



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 6, 2017 – 06:43 PM EST

PDB ID : 4H8W  
Title : Crystal structure of non-neutralizing and ADCC-potent antibody N5-i5 in complex with HIV-1 clade A/E gp120 and sCD4.  
Authors : Tolbert, W.D.; Acharya, P.; Kwong, P.D.; Pazgier, M.  
Deposited on : unknown  
Resolution : 1.85 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

<http://wwpdb.org/validation/2016/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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20030345  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20030345

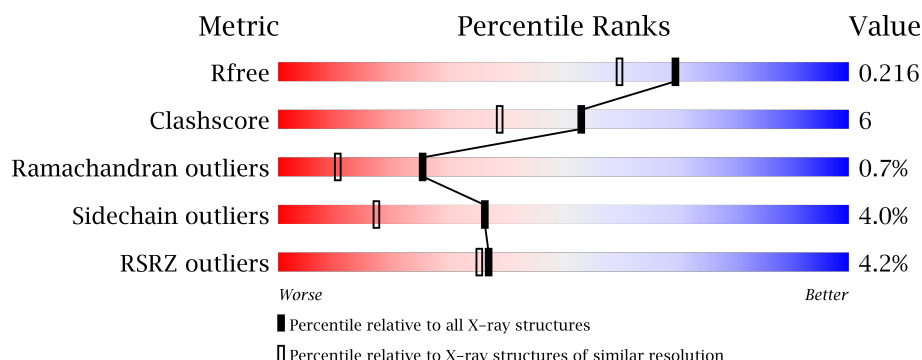
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 1.85 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))
$R_{free}$	100719	1923 (1.86-1.86)
Clashscore	112137	2083 (1.86-1.86)
Ramachandran outliers	110173	2060 (1.86-1.86)
Sidechain outliers	110143	2060 (1.86-1.86)
RSRZ outliers	101464	1932 (1.86-1.86)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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  $\leq 5\%$ . 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	<div> <div>2%</div> <div>82%</div> <div>12%</div> <div>5%</div> </div>
2	H	226	<div> <div>8%</div> <div>81%</div> <div>11%</div> <div>6%</div> </div>
3	L	217	<div> <div>5%</div> <div>82%</div> <div>14%</div> <div>..</div> </div>
4	C	185	<div> <div>2%</div> <div>84%</div> <div>10%</div> <div>5%</div> </div>

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
5	NAG	G	502	-	-	-	X
5	NAG	G	510	-	-	-	X
6	MRD	G	511	-	-	-	X
6	MRD	G	512	-	-	-	X
6	MRD	L	301	-	-	-	X

## 2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 8180 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 93TH057 (H375S) GP120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	337	Total	C	N	O	S	0	6	0
			2667	1681	457	507	22			

- Molecule 2 is a protein called FAB HEAVY CHAIN OF ADCC AND NON-NEUTRALIZING ANTI-HIV-1 ANTIBODY N5-I5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	212	Total	C	N	O	S	0	6	0
			1601	1005	272	318	6			

- Molecule 3 is a protein called FAB LIGHT CHAIN OF ADCC AND NON-NEUTRALIZING ANTI-HIV-1 ANTIBODY N5-I5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	212	Total	C	N	O	S	0	4	0
			1604	997	271	331	5			

- Molecule 4 is a protein called T-cell surface glycoprotein CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	C	176	Total	C	N	O	S	0	2	0
			1376	860	240	272	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	184	ASN	-	EXPRESSION TAG	UNP P01730
C	185	THR	-	EXPRESSION TAG	UNP P01730

- Molecule 5 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

- Molecule 6 is (4R)-2-METHYLPENTANE-2,4-DIOL (three-letter code: MRD) (formula:  $C_6H_{14}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	G	1	Total	C	O	0	0
			8	6	2		
6	G	1	Total	C	O	0	0
			8	6	2		
6	L	1	Total	C	O	0	0
			8	6	2		

