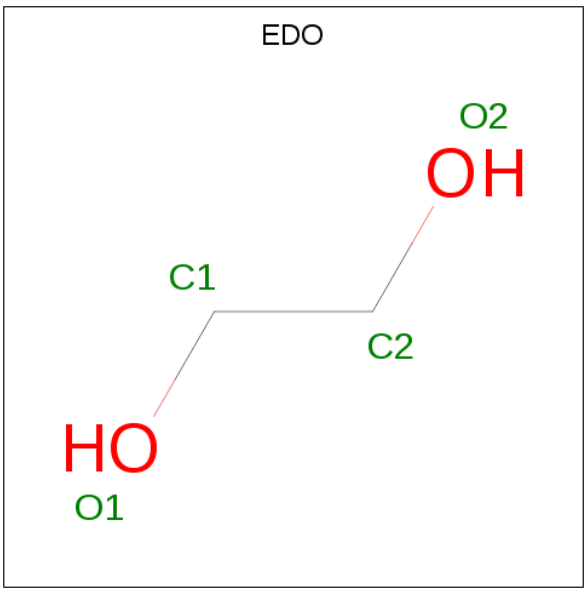
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0
4	G	1	14	8	1	5	0	0

- Molecule 5 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		
5	G	1	15	8	2	4	1	0	0

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



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	G	1	Total C O 4 2 2	0	0
6	G	1	Total C O 4 2 2	0	0
6	H	1	Total C O 4 2 2	0	0
6	H	1	Total C O 4 2 2	0	0
6	L	1	Total C O 4 2 2	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	H	1	Total Cl 1 1	0	0

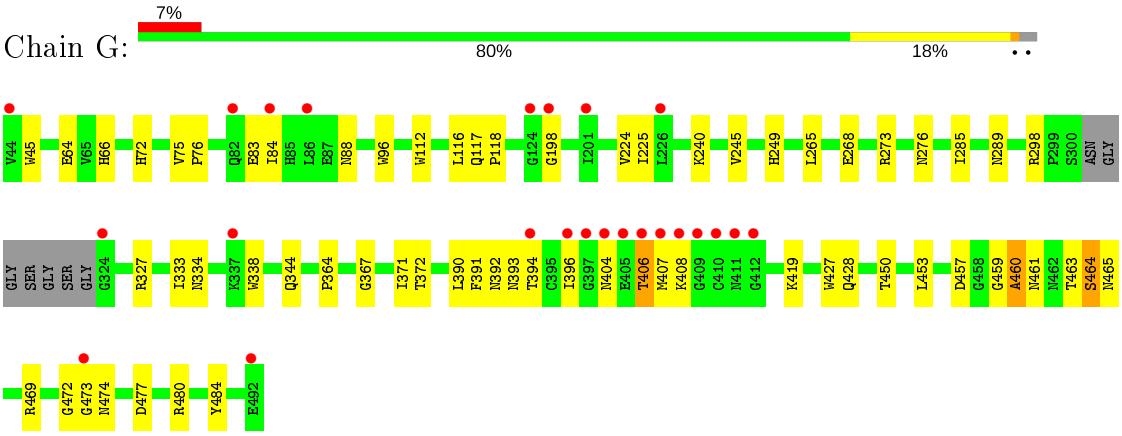
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	G	93	Total O 93 93	0	0
8	H	84	Total O 84 84	0	0
8	L	57	Total O 57 57	0	0

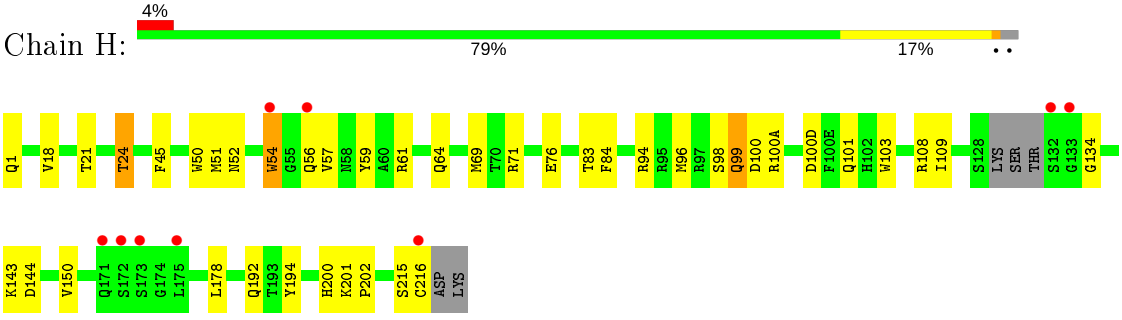
3 Residue-property plots (i)

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.

