



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 10, 2020 – 12:52 PM BST

PDB ID : 6W4M  
Title : CRYSTAL STRUCTURE OF THE ADCC-POTENT, WEAKLY NEUTRALIZING HIV ENV CO-RECEPTOR BINDING SITE ANTIBODY N12-I2 FAB IN COMPLEX WITH HIV-1 CLADE A/E GP120 AND M48U1  
Authors : Tolbert, W.D.; Gohain, N.; Pazgier, M.  
Deposited on : 2020-03-11  
Resolution : 3.20 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

<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

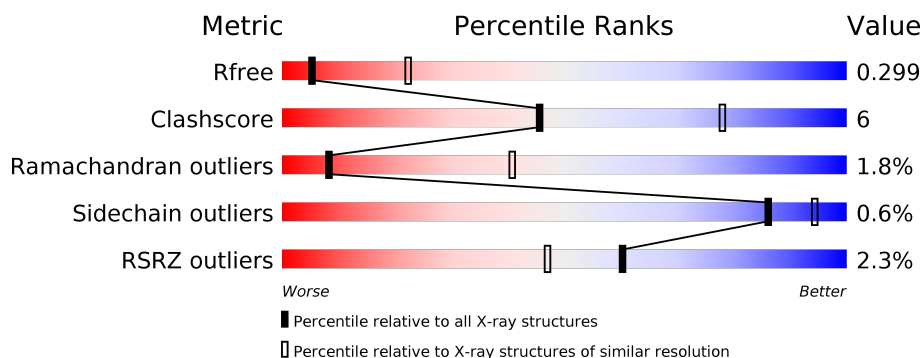
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.20 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 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  $\leq 5\%$ . 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>3%</div> <div> <div></div> <div>82%</div> <div>13%</div> <div>• •</div> </div> </div>
2	L	215	<div> <div>%</div> <div> <div></div> <div>83%</div> <div>16%</div> <div>•</div> </div> </div>
3	H	239	<div> <div>3%</div> <div> <div></div> <div>74%</div> <div>18%</div> <div>• 5%</div> </div> </div>
4	N	28	<div> <div></div> <div> <div></div> <div>71%</div> <div>25%</div> <div>•</div> </div> </div>

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6320 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 clade A/E 93TH057 HIV-1 gp120 core.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	338	Total	C	N	O	S	0	0	0
			2637	1653	456	505	23			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	375	SER	HIS	engineered mutation	UNP A0A0M3KKW9

- Molecule 2 is a protein called ANTI-HIV ANTIBODY N12-I2 FAB LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	213	Total	C	N	O	S	0	0	0
			1630	1017	276	331	6			

- Molecule 3 is a protein called ANTI-HIV ANTIBODY N12-I2 FAB HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	226	Total	C	N	O	S	0	0	0
			1718	1079	283	345	11			

- Molecule 4 is a protein called M48U1 CD4 MIMETIC PEPTIDE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	N	28	Total	C	N	O	S	0	0	1
			209	133	38	32	6			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).

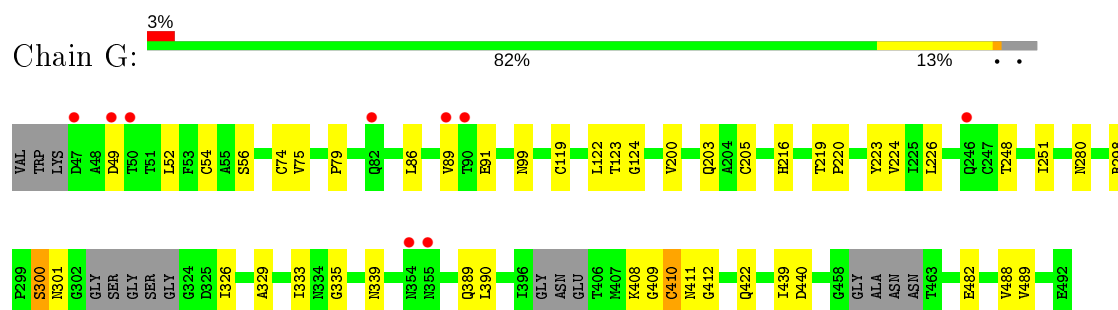


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1	Total	C	N	O	0	0
			14	8	1	5		
5	G	1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