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



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

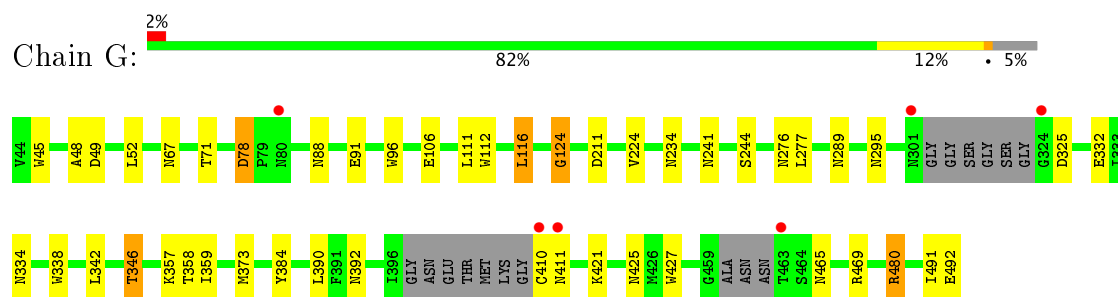
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	G	318	Total 318	O 318	0	0
8	H	180	Total 180	O 180	0	0
8	L	162	Total 162	O 162	0	0
8	C	102	Total 102	O 102	0	0

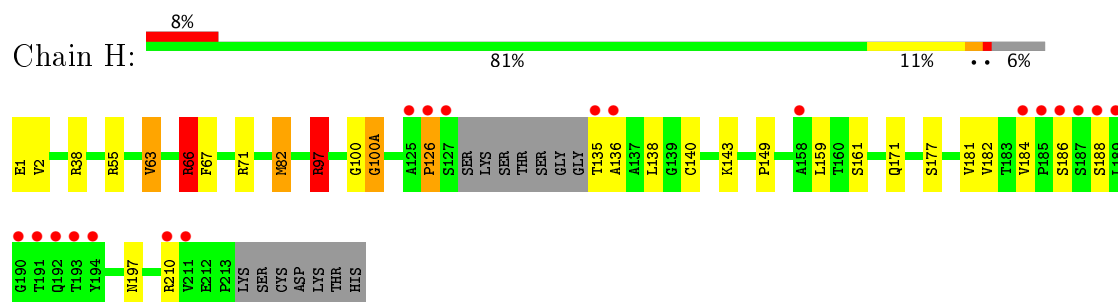
### 3 Residue-property plots [i](#)

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.

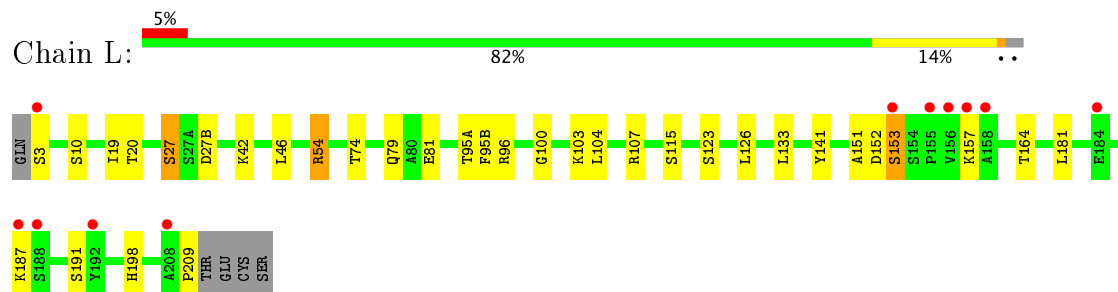
- Molecule 1: HIV-1 CLADE A/E 93TH057 (H375S) GP120



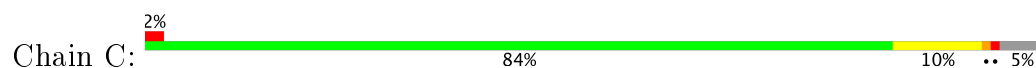
- Molecule 2: FAB HEAVY CHAIN OF ADCC AND NON-NEUTRALIZING ANTI-HIV-1 ANTIBODY N5-I5



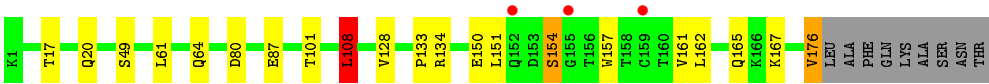
- Molecule 3: FAB LIGHT CHAIN OF ADCC AND NON-NEUTRALIZING ANTI-HIV-1 ANTIBODY N5-I5