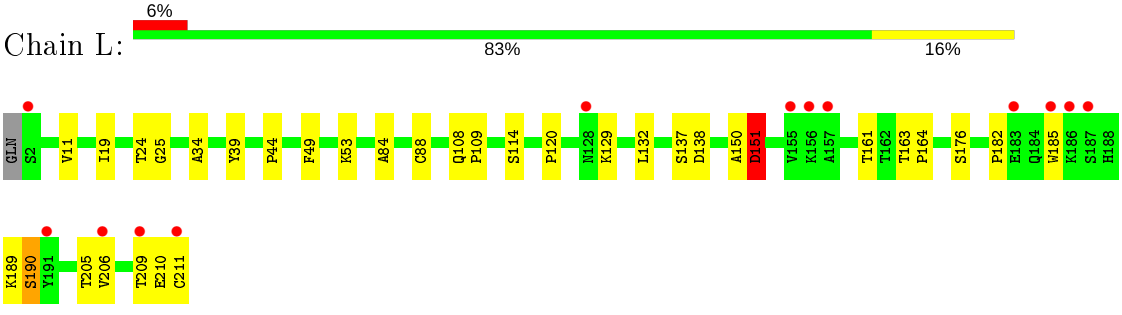
- Molecule 1: HIV-1 CLADE A/E STRAIN 93TH057 GP120



- Molecule 2: HEAVY CHAIN OF ANTIBODY VRC-PG20



- Molecule 3: LIGHT CHAIN OF ANTIBODY VRC-PG20



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	164.28Å 66.31Å 95.11Å 90.00° 108.34° 90.00°	Depositor
Resolution (Å)	45.14 – 2.30 45.14 – 2.29	Depositor EDS
% Data completeness (in resolution range)	89.7 (45.14-2.30) 89.7 (45.14-2.29)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.03 (at 2.29Å)	Xtriage
Refinement program	PHENIX dev_998	Depositor
R, R_{free}	0.179 , 0.234 0.178 , 0.234	Depositor DCC
R_{free} test set	1973 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å^2)	51.9	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 61.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	6365	wwPDB-VP
Average B, all atoms (Å^2)	77.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.75% 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: EPE, EDO, NAG, PCA, CL

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	G	0.41	0/2762	0.57	1/3749 (0.0%)
2	H	0.43	0/1786	0.60	0/2429
3	L	0.38	0/1545	0.52	0/2107
All	All	0.41	0/6093	0.57	1/8285 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	G	408	LYS	C-N-CA	8.18	139.48	122.30

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	G	2705	0	2636	60	0
2	H	1742	0	1687	43	0
3	L	1508	0	1450	26	0
4	G	140	0	130	3	0
5	G	15	0	17	1	0
6	G	8	0	12	6	0
6	H	8	0	12	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
6	L	4	0	6	0	0
7	H	1	0	0	1	0
8	G	93	0	0	10	0
8	H	84	0	0	7	0
8	L	57	0	0	1	0
All	All	6365	0	5950	125	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 10.

