



# wwPDB X-ray Structure Validation Summary Report

Sep 17, 2023 – 10:54 PM EDT

PDB ID : 4YDJ  
Title : Crystal structure of broadly and potently neutralizing antibody 44-VRC13.01  
in complex with HIV-1 clade AE strain 93TH057 gp120  
Authors : Zhou, T.; Moquin, S.; Kwong, P.D.  
Deposited on : 2015-02-22  
Resolution : 2.31 Å (reported)

This is a wwPDB X-ray Structure Validation Summary 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

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

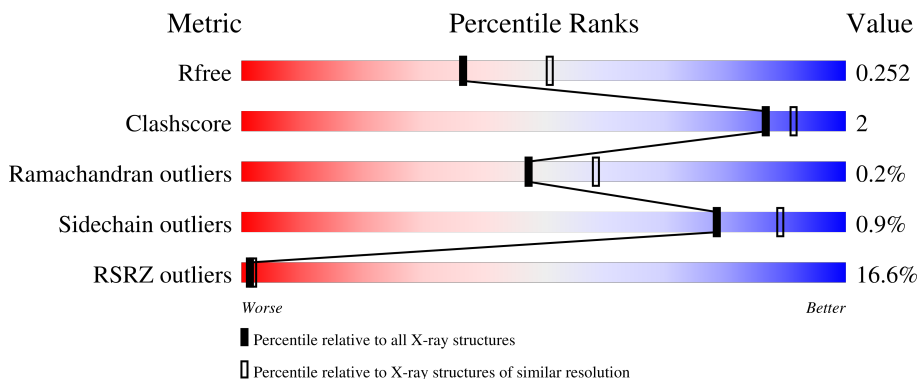
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.31 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 of 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	 28% 79% 5% 16%
1	I	353	 27% 76% 6% 17%
2	A	238	 5% 97%
2	H	238	 7% 95%
3	B	206	 4% 94% 5%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	L	206	<p>7% 94% 5%</p>

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
4	NAG	G	502	-	-	-	X
4	NAG	I	506	-	-	-	X
4	NAG	L	301	-	-	-	X
8	PEG	B	304	-	-	-	X
8	PEG	B	305	-	-	-	X

## 2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 23589 atoms, of which 11493 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 Envelope glycoprotein gp160,Envelope glycoprotein gp160.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
1	G	298	4646	1476	2295	411	445	19	0	0	0
1	I	292	4555	1451	2248	402	436	18	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	197	GLY	-	linker	UNP Q0ED31
G	198	GLY	-	linker	UNP Q0ED31
G	318	GLY	-	linker	UNP Q0ED31
G	319	GLY	-	linker	UNP Q0ED31
G	320	SER	-	linker	UNP Q0ED31
G	321	GLY	-	linker	UNP Q0ED31
G	322	SER	-	linker	UNP Q0ED31
G	323	GLY	-	linker	UNP Q0ED31
I	197	GLY	-	linker	UNP Q0ED31
I	198	GLY	-	linker	UNP Q0ED31
I	318	GLY	-	linker	UNP Q0ED31
I	319	GLY	-	linker	UNP Q0ED31
I	320	SER	-	linker	UNP Q0ED31
I	321	GLY	-	linker	UNP Q0ED31
I	322	SER	-	linker	UNP Q0ED31
I	323	GLY	-	linker	UNP Q0ED31