- Molecule 4: T-cell surface glycoprotein CD4







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	92.25Å 103.91Å 134.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.24 – 1.85 41.20 – 1.85	Depositor EDS
% Data completeness (in resolution range)	99.6 (41.24-1.85) 99.6 (41.20-1.85)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.17 (at 1.86Å)	Xtriage
Refinement program	PHENIX 5.7.0029, REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.182 , 0.215 0.183 , 0.216	Depositor DCC
$R_{free}$ test set	5508 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	31.0	Xtriage
Anisotropy	0.401	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 50.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	8180	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.45% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: MRD, GOL, NAG

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	1.08	8/2736 (0.3%)	1.16	12/3714 (0.3%)
2	H	1.02	1/1647 (0.1%)	1.38	16/2241 (0.7%)
3	L	1.12	6/1642 (0.4%)	1.17	10/2236 (0.4%)
4	C	0.94	0/1401	1.01	3/1889 (0.2%)
All	All	1.05	15/7426 (0.2%)	1.19	41/10080 (0.4%)

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	G	0	2
2	H	0	2
3	L	0	1
All	All	0	5

All (15) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	96[A]	ARG	CZ-NH1	-7.11	1.23	1.33
3	L	96[B]	ARG	CZ-NH1	-7.11	1.23	1.33
1	G	91	GLU	CD-OE1	7.09	1.33	1.25
3	L	54	ARG	CZ-NH1	6.76	1.41	1.33
2	H	66	ARG	CD-NE	-6.46	1.35	1.46
3	L	95(A)	THR	C-N	5.99	1.47	1.34
1	G	346	THR	CB-CG2	-5.80	1.33	1.52
3	L	10	SER	CB-OG	-5.64	1.34	1.42
1	G	78	ASP	CB-CG	-5.54	1.40	1.51
1	G	492	GLU	C-O	5.32	1.33	1.23
1	G	106	GLU	CD-OE2	-5.31	1.19	1.25

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	100	GLY	N-CA	5.20	1.53	1.46
1	G	45	TRP	CG-CD1	-5.16	1.29	1.36
1	G	480	ARG	CZ-NH2	-5.11	1.26	1.33
1	G	332	GLU	CD-OE2	-5.01	1.20	1.25

All (41) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	66	ARG	NE-CZ-NH2	-24.05	108.28	120.30
2	H	66	ARG	NE-CZ-NH1	20.07	130.33	120.30
3	L	54	ARG	NE-CZ-NH2	-19.66	110.47	120.30
1	G	480	ARG	NE-CZ-NH2	-16.27	112.17	120.30
2	H	97	ARG	NE-CZ-NH2	-14.05	113.28	120.30
1	G	480	ARG	NE-CZ-NH1	13.29	126.95	120.30
2	H	82	MET	CG-SD-CE	-11.72	81.44	100.20
2	H	66	ARG	CD-NE-CZ	11.66	139.93	123.60
3	L	54	ARG	NE-CZ-NH1	10.75	125.68	120.30
3	L	96[A]	ARG	NE-CZ-NH1	-10.65	114.97	120.30
3	L	96[B]	ARG	NE-CZ-NH1	-10.65	114.97	120.30
3	L	27	SER	O-C-N	-10.26	106.28	122.70
2	H	97	ARG	NE-CZ-NH1	9.99	125.30	120.30
3	L	96[A]	ARG	NE-CZ-NH2	9.90	125.25	120.30
3	L	96[B]	ARG	NE-CZ-NH2	9.90	125.25	120.30
2	H	63	VAL	CG1-CB-CG2	9.83	126.62	110.90
1	G	124	GLY	O-C-N	-9.79	106.56	123.20
1	G	469	ARG	NE-CZ-NH1	8.36	124.48	120.30
2	H	55	ARG	NE-CZ-NH1	-8.31	116.15	120.30
1	G	116	LEU	CA-CB-CG	-7.79	97.39	115.30
4	C	80	ASP	CB-CG-OD1	7.35	124.91	118.30
1	G	49	ASP	CB-CG-OD2	-7.21	111.81	118.30
1	G	373	MET	CG-SD-CE	-7.16	88.74	100.20
2	H	97	ARG	CD-NE-CZ	6.82	133.15	123.60
2	H	38	ARG	NE-CZ-NH2	-6.54	117.03	120.30
1	G	469	ARG	NE-CZ-NH2	-6.40	117.10	120.30
2	H	71	ARG	NE-CZ-NH2	-6.20	117.20	120.30
2	H	100(A)	GLY	O-C-N	-6.17	112.83	122.70
1	G	325	ASP	CB-CG-OD1	6.07	123.77	118.30
3	L	27	SER	CA-C-N	5.88	130.12	117.20
2	H	82	MET	O-C-N	-5.87	113.31	122.70
4	C	80	ASP	CB-CG-OD2	-5.85	113.03	118.30
1	G	211	ASP	CB-CG-OD1	5.79	123.51	118.30
2	H	63	VAL	CA-CB-CG1	5.76	119.55	110.90