All (125) 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)
7:H:303:CL:CL	8:H:466:HOH:O	2.22	0.93
2:H:216:CYS:SG	3:L:211:CYS:N	2.45	0.89
1:G:88:ASN:ND2	8:G:631:HOH:O	2.08	0.87
2:H:100(A):ARG:NH1	8:H:464:HOH:O	2.18	0.77
1:G:240:LYS:NZ	8:G:608:HOH:O	2.17	0.76
1:G:473:GLY:HA3	1:G:477:ASP:OD2	1.90	0.72
3:L:108:GLN:HB2	3:L:109:PRO:HD2	1.70	0.71
2:H:52:ASN:HB3	2:H:56[B]:GLN:HG2	1.74	0.69
1:G:463:THR:HA	1:G:464:SER:CB	2.23	0.69
1:G:406:THR:HG23	1:G:406:THR:O	1.92	0.68
8:H:480:HOH:O	3:L:211:CYS:SG	2.53	0.67
1:G:463:THR:HA	1:G:464:SER:HB2	1.77	0.66
1:G:480:ARG:NH2	8:G:613:HOH:O	2.22	0.65
1:G:198:GLY:O	8:G:626:HOH:O	2.15	0.64
1:G:463:THR:HB	1:G:464:SER:HB2	1.80	0.63
1:G:406:THR:CG2	1:G:406:THR:O	2.47	0.62
1:G:463:THR:CA	1:G:464:SER:HB2	2.30	0.61
1:G:474:ASN:HA	6:G:513:EDO:C2	2.31	0.61
1:G:480:ARG:NE	8:G:613:HOH:O	2.20	0.61
1:G:394:THR:CG2	1:G:407:MET:SD	2.90	0.59
1:G:428:GLN:HG2	1:G:428:GLN:O	2.01	0.58
2:H:108:ARG:NH1	8:H:402:HOH:O	2.36	0.58
2:H:50:TRP:O	2:H:69:MET:HE3	2.04	0.57
1:G:474:ASN:CA	6:G:513:EDO:H22	2.34	0.57
1:G:463:THR:CB	1:G:464:SER:HB2	2.35	0.56
1:G:367:GLY:HA3	1:G:371:ILE:HD11	1.87	0.56
1:G:393:ASN:HA	1:G:396:ILE:HD13	1.88	0.56
1:G:463:THR:CA	1:G:464:SER:CB	2.84	0.55