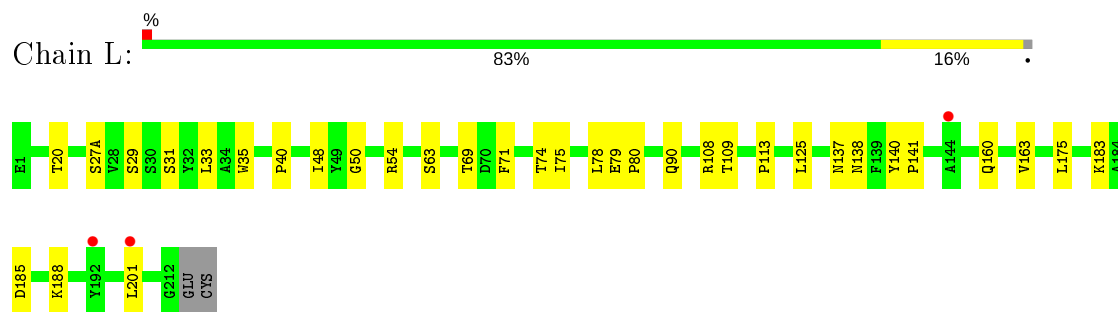
### 3 Residue-property plots [i](#)

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.

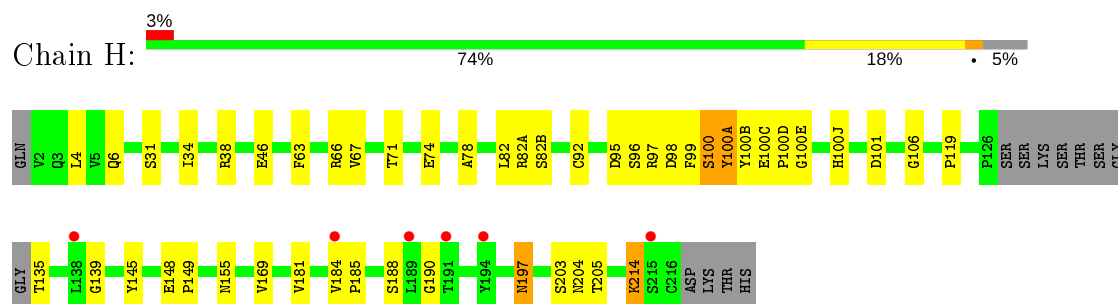
- Molecule 1: clade A/E 93TH057 HIV-1 gp120 core



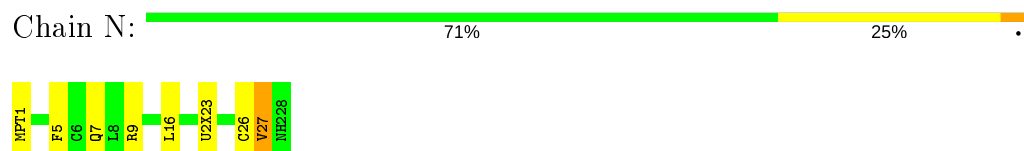
- Molecule 2: ANTI-HIV ANTIBODY N12-I2 FAB LIGHT CHAIN



- Molecule 3: ANTI-HIV ANTIBODY N12-I2 FAB HEAVY CHAIN



- Molecule 4: M48U1 CD4 MIMETIC PEPTIDE



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	52.69Å 69.52Å 213.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.34 – 3.20 42.34 – 3.20	Depositor EDS
% Data completeness (in resolution range)	88.5 (42.34-3.20) 88.5 (42.34-3.20)	Depositor EDS
$R_{merge}$	0.32	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.26 (at 3.19Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660, REFMAC 5.7.0032	Depositor
R, $R_{free}$	0.241 , 0.301 0.242 , 0.299	Depositor DCC
$R_{free}$ test set	613 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	86.0	Xtriage
Anisotropy	0.167	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 19.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	6320	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.33% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: DPR, NAG, MPT, NH2, U2X, TYS