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	C	108	LEU	CB-CG-CD1	5.67	120.65	111.00
2	H	71	ARG	NE-CZ-NH1	5.48	123.04	120.30
2	H	140	CYS	CA-CB-SG	-5.42	104.25	114.00
3	L	107	ARG	NE-CZ-NH1	5.29	122.95	120.30
3	L	46	LEU	CB-CG-CD2	-5.16	102.23	111.00
1	G	52[A]	LEU	CB-CG-CD1	-5.10	102.33	111.00
1	G	52[B]	LEU	CB-CG-CD1	-5.10	102.33	111.00

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	G	124	GLY	Mainchain
1	G	48	ALA	Mainchain
2	H	100(A)	GLY	Mainchain
2	H	66	ARG	Sidechain
3	L	27	SER	Mainchain

## 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	2667	0	2631	47	0
2	H	1601	0	1579	13	0
3	L	1604	0	1543	12	0
4	C	1376	0	1408	9	0
5	G	140	0	130	32	0
6	G	16	0	28	6	0
6	L	8	0	14	0	0
7	C	6	0	8	0	0
8	C	102	0	0	1	0
8	G	318	0	0	2	0
8	H	180	0	0	4	0
8	L	162	0	0	3	0
All	All	8180	0	7341	83	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 6.