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:103:TRP:O	6:H:302:EDO:H21	2.07	0.55
1:G:116:LEU:O	1:G:118:PRO:HD3	2.08	0.54
2:H:50:TRP:O	2:H:69:MET:CE	2.56	0.54
2:H:59:TYR:CE1	2:H:69:MET:HE2	2.42	0.54
1:G:344:GLN:HB2	4:G:505:NAG:H81	1.89	0.53
6:G:513:EDO:H11	8:G:617:HOH:O	2.07	0.53
3:L:120:PRO:HD3	3:L:132:LEU:HD23	1.91	0.53
2:H:50:TRP:C	2:H:69:MET:HE3	2.28	0.53
1:G:474:ASN:HA	6:G:513:EDO:H22	1.91	0.52
2:H:83:THR:HG22	2:H:84:PHE:N	2.25	0.52
2:H:216:CYS:SG	3:L:211:CYS:C	2.87	0.52
1:G:96:TRP:CE3	1:G:480:ARG:HD3	2.44	0.52
1:G:477:ASP:OD1	1:G:480:ARG:NH1	2.43	0.52
3:L:161:THR:OG1	3:L:176:SER:OG	2.27	0.51
1:G:64:GLU:OE1	1:G:66:HIS:HB2	2.09	0.51
2:H:100(A):ARG:HD3	2:H:100(D):ASP:OD2	2.11	0.51
2:H:18:VAL:HG11	2:H:109:ILE:HD13	1.93	0.50
3:L:163:THR:HG23	3:L:164:PRO:HD2	1.94	0.50
1:G:327:ARG:HB2	1:G:419:LYS:HE3	1.94	0.50
2:H:201:LYS:N	2:H:202:PRO:HD2	2.26	0.50
1:G:391:PHE:HB2	8:G:642:HOH:O	2.10	0.49
2:H:200:HIS:CE1	2:H:202:PRO:HG2	2.47	0.49
2:H:24:THR:HB	2:H:76:GLU:O	2.11	0.49
2:H:98:SER:N	8:H:445:HOH:O	2.44	0.49
3:L:11:VAL:CG1	3:L:19:ILE:HD11	2.42	0.49
1:G:459:GLY:O	1:G:461:ASN:N	2.45	0.49
2:H:200:HIS:NE2	2:H:202:PRO:HG2	2.27	0.49
6:G:513:EDO:C1	8:G:617:HOH:O	2.61	0.48
2:H:201:LYS:N	2:H:202:PRO:CD	2.77	0.48
2:H:99:GLN:O	2:H:100:ASP:CB	2.61	0.48
1:G:72:HIS:C	1:G:72:HIS:ND1	2.66	0.48
2:H:143:LYS:HE3	3:L:129:LYS:CE	2.44	0.48
2:H:83:THR:HG22	2:H:84:PHE:CD1	2.48	0.47
1:G:268:GLU:O	1:G:289:ASN:OD1	2.33	0.47
1:G:396:ILE:HD12	1:G:396:ILE:N	2.29	0.47
1:G:298:ARG:C	1:G:298:ARG:HD2	2.34	0.47
2:H:59:TYR:HE1	2:H:69:MET:HE2	1.79	0.47
1:G:396:ILE:C	1:G:404:ASN:H	2.19	0.47
2:H:45:PHE:CZ	3:L:44:PRO:HG3	2.50	0.47
1:G:112:TRP:CE3	1:G:116:LEU:HD22	2.50	0.46
1:G:371:ILE:HG13	1:G:372:THR:N	2.30	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:344:GLN:NE2	8:G:669:HOH:O	2.49	0.46
2:H:96:MET:HG2	2:H:101:GLN:HB2	1.97	0.46
2:H:98:SER:O	2:H:99:GLN:C	2.53	0.46
1:G:392:ASN:O	1:G:396:ILE:CD1	2.63	0.46
1:G:457:ASP:OD2	1:G:469:ARG:NH1	2.49	0.46
3:L:24:THR:HG22	3:L:25:GLY:N	2.31	0.46
1:G:75:VAL:HB	1:G:76:PRO:CD	2.47	0.45
3:L:150:ALA:O	3:L:151:ASP:HB2	2.17	0.45
3:L:137:SER:OG	3:L:138:ASP:OD2	2.33	0.45
2:H:21:THR:HG23	8:H:482:HOH:O	2.16	0.45
3:L:108:GLN:HB2	3:L:109:PRO:CD	2.42	0.45
3:L:34:ALA:O	3:L:88:CYS:HA	2.17	0.44
1:G:338:TRP:CE2	1:G:390:LEU:HD22	2.52	0.44
2:H:98:SER:O	2:H:100:ASP:N	2.50	0.44
1:G:224:VAL:HG22	1:G:225:ILE:N	2.32	0.44
2:H:100:ASP:OD1	2:H:100(A):ARG:N	2.50	0.44
2:H:51:MET:HE3	2:H:52:ASN:C	2.37	0.44
1:G:427:TRP:HB3	2:H:54:TRP:CH2	2.52	0.44
1:G:463:THR:HA	1:G:464:SER:OG	2.17	0.44
2:H:99:GLN:O	2:H:100:ASP:CG	2.55	0.44
3:L:39:TYR:CD1	3:L:84:ALA:HB2	2.53	0.44
2:H:143:LYS:HE3	3:L:129:LYS:HE3	2.00	0.43
2:H:216:CYS:SG	3:L:210:GLU:HB3	2.58	0.43
2:H:56[A]:GLN:CG	2:H:57:VAL:N	2.81	0.43
3:L:209:THR:O	3:L:210:GLU:HB2	2.17	0.43
1:G:460:ALA:CB	2:H:61:ARG:HB2	2.47	0.43
2:H:150:VAL:CG2	2:H:178:LEU:HD21	2.48	0.43
1:G:72:HIS:C	1:G:72:HIS:HD1	2.22	0.43
1:G:117:GLN:HA	1:G:117:GLN:OE1	2.19	0.43
1:G:265:LEU:HD21	1:G:450:THR:HG22	2.00	0.43
3:L:49:PHE:O	3:L:53:LYS:HB2	2.18	0.43
2:H:178:LEU:HD12	2:H:178:LEU:C	2.39	0.42
1:G:83:GLU:CB	1:G:245:VAL:HG12	2.50	0.42
1:G:276:ASN:OD1	4:G:504:NAG:O5	2.38	0.42
5:G:511:EPE:H52	5:G:511:EPE:O8	2.20	0.42
3:L:11:VAL:HG12	3:L:19:ILE:HD11	2.01	0.42
1:G:333:ILE:HD13	1:G:390:LEU:HD21	2.02	0.42
1:G:285:ILE:HG12	1:G:453:LEU:HD23	2.01	0.42
2:H:100:ASP:HB3	8:H:468:HOH:O	2.20	0.42
1:G:364:PRO:HG2	1:G:372:THR:HA	2.01	0.42
1:G:404:ASN:OD1	1:G:404:ASN:C	2.57	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:L:189:LYS:O	3:L:190:SER:HB3	2.20	0.41
2:H:192:GLN:HG2	2:H:194:TYR:CZ	2.56	0.41
2:H:45:PHE:HZ	3:L:44:PRO:HG3	1.85	0.41
1:G:463:THR:HB	1:G:465:ASN:H	1.84	0.41
4:G:505:NAG:H2	4:G:505:NAG:H83	1.92	0.41
1:G:474:ASN:CA	6:G:513:EDO:C2	2.95	0.41
1:G:472:GLY:CA	8:G:683:HOH:O	2.69	0.41
2:H:100:ASP:OD1	2:H:100:ASP:C	2.58	0.41
1:G:273:ARG:NH2	1:G:484:TYR:CE2	2.89	0.41
1:G:45:TRP:CD1	1:G:45:TRP:C	2.95	0.40
1:G:83:GLU:HB3	1:G:245:VAL:HG12	2.03	0.40
3:L:205:THR:HG22	3:L:206:VAL:N	2.36	0.40
2:H:215:SER:O	2:H:216:CYS:C	2.60	0.40
3:L:182:PRO:O	3:L:185:TRP:N	2.54	0.40
3:L:24:THR:CG2	8:L:455:HOH:O	2.68	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	G	342/353 (97%)	320 (94%)	18 (5%)	4 (1%)	13	14
2	H	219/227 (96%)	205 (94%)	11 (5%)	3 (1%)	11	11
3	L	201/204 (98%)	185 (92%)	14 (7%)	2 (1%)	15	17
All	All	762/784 (97%)	710 (93%)	43 (6%)	9 (1%)	13	14