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	0.26	0/2689	0.46	0/3646
2	L	0.26	0/1665	0.47	0/2258
3	H	0.29	0/1725	0.51	0/2349
4	N	0.39	0/176	0.63	0/231
All	All	0.27	0/6255	0.48	0/8484

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
4	N	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	N	1	MPT	Mainchain

### 5.2 Too-close contacts [i](#)

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

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	G	2637	0	2576	26	0
2	L	1630	0	1582	18	0
3	H	1718	0	1650	34	0
4	N	209	0	212	4	0
5	G	126	0	117	1	0
All	All	6320	0	6137	80	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 (80) 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:56:SER:HB3	1:G:74:CYS:HB3	1.52	0.92
3:H:99:PHE:HD1	3:H:100:SER:H	1.22	0.87
3:H:98:ASP:O	3:H:99:PHE:CG	2.34	0.79
3:H:6:GLN:HG3	3:H:106:GLY:H	1.49	0.76
3:H:66:ARG:HA	3:H:82(A):ARG:NH2	2.03	0.74
3:H:66:ARG:HA	3:H:82(A):ARG:HH22	1.53	0.73
1:G:422:GLN:NE2	3:H:31:SER:OG	2.23	0.71
1:G:119:CYS:N	1:G:205:CYS:SG	2.63	0.71
1:G:216:HIS:ND1	1:G:248:THR:O	2.23	0.69
3:H:98:ASP:O	3:H:99:PHE:CD1	2.47	0.68
3:H:99:PHE:HD1	3:H:100:SER:N	1.92	0.67
1:G:410:CYS:O	1:G:412:GLY:N	2.27	0.66
3:H:99:PHE:CD1	3:H:100:SER:N	2.62	0.66
1:G:86:LEU:HB3	1:G:89:VAL:HG21	1.76	0.66
3:H:67:VAL:HG22	3:H:82:LEU:HD13	1.80	0.62
1:G:408:LYS:CD	1:G:409:GLY:H	2.13	0.61
1:G:251:ILE:HD12	1:G:482:GLU:HB3	1.82	0.61
2:L:160:GLN:HB3	3:H:169:VAL:HG11	1.83	0.60
2:L:137:ASN:OD1	2:L:138:ASN:ND2	2.35	0.59
1:G:300:SER:OG	1:G:301:ASN:N	2.32	0.59
2:L:27(A):SER:HA	2:L:69:THR:HG22	1.84	0.58
1:G:91:GLU:HG3	1:G:226:LEU:HD13	1.87	0.56
2:L:185:ASP:HA	2:L:188:LYS:HE2	1.87	0.55
3:H:100(C):GLU:HG3	3:H:100(D):PRO:HD2	1.89	0.55
3:H:34:ILE:HD11	3:H:92:CYS:SG	2.48	0.54
3:H:99:PHE:O	3:H:100(A):TYS:N	2.40	0.54
2:L:48:ILE:HD13	2:L:54:ARG:HA	1.89	0.54
4:N:7:GLN:HE22	4:N:16:LEU:HA	1.72	0.54
1:G:219:THR:HG21	1:G:224:VAL:HA	1.89	0.53