All (83) 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:G:334:ASN:HD21	5:G:508:NAG:C1	1.03	1.63
1:G:392:ASN:HD21	5:G:510:NAG:C1	1.09	1.60
1:G:289:ASN:HD21	5:G:506:NAG:C1	1.07	1.57
1:G:276:ASN:HD21	5:G:505:NAG:C1	0.93	1.56
1:G:234:ASN:HD21	5:G:502:NAG:C1	1.11	1.54
1:G:241:ASN:HD21	5:G:503:NAG:C1	0.91	1.54
1:G:88:ASN:HD21	5:G:501:NAG:C1	0.92	1.52
1:G:289:ASN:HD21	5:G:506:NAG:C2	1.74	0.99
1:G:334:ASN:HD21	5:G:508:NAG:C2	1.80	0.94
3:L:209:PRO:O	8:L:548:HOH:O	1.86	0.93
1:G:392:ASN:HD21	5:G:510:NAG:C2	1.83	0.91
1:G:234:ASN:HD21	5:G:502:NAG:C2	1.83	0.91
2:H:66:ARG:HD3	8:H:463:HOH:O	1.83	0.78
6:G:512:MRD:O2	8:G:889:HOH:O	2.01	0.78
1:G:392:ASN:CG	5:G:510:NAG:C1	2.52	0.77
1:G:234:ASN:ND2	5:G:502:NAG:C2	2.44	0.75
1:G:88:ASN:ND2	5:G:501:NAG:C2	2.53	0.71
3:L:103:LYS:NZ	8:L:562:HOH:O	2.19	0.71
1:G:295:ASN:CG	5:G:507:NAG:C1	2.57	0.71
1:G:67:ASN:O	1:G:71[B]:THR:HG22	1.91	0.70
1:G:334:ASN:CG	5:G:508:NAG:C1	2.58	0.69
1:G:410:CYS:N	5:G:510:NAG:HO6	1.90	0.69
2:H:66:ARG:CD	8:H:463:HOH:O	2.37	0.69
1:G:295:ASN:ND2	5:G:507:NAG:O5	2.26	0.68
1:G:241:ASN:CG	5:G:503:NAG:C1	2.63	0.64
4:C:151:LEU:HD12	4:C:176:VAL:HG21	1.79	0.64
4:C:17:THR:OG1	8:C:391:HOH:O	2.05	0.63
1:G:67:ASN:O	1:G:71[A]:THR:HG23	1.99	0.63
1:G:234:ASN:CG	5:G:502:NAG:C1	2.68	0.62
1:G:295:ASN:ND2	5:G:507:NAG:C2	2.58	0.59
3:L:20[A]:THR:HG22	3:L:74:THR:OG1	2.03	0.58
3:L:19:ILE:HD12	3:L:19:ILE:C	2.25	0.57
1:G:224[A]:VAL:HG11	6:G:511:MRD:H5C2	1.90	0.54
1:G:392:ASN:OD1	5:G:510:NAG:C1	2.56	0.54
4:C:128:VAL:HG22	4:C:161:VAL:HG22	1.90	0.53
1:G:334:ASN:ND2	5:G:508:NAG:C2	2.56	0.53
3:L:79:GLN:OE1	3:L:81:GLU:OE1	2.27	0.53
1:G:289:ASN:CG	5:G:506:NAG:C1	2.65	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:L:133:LEU:HD12	3:L:133:LEU:N	2.25	0.52
3:L:151:ALA:O	3:L:153:SER:N	2.34	0.52
1:G:427:TRP:H	6:G:512:MRD:H5C3	1.75	0.52
1:G:241:ASN:ND2	5:G:503:NAG:C2	2.66	0.51
1:G:425:ASN:O	6:G:512:MRD:H5C1	2.09	0.51
2:H:135:THR:O	2:H:184:VAL:O	2.29	0.51
2:H:97:ARG:HD3	8:H:307:HOH:O	2.11	0.50
1:G:276:ASN:ND2	5:G:505:NAG:C2	2.70	0.50
1:G:427:TRP:N	6:G:512:MRD:H5C3	2.27	0.49
2:H:67:PHE:CZ	2:H:82:MET:HE2	2.48	0.49
3:L:3:SER:N	8:L:545:HOH:O	2.45	0.48
2:H:138:LEU:O	2:H:181:VAL:HG23	2.14	0.48
2:H:1:GLU:HG2	2:H:2:VAL:N	2.29	0.47
1:G:276:ASN:CG	5:G:505:NAG:C1	2.73	0.47
1:G:88:ASN:CG	5:G:501:NAG:C1	2.72	0.47
2:H:1:GLU:HG2	2:H:2:VAL:H	1.79	0.47
3:L:103:LYS:HE3	3:L:104:LEU:O	2.14	0.47
1:G:289:ASN:ND2	5:G:506:NAG:C2	2.49	0.47
1:G:224[B]:VAL:CG1	1:G:244:SER:HB2	2.44	0.47
6:G:512:MRD:H5C2	8:G:632:HOH:O	2.15	0.47
1:G:491:ILE:HG22	1:G:491:ILE:O	2.16	0.46
2:H:143:LYS:HE3	2:H:171[A]:GLN:HE22	1.80	0.46
1:G:346:THR:HG23	1:G:359:ILE:HB	1.98	0.46
1:G:88:ASN:ND2	5:G:501:NAG:O5	2.36	0.45
4:C:20:GLN:HG2	4:C:87:GLU:OE1	2.17	0.44
4:C:150:GLU:O	4:C:176:VAL:HG11	2.18	0.43
1:G:96:TRP:CZ3	1:G:480:ARG:HD3	2.54	0.43
2:H:143:LYS:HD2	2:H:177:SER:OG	2.18	0.43
3:L:42:LYS:HE3	3:L:42:LYS:HB3	1.71	0.43
1:G:112:TRP:CE3	1:G:116:LEU:HD13	2.54	0.43
1:G:342:LEU:HD12	1:G:342:LEU:HA	1.78	0.42
2:H:97:ARG:HG3	2:H:100:GLY:HA3	2.02	0.42
4:C:133:PRO:HD3	4:C:157:TRP:CE2	2.54	0.42
1:G:241:ASN:ND2	5:G:503:NAG:O5	2.32	0.42
2:H:181:VAL:HG22	2:H:182:VAL:N	2.35	0.41
3:L:123:SER:HA	3:L:126:LEU:HD12	2.01	0.41
4:C:151:LEU:HA	4:C:176:VAL:HG21	2.03	0.41
4:C:108:LEU:O	4:C:176:VAL:HA	2.20	0.41
2:H:66:ARG:HD2	8:H:463:HOH:O	2.09	0.41
4:C:151:LEU:O	4:C:154:SER:HB2	2.20	0.41
3:L:141:TYR:O	3:L:198:HIS:HE1	2.03	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:338:TRP:CE2	1:G:390:LEU:HD22	2.56	0.40
1:G:334:ASN:OD1	5:G:508:NAG:C1	2.68	0.40
1:G:384:TYR:CD2	1:G:421:LYS:HD3	2.56	0.40
1:G:358:THR:HB	1:G:465:ASN:HB3	2.04	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	335/353 (95%)	325 (97%)	9 (3%)	1 (0%)	44	29
2	H	214/226 (95%)	206 (96%)	6 (3%)	2 (1%)	20	7
3	L	214/217 (99%)	202 (94%)	9 (4%)	3 (1%)	13	3
4	C	176/185 (95%)	173 (98%)	3 (2%)	0	100	100
All	All	939/981 (96%)	906 (96%)	27 (3%)	6 (1%)	25	13