All (9) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	460	ALA
1	G	464	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	H	99	GLN
1	G	406	THR
2	H	144	ASP
3	L	190	SER
1	G	334	ASN
3	L	151	ASP
2	H	134	GLY

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	G	308/311 (99%)	306 (99%)	2 (1%)	86 94
2	H	192/196 (98%)	187 (97%)	5 (3%)	46 63
3	L	170/171 (99%)	168 (99%)	2 (1%)	71 84
All	All	670/678 (99%)	661 (99%)	9 (1%)	69 82

All (9) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	84	ILE
1	G	249	HIS
2	H	24	THR
2	H	54	TRP
2	H	64	GLN
2	H	71	ARG
2	H	94	ARG
3	L	114	SER
3	L	151	ASP

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

Mol	Chain	Res	Type
2	H	102	HIS

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	$\#\|Z\| > 2$
2	PCA	H	1	2	7,8,9	2.19	2 (28%)	9,10,12	2.33	5 (55%)

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	PCA	H	1	2	-	0/0/11/13	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	1	PCA	CD-N	4.86	1.47	1.34
2	H	1	PCA	CA-N	3.00	1.50	1.46

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	H	1	PCA	CB-CA-C	-3.69	107.62	112.70
2	H	1	PCA	CA-N-CD	-3.13	102.86	113.58
2	H	1	PCA	OE-CD-CG	-3.10	121.35	126.76
2	H	1	PCA	CB-CA-N	2.85	111.49	103.30
2	H	1	PCA	CG-CD-N	2.51	114.90	108.39

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 17 ligands modelled in this entry, 1 is monoatomic - leaving 16 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	$\#\|Z\| > 2$
4	NAG	G	509	1	14,14,15	0.51	0	17,19,21	1.65	2 (11%)
6	EDO	H	302	-	3,3,3	0.52	0	2,2,2	0.15	0
4	NAG	G	505	1	14,14,15	0.53	0	17,19,21	1.26	2 (11%)
5	EPE	G	511	-	15,15,15	0.65	1 (6%)	18,20,20	2.11	6 (33%)
6	EDO	G	513	-	3,3,3	0.50	0	2,2,2	0.10	0
4	NAG	G	501	1	14,14,15	0.54	0	17,19,21	0.88	1 (5%)
4	NAG	G	506	1	14,14,15	0.42	0	17,19,21	1.73	4 (23%)
4	NAG	G	508	1	14,14,15	0.65	0	17,19,21	0.79	0
4	NAG	G	504	1	14,14,15	0.62	0	17,19,21	1.22	2 (11%)
4	NAG	G	510	1	14,14,15	0.54	0	17,19,21	1.06	1 (5%)
4	NAG	G	507	1	14,14,15	0.50	0	17,19,21	0.92	0
6	EDO	L	301	-	3,3,3	0.44	0	2,2,2	0.26	0
4	NAG	G	503	1	14,14,15	0.48	0	17,19,21	1.46	2 (11%)
6	EDO	H	301	-	3,3,3	0.55	0	2,2,2	0.18	0
6	EDO	G	512	-	3,3,3	0.36	0	2,2,2	0.57	0
4	NAG	G	502	1	14,14,15	0.54	0	17,19,21	1.78	2 (11%)