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:49:ASP:OD2	1:G:99:ASN:ND2	2.41	0.53
3:H:197:ASN:O	3:H:197:ASN:ND2	2.41	0.53
1:G:52:LEU:HD11	1:G:488:VAL:HG21	1.89	0.53
2:L:63:SER:HB2	2:L:74:THR:HB	1.91	0.52
3:H:100(D):PRO:N	3:H:100(E):GLY:HA3	2.25	0.52
3:H:203:SER:O	3:H:205:THR:N	2.43	0.51
3:H:63:PHE:HB3	3:H:67:VAL:HG21	1.93	0.51
1:G:54:CYS:HB3	1:G:74:CYS:HA	1.93	0.51
3:H:96:SER:HB3	3:H:101:ASP:OD2	2.11	0.51
2:L:163:VAL:HG22	2:L:175:LEU:HD12	1.93	0.50
1:G:333:ILE:HD13	1:G:390:LEU:HD11	1.93	0.50
3:H:98:ASP:C	3:H:99:PHE:CG	2.85	0.50
1:G:408:LYS:HD3	1:G:409:GLY:H	1.76	0.49
3:H:71:THR:HA	3:H:78:ALA:HA	1.93	0.49
1:G:56:SER:N	1:G:75:VAL:O	2.34	0.48
2:L:35:TRP:HB2	2:L:48:ILE:HB	1.96	0.48
2:L:140:TYR:CD1	2:L:141:PRO:HA	2.49	0.48
4:N:5:PHE:CE2	4:N:9:ARG:HD2	2.50	0.47
3:H:139:GLY:HA3	3:H:181:VAL:HG12	1.96	0.47
3:H:100:SER:OG	3:H:100(A):TYS:N	2.48	0.47
1:G:280:ASN:N	1:G:280:ASN:OD1	2.47	0.46
1:G:123:THR:HG22	1:G:124:GLY:H	1.80	0.46
3:H:99:PHE:C	3:H:100:SER:HG	2.18	0.46
2:L:108:ARG:HG3	2:L:109:THR:H	1.80	0.45
2:L:20:THR:HG22	2:L:74:THR:HG23	1.98	0.45
3:H:38:ARG:NE	3:H:46:GLU:OE1	2.49	0.45
3:H:119:PRO:HB3	3:H:145:TYR:HB3	1.98	0.45
3:H:95:ASP:OD1	3:H:96:SER:N	2.50	0.45
1:G:335:GLY:O	1:G:339:ASN:ND2	2.50	0.45
2:L:113:PRO:HD2	2:L:201:LEU:HD21	1.98	0.44
2:L:31:SER:O	2:L:50:GLY:HA2	2.17	0.44
2:L:33:LEU:HD22	2:L:71:PHE:CG	2.52	0.44
3:H:74:GLU:N	3:H:74:GLU:OE1	2.50	0.43
3:H:67:VAL:CG2	3:H:82:LEU:HD13	2.47	0.43
2:L:125:LEU:HD22	2:L:183:LYS:HG3	2.00	0.43
3:H:148:GLU:HG2	3:H:149:PRO:HA	2.01	0.42
1:G:298:ARG:NH1	1:G:326:ILE:O	2.49	0.42
1:G:389:GLN:HA	5:G:509:NAG:O7	2.20	0.42
3:H:97:ARG:NH2	3:H:100(J):HIS:O	2.47	0.42
1:G:122:LEU:HG	1:G:200:VAL:HG12	2.02	0.42
2:L:75:ILE:HG21	2:L:78:LEU:HD23	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
4:N:26:CYS:O	4:N:27:VAL:HG13	2.20	0.42
1:G:298:ARG:HG2	1:G:329:ALA:HB2	2.02	0.41
3:H:135:THR:HB	3:H:184:VAL:O	2.21	0.41
2:L:140:TYR:CG	2:L:141:PRO:HA	2.55	0.41
3:H:4:LEU:HD23	3:H:92:CYS:SG	2.61	0.41
3:H:185:PRO:HG2	3:H:188:SER:HB2	2.02	0.41
1:G:439:ILE:HG22	1:G:440:ASP:H	1.86	0.40
1:G:220:PRO:HG2	1:G:223:TYR:CD2	2.57	0.40
2:L:79:GLU:HB3	2:L:80:PRO:HD2	2.03	0.40
4:N:16:LEU:N	4:N:27:VAL:O	2.41	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	330/353 (94%)	294 (89%)	30 (9%)	6 (2%)	8	41
2	L	211/215 (98%)	196 (93%)	13 (6%)	2 (1%)	17	56
3	H	220/239 (92%)	198 (90%)	17 (8%)	5 (2%)	6	34
4	N	24/28 (86%)	22 (92%)	1 (4%)	1 (4%)	3	20
All	All	785/835 (94%)	710 (90%)	61 (8%)	14 (2%)	8	41