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	L	152	ASP
2	H	126	PRO
3	L	153	SER
2	H	136	ALA
3	L	95(B)	PHE
1	G	411	ASN

### 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%)	304 (99%)	4 (1%)	73	63
2	H	183/189 (97%)	173 (94%)	10 (6%)	25	9
3	L	181/182 (100%)	173 (96%)	8 (4%)	33	14
4	C	162/167 (97%)	151 (93%)	11 (7%)	18	5
All	All	834/849 (98%)	801 (96%)	33 (4%)	36	17

All (33) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	78	ASP
1	G	111	LEU
1	G	277	LEU
1	G	357	LYS
2	H	63	VAL
2	H	97	ARG
2	H	126	PRO
2	H	149	PRO
2	H	159	LEU
2	H	161	SER
2	H	186	SER
2	H	188	SER
2	H	197	ASN
2	H	210	ARG
3	L	27(B)	ASP
3	L	54	ARG
3	L	115	SER
3	L	157	LYS
3	L	164	THR
3	L	181	LEU
3	L	187	LYS
3	L	191	SER
4	C	49	SER
4	C	61	LEU
4	C	64	GLN
4	C	101	THR
4	C	108	LEU
4	C	134	ARG

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
4	C	154	SER
4	C	162	LEU
4	C	165	GLN
4	C	167	LYS
4	C	176	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	G	88	ASN
1	G	234	ASN
1	G	241	ASN
1	G	276	ASN
1	G	289	ASN
1	G	334	ASN
1	G	392	ASN
1	G	444	ASN
3	L	79	GLN
4	C	110	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](#)