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	NAG	G	509	1	-	3/6/23/26	0/1/1/1
6	EDO	H	302	-	-	0/1/1/1	-
4	NAG	G	505	1	-	3/6/23/26	0/1/1/1
5	EPE	G	511	-	-	1/9/19/19	0/1/1/1
6	EDO	G	513	-	-	0/1/1/1	-
4	NAG	G	501	1	-	0/6/23/26	0/1/1/1
4	NAG	G	506	1	-	0/6/23/26	0/1/1/1
4	NAG	G	508	1	-	1/6/23/26	0/1/1/1
4	NAG	G	504	1	-	6/6/23/26	0/1/1/1
4	NAG	G	510	1	-	0/6/23/26	0/1/1/1
4	NAG	G	507	1	-	2/6/23/26	0/1/1/1
6	EDO	L	301	-	-	0/1/1/1	-
4	NAG	G	503	1	-	0/6/23/26	0/1/1/1
6	EDO	H	301	-	-	1/1/1/1	-
6	EDO	G	512	-	-	1/1/1/1	-
4	NAG	G	502	1	-	1/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	511	EPE	C10-S	2.11	1.80	1.77

All (22) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	G	506	NAG	C1-O5-C5	5.76	120.00	112.19
5	G	511	EPE	O2S-S-C10	5.73	113.82	106.92
4	G	509	NAG	C1-O5-C5	5.01	118.98	112.19
4	G	502	NAG	C1-O5-C5	4.98	118.94	112.19
4	G	502	NAG	O5-C1-C2	4.70	118.70	111.29
4	G	503	NAG	C1-O5-C5	4.45	118.23	112.19
5	G	511	EPE	C5-N4-C3	3.99	117.82	108.83
4	G	510	NAG	C1-O5-C5	3.13	116.44	112.19
4	G	504	NAG	C1-O5-C5	2.99	116.24	112.19
4	G	509	NAG	C4-C3-C2	-2.88	106.79	111.02
5	G	511	EPE	C7-N4-C5	2.85	118.52	111.23
5	G	511	EPE	C7-N4-C3	2.61	117.91	111.23
4	G	501	NAG	C1-O5-C5	2.55	115.64	112.19
5	G	511	EPE	O2S-S-O1S	-2.52	105.21	113.95

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	G	505	NAG	C3-C4-C5	2.48	114.66	110.24
4	G	504	NAG	O5-C5-C6	2.40	110.97	107.20
4	G	506	NAG	C4-C3-C2	-2.19	107.80	111.02
5	G	511	EPE	C5-C6-N1	-2.11	106.30	110.64
4	G	505	NAG	C1-O5-C5	2.08	115.01	112.19
4	G	506	NAG	O5-C5-C4	2.06	115.83	110.83
4	G	503	NAG	C4-C3-C2	-2.02	108.05	111.02
4	G	506	NAG	C3-C4-C5	2.02	113.85	110.24

There are no chirality outliers.