All (14) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	79	PRO
3	H	204	ASN
4	N	27	VAL
1	G	410	CYS
3	H	155	ASN

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	300	SER
2	L	29	SER
3	H	100	SER
3	H	214	LYS
1	G	203	GLN
1	G	411	ASN
2	L	40	PRO
3	H	190	GLY
1	G	489	VAL

### 5.3.2 Protein sidechains ⓘ

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	302/311 (97%)	302 (100%)	0	100	100
2	L	184/186 (99%)	183 (100%)	1 (0%)	88	95
3	H	191/202 (95%)	188 (98%)	3 (2%)	62	84
4	N	20/20 (100%)	20 (100%)	0	100	100
All	All	697/719 (97%)	693 (99%)	4 (1%)	86	94

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	L	90	GLN
3	H	82(B)	SER
3	H	197	ASN
3	H	214	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	G	301	ASN
2	L	38	GLN
2	L	138	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	H	39	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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$
3	TYS	H	100(A)	3	15,16,17	1.58	3 (20%)	18,22,24	0.95	1 (5%)
4	U2X	N	23	4	19,20,21	1.31	2 (10%)	22,25,27	1.52	5 (22%)
3	TYS	H	100(B)	3	15,16,17	1.63	3 (20%)	18,22,24	0.82	1 (5%)

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	TYS	H	100(A)	3	-	6/10/11/13	0/1/1/1
4	U2X	N	23	4	-	4/10/19/21	0/2/2/2
3	TYS	H	100(B)	3	-	5/10/11/13	0/1/1/1

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	100(A)	TYS	O2-S	4.46	1.64	1.45
3	H	100(B)	TYS	O1-S	4.38	1.64	1.45
4	N	23	U2X	OH-CZ	3.14	1.44	1.37

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	H	100(B)	TYS	OH-CZ	-2.89	1.37	1.42
3	H	100(A)	TYS	OH-CZ	-2.82	1.38	1.42
3	H	100(B)	TYS	OH-S	-2.66	1.54	1.58
3	H	100(A)	TYS	OH-S	-2.20	1.54	1.58
4	N	23	U2X	CA-N	-2.03	1.42	1.48

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	N	23	U2X	CG-CB-CA	4.21	122.62	114.10
3	H	100(A)	TYS	CG-CB-CA	-2.68	108.68	114.10
4	N	23	U2X	C6-C5-C4	2.56	116.64	111.42
4	N	23	U2X	CB-CA-C	-2.26	107.24	111.47
4	N	23	U2X	C6-C1-C2	2.24	115.97	111.42
3	H	100(B)	TYS	O3-S-OH	2.23	111.19	105.83
4	N	23	U2X	C5-C6-C1	2.08	117.57	111.18

There are no chirality outliers.

All (15) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	H	100(A)	TYS	C-CA-CB-CG
3	H	100(A)	TYS	O-C-CA-CB
3	H	100(A)	TYS	CZ-OH-S-O2
3	H	100(B)	TYS	N-CA-CB-CG
3	H	100(B)	TYS	C-CA-CB-CG
3	H	100(B)	TYS	CZ-OH-S-O1
3	H	100(B)	TYS	CZ-OH-S-O3
4	N	23	U2X	CE1-CZ-OH-C7
4	N	23	U2X	CE2-CZ-OH-C7
3	H	100(A)	TYS	N-CA-CB-CG
4	N	23	U2X	C4-C3-C7-OH
3	H	100(A)	TYS	CZ-OH-S-O3
3	H	100(A)	TYS	CZ-OH-S-O1
3	H	100(B)	TYS	CZ-OH-S-O2
4	N	23	U2X	C2-C3-C7-OH

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	H	100(A)	TYS	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