14 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$
7	GOL	C	201	-	5,5,5	0.57	0	5,5,5	0.59	0
5	NAG	G	501	1	14,14,15	0.70	0	15,19,21	2.31	3 (20%)
5	NAG	G	502	1	14,14,15	1.07	0	15,19,21	2.43	6 (40%)
5	NAG	G	503	1	14,14,15	0.71	0	15,19,21	1.82	4 (26%)
5	NAG	G	504	1	14,14,15	1.30	1 (7%)	15,19,21	1.90	6 (40%)
5	NAG	G	505	1	14,14,15	0.71	0	15,19,21	1.33	3 (20%)
5	NAG	G	506	1	14,14,15	1.59	2 (14%)	15,19,21	2.11	4 (26%)
5	NAG	G	507	1	14,14,15	1.00	1 (7%)	15,19,21	1.02	1 (6%)
5	NAG	G	508	1	14,14,15	0.84	1 (7%)	15,19,21	1.42	1 (6%)
5	NAG	G	509	1	14,14,15	1.06	2 (14%)	15,19,21	2.48	5 (33%)
5	NAG	G	510	1	14,14,15	0.58	0	15,19,21	1.34	1 (6%)
6	MRD	G	511	-	7,7,7	0.25	0	9,10,10	0.31	0
6	MRD	G	512	-	7,7,7	0.27	0	9,10,10	0.31	0
6	MRD	L	301	-	7,7,7	0.25	0	9,10,10	0.30	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
7	GOL	C	201	-	-	0/4/4/4	0/0/0/0
5	NAG	G	501	1	-	0/6/23/26	0/1/1/1
5	NAG	G	502	1	-	0/6/23/26	0/1/1/1
5	NAG	G	503	1	-	0/6/23/26	0/1/1/1
5	NAG	G	504	1	-	0/6/23/26	0/1/1/1
5	NAG	G	505	1	-	0/6/23/26	0/1/1/1
5	NAG	G	506	1	-	0/6/23/26	0/1/1/1
5	NAG	G	507	1	-	0/6/23/26	0/1/1/1
5	NAG	G	508	1	-	0/6/23/26	0/1/1/1
5	NAG	G	509	1	-	0/6/23/26	0/1/1/1
5	NAG	G	510	1	-	0/6/23/26	0/1/1/1
6	MRD	G	511	-	-	0/5/5/5	0/0/0/0
6	MRD	G	512	-	-	0/5/5/5	0/0/0/0
6	MRD	L	301	-	-	0/5/5/5	0/0/0/0

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	G	506	NAG	O5-C1	-4.55	1.36	1.43
5	G	504	NAG	O5-C1	-3.96	1.37	1.43
5	G	509	NAG	C1-C2	-2.67	1.48	1.52
5	G	506	NAG	O3-C3	-2.48	1.37	1.43
5	G	508	NAG	O5-C1	-2.37	1.39	1.43
5	G	507	NAG	C1-C2	-2.30	1.49	1.52
5	G	509	NAG	O3-C3	-2.11	1.38	1.43

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	509	NAG	C1-O5-C5	-6.53	103.16	112.17
5	G	502	NAG	C1-O5-C5	-5.70	104.31	112.17
5	G	501	NAG	C1-O5-C5	-4.86	105.47	112.17
5	G	508	NAG	C1-C2-N2	-4.67	102.50	110.49
5	G	502	NAG	C6-C5-C4	-3.52	104.76	113.00
5	G	509	NAG	C3-C4-C5	-3.21	104.56	110.22
5	G	502	NAG	O6-C6-C5	-2.97	101.36	111.34
5	G	501	NAG	O4-C4-C3	-2.92	104.01	110.36
5	G	506	NAG	O3-C3-C4	-2.88	104.08	110.36
5	G	504	NAG	O5-C1-C2	-2.86	107.50	111.47
5	G	503	NAG	O4-C4-C3	-2.85	104.15	110.36
5	G	506	NAG	C1-C2-N2	-2.84	105.64	110.49
5	G	505	NAG	C1-C2-N2	-2.76	105.78	110.49
5	G	507	NAG	O4-C4-C3	-2.70	104.47	110.36
5	G	510	NAG	C1-O5-C5	-2.69	108.46	112.17
5	G	509	NAG	O6-C6-C5	-2.50	102.92	111.34
5	G	504	NAG	C4-C3-C2	-2.32	107.61	111.02
5	G	509	NAG	O3-C3-C2	-2.29	104.49	109.39
5	G	502	NAG	O4-C4-C3	-2.17	105.63	110.36
5	G	505	NAG	O7-C7-C8	-2.16	118.13	122.06
5	G	502	NAG	C4-C3-C2	-2.09	107.95	111.02
5	G	504	NAG	O6-C6-C5	-2.05	104.43	111.34
5	G	505	NAG	C2-N2-C7	2.12	126.04	122.94
5	G	504	NAG	C2-N2-C7	2.27	126.26	122.94
5	G	503	NAG	C3-C4-C5	2.71	115.00	110.22
5	G	504	NAG	C6-C5-C4	2.76	119.45	113.00
5	G	503	NAG	C4-C3-C2	2.88	115.24	111.02
5	G	504	NAG	C1-O5-C5	3.01	116.31	112.17
5	G	503	NAG	O5-C1-C2	3.46	116.29	111.47
5	G	509	NAG	O3-C3-C4	3.63	118.25	110.36
5	G	502	NAG	C1-C2-N2	4.16	117.59	110.49