All (19) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	G	509	NAG	C3-C2-N2-C7
5	G	511	EPE	C8-C7-N4-C5
4	G	504	NAG	C8-C7-N2-C2
4	G	504	NAG	O7-C7-N2-C2
4	G	509	NAG	C8-C7-N2-C2
4	G	504	NAG	O5-C5-C6-O6
4	G	509	NAG	O7-C7-N2-C2
4	G	505	NAG	C8-C7-N2-C2
4	G	505	NAG	O7-C7-N2-C2
4	G	504	NAG	C4-C5-C6-O6
6	H	301	EDO	O1-C1-C2-O2
4	G	502	NAG	O5-C5-C6-O6
4	G	507	NAG	C8-C7-N2-C2
4	G	507	NAG	O7-C7-N2-C2
4	G	504	NAG	C1-C2-N2-C7
6	G	512	EDO	O1-C1-C2-O2
4	G	504	NAG	C3-C2-N2-C7
4	G	508	NAG	C8-C7-N2-C2
4	G	505	NAG	C4-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	H	302	EDO	1	0
4	G	505	NAG	2	0
5	G	511	EPE	1	0
6	G	513	EDO	6	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	504	NAG	1	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	G	346/353 (98%)	0.39	24 (6%)	16	22	39, 78, 133, 191	0
2	H	221/227 (97%)	0.06	9 (4%)	37	44	39, 65, 119, 165	0
3	L	203/204 (99%)	0.08	13 (6%)	19	25	39, 71, 132, 183	0
All	All	770/784 (98%)	0.21	46 (5%)	21	28	39, 73, 129, 191	0

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	404	ASN	5.5
1	G	396	ILE	4.7
1	G	324	GLY	4.7
1	G	407	MET	4.5
1	G	408	LYS	4.4
1	G	84	ILE	4.4
1	G	44	VAL	4.3
1	G	86	LEU	4.0
3	L	2	SER	3.7
3	L	209	THR	3.5
3	L	155	VAL	3.5
1	G	492	GLU	3.4
2	H	54	TRP	3.4
1	G	410	CYS	3.4
1	G	198	GLY	3.3
1	G	405	GLU	3.3
1	G	406	THR	3.3
1	G	409	GLY	3.2
1	G	397	GLY	3.0
1	G	82	GLN	3.0
3	L	211	CYS	2.9
3	L	128	ASN	2.9
2	H	56[A]	GLN	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	H	216	CYS	2.8
1	G	394	THR	2.7
3	L	191	TYR	2.7
2	H	133	GLY	2.7
3	L	206	VAL	2.6
2	H	172	SER	2.6
3	L	185	TRP	2.6
3	L	156	LYS	2.6
1	G	124	GLY	2.6
1	G	337	LYS	2.6
2	H	175	LEU	2.5
2	H	173	SER	2.4
3	L	186	LYS	2.3
3	L	183	GLU	2.2
1	G	411	ASN	2.2
1	G	201	ILE	2.2
3	L	187	SER	2.2
1	G	412	GLY	2.2
3	L	157	ALA	2.1
1	G	226	LEU	2.1
2	H	171	GLN	2.1
2	H	132	SER	2.1
1	G	473	GLY	2.0

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

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	PCA	H	1	8/9	0.95	0.12	57,65,76,84	0

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
6	EDO	G	513	4/4	0.83	0.39	76,85,86,89	0
4	NAG	G	508	14/15	0.83	0.19	100,121,151,155	0
4	NAG	G	510	14/15	0.83	0.34	70,114,133,137	0
6	EDO	H	301	4/4	0.87	0.17	71,72,80,84	0
4	NAG	G	507	14/15	0.88	0.26	95,118,122,128	0
6	EDO	H	302	4/4	0.89	0.16	74,86,93,96	0
4	NAG	G	504	14/15	0.90	0.17	87,97,108,109	0
7	CL	H	303	1/1	0.91	0.13	108,108,108,108	0
4	NAG	G	502	14/15	0.91	0.22	89,120,127,130	0
4	NAG	G	506	14/15	0.92	0.18	81,95,108,112	0
4	NAG	G	505	14/15	0.93	0.19	93,99,113,126	0
4	NAG	G	503	14/15	0.95	0.10	58,69,85,86	0
4	NAG	G	509	14/15	0.96	0.14	55,75,91,112	0
4	NAG	G	501	14/15	0.96	0.09	58,76,86,86	0
6	EDO	L	301	4/4	0.97	0.12	55,64,67,77	0
5	EPE	G	511	15/15	0.97	0.15	37,61,78,79	0
6	EDO	G	512	4/4	0.99	0.32	47,54,58,72	0

6.5 Other polymers (i)

There are no such residues in this entry.