9 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	NAG	G	504	1	14,14,15	0.29	0	17,19,21	0.41	0
5	NAG	G	507	1	14,14,15	0.22	0	17,19,21	0.33	0
5	NAG	G	503	1	14,14,15	0.45	0	17,19,21	0.51	0
5	NAG	G	502	1	14,14,15	0.62	0	17,19,21	0.63	0
5	NAG	G	509	1	14,14,15	0.66	1 (7%)	17,19,21	0.58	0
5	NAG	G	505	1	14,14,15	0.28	0	17,19,21	0.51	0
5	NAG	G	501	1	14,14,15	0.28	0	17,19,21	0.36	0
5	NAG	G	506	1	14,14,15	0.17	0	17,19,21	0.48	0
5	NAG	G	508	1	14,14,15	0.31	0	17,19,21	0.45	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
5	NAG	G	504	1	-	0/6/23/26	0/1/1/1
5	NAG	G	507	1	-	2/6/23/26	0/1/1/1
5	NAG	G	503	1	-	0/6/23/26	0/1/1/1
5	NAG	G	502	1	-	1/6/23/26	0/1/1/1
5	NAG	G	509	1	-	1/6/23/26	0/1/1/1
5	NAG	G	505	1	-	0/6/23/26	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	NAG	G	501	1	-	2/6/23/26	0/1/1/1
5	NAG	G	506	1	-	2/6/23/26	0/1/1/1
5	NAG	G	508	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	509	NAG	C1-C2	2.12	1.55	1.52

There are no bond angle outliers.

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	507	NAG	O5-C5-C6-O6
5	G	501	NAG	O5-C5-C6-O6
5	G	507	NAG	C4-C5-C6-O6
5	G	501	NAG	C4-C5-C6-O6
5	G	508	NAG	O5-C5-C6-O6
5	G	502	NAG	O5-C5-C6-O6
5	G	506	NAG	C4-C5-C6-O6
5	G	509	NAG	C4-C5-C6-O6
5	G	506	NAG	O5-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	G	509	NAG	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

### 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	338/353 (95%)	-0.02	9 (2%) 54 39	47, 78, 116, 145	0
2	L	213/215 (99%)	0.09	3 (1%) 75 63	50, 86, 124, 138	0
3	H	224/239 (93%)	-0.12	6 (2%) 54 39	50, 71, 106, 127	0
4	N	24/28 (85%)	-0.30	0 100 100	66, 88, 106, 120	0
All	All	799/835 (95%)	-0.03	18 (2%) 60 47	47, 78, 118, 145	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	49	ASP	8.5
1	G	354	ASN	4.8
1	G	82	GLN	3.8
1	G	90	THR	3.7
2	L	201	LEU	3.7
1	G	50	THR	3.4
1	G	47	ASP	3.2
1	G	89	VAL	3.1
3	H	138	LEU	3.0
1	G	246	GLN	2.8
3	H	194	TYR	2.6
2	L	144	ALA	2.3
3	H	215	SER	2.3
3	H	189	LEU	2.2
1	G	355	ASN	2.2
3	H	184	VAL	2.1
3	H	191	THR	2.0
2	L	192	TYR	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<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	B-factors(Å <sup>2</sup> )	Q<0.9
3	TYS	H	100(A)	16/17	0.87	0.25	73,83,129,155	0
3	TYS	H	100(B)	16/17	0.89	0.21	73,82,94,97	0
4	U2X	N	23	19/20	0.95	0.20	37,51,66,68	0
4	DPR	N	21	7/8	0.97	0.17	63,69,76,80	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<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	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	G	509	14/15	0.78	0.31	71,87,104,113	0
5	NAG	G	502	14/15	0.85	0.33	77,84,93,97	0
5	NAG	G	501	14/15	0.88	0.32	63,76,92,106	0
5	NAG	G	507	14/15	0.91	0.21	51,71,84,92	0
5	NAG	G	506	14/15	0.91	0.21	62,69,78,91	0
5	NAG	G	508	14/15	0.93	0.20	46,57,78,81	0
5	NAG	G	505	14/15	0.94	0.18	46,69,83,87	0
5	NAG	G	504	14/15	0.94	0.17	52,69,81,93	0
5	NAG	G	503	14/15	0.96	0.14	42,50,60,69	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.