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	506	NAG	O5-C1-C2	4.20	117.32	111.47
5	G	506	NAG	C1-O5-C5	4.66	118.59	112.17
5	G	501	NAG	O5-C1-C2	5.54	119.19	111.47

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

10 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	501	NAG	4	0
5	G	502	NAG	4	0
5	G	503	NAG	4	0
5	G	505	NAG	3	0
5	G	506	NAG	4	0
5	G	507	NAG	3	0
5	G	508	NAG	5	0
5	G	510	NAG	5	0
6	G	511	MRD	1	0
6	G	512	MRD	5	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 ⓘ

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	G	337/353 (95%)	-0.10	6 (1%) 69 69	23, 34, 55, 76	0
2	H	212/226 (93%)	0.24	19 (8%) 10 10	24, 36, 79, 109	0
3	L	212/217 (97%)	0.09	11 (5%) 28 27	23, 42, 85, 94	0
4	C	176/185 (95%)	0.06	3 (1%) 70 70	27, 44, 64, 83	0
All	All	937/981 (95%)	0.05	39 (4%) 37 35	23, 38, 73, 109	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	187	SER	5.6
2	H	127	SER	5.4
2	H	184	VAL	5.1
2	H	190	GLY	5.0
2	H	191	THR	4.9
2	H	188	SER	4.3
2	H	185	PRO	4.3
2	H	189	LEU	3.9
2	H	192	GLN	3.8
2	H	193	THR	3.6
2	H	194	TYR	3.5
1	G	410	CYS	3.5
2	H	135	THR	3.5
2	H	186	SER	3.4
3	L	157	LYS	3.3
2	H	210	ARG	3.2
3	L	158	ALA	3.1
2	H	125	ALA	3.1
3	L	192	TYR	3.1
1	G	411	ASN	2.9
2	H	126	PRO	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	H	136	ALA	2.9
3	L	155	PRO	2.8
3	L	156	VAL	2.8
1	G	324	GLY	2.7
3	L	3	SER	2.6
3	L	187	LYS	2.5
3	L	208	ALA	2.4
1	G	463	THR	2.4
4	C	152	GLN	2.4
1	G	80	ASN	2.3
4	C	155	GLY	2.3
3	L	153	SER	2.2
4	C	159	CYS	2.2
2	H	158	ALA	2.2
1	G	301	ASN	2.2
3	L	188	SER	2.1
2	H	211	VAL	2.1
3	L	184	GLU	2.1

## 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
6	MRD	G	511	8/8	0.76	0.39	15.99	66,75,79,80	0
6	MRD	G	512	8/8	0.74	0.32	8.72	22,40,45,58	0
5	NAG	G	502	14/15	0.95	0.15	3.48	49,55,57,63	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	MRD	L	301	8/8	0.84	0.20	2.95	63,78,81,82	0
5	NAG	G	510	14/15	0.85	0.29	2.67	60,70,85,89	0
5	NAG	G	507	14/15	0.95	0.14	1.59	46,53,61,65	0
5	NAG	G	504	14/15	0.96	0.11	0.42	30,33,42,47	0
5	NAG	G	506	14/15	0.95	0.10	-0.40	35,46,55,62	0
5	NAG	G	505	14/15	0.93	0.07	-0.70	45,49,54,54	0
5	NAG	G	501	14/15	0.90	0.11	-0.91	50,57,66,69	0
5	NAG	G	503	14/15	0.88	0.26	-	47,65,71,75	0
5	NAG	G	508	14/15	0.90	0.23	-	64,68,75,83	0
7	GOL	C	201	6/6	0.73	0.14	-	63,71,74,76	0
5	NAG	G	509	14/15	0.95	0.07	-	45,48,54,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.