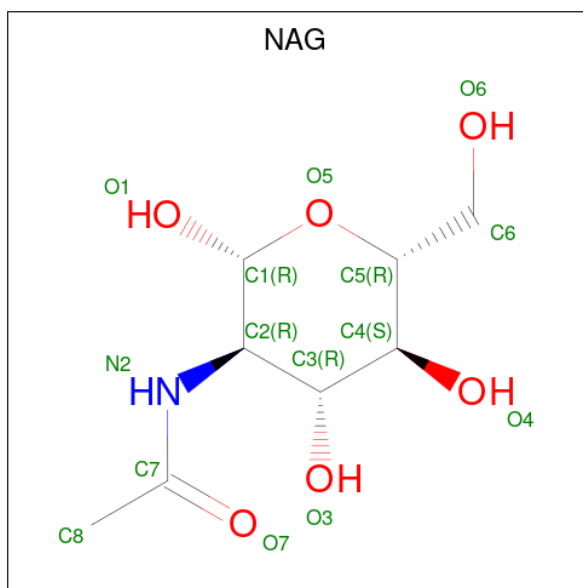
- Molecule 2 is a protein called HEAVY CHAIN OF ANTIBODY 44-VRC13.01.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	H	N	O	S			
2	H	238	3583	1141	1787	308	339	8	0	0	0
2	A	238	3582	1141	1786	308	339	8	0	0	0

- Molecule 3 is a protein called LIGHT CHAIN OF ANTIBODY 44-VRC13.01.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace	
			Total	C	H	N	O				S
3	L	205	Total	C	H	N	O	S	0	0	0
			3052	967	1500	266	313	6			
3	B	205	Total	C	H	N	O	S	0	0	0
			3049	967	1497	266	313	6			

- Molecule 4 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>).



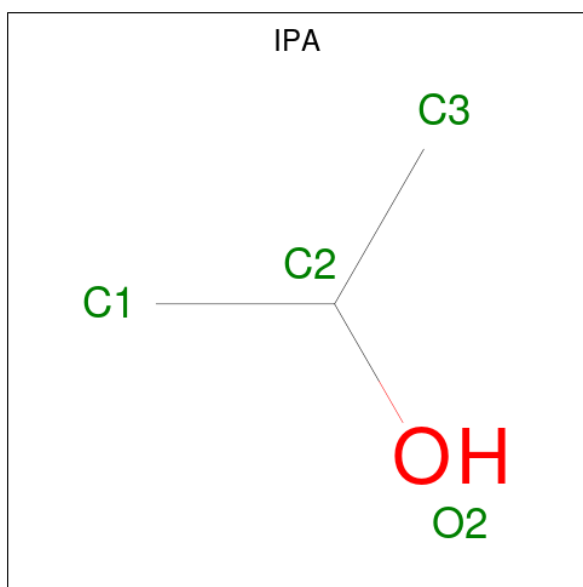
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	H	N	O		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

Continued on next page...

*Continued from previous page...*

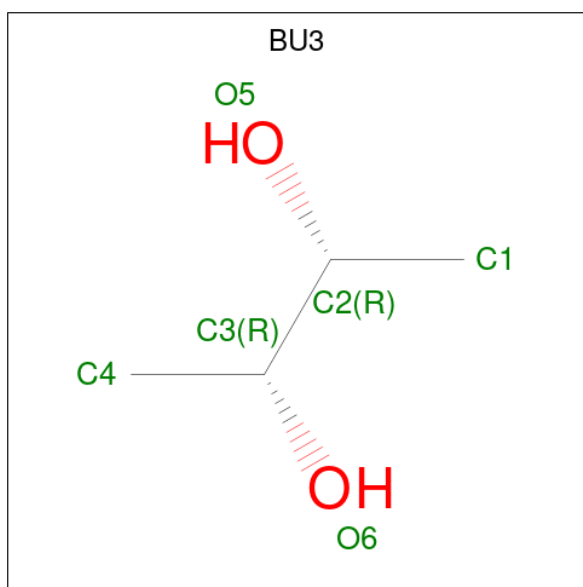
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	G	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	H	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	L	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	I	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	A	1	Total	C	H	N	O	0	0
			27	8	13	1	5		
4	B	1	Total	C	H	N	O	0	0
			27	8	13	1	5		

- Molecule 5 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C<sub>3</sub>H<sub>8</sub>O).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	H	1	Total	C	H	O	0	0
			12	3	8	1		
5	L	1	Total	C	H	O	0	0
			12	3	8	1		
5	A	1	Total	C	H	O	0	0
			12	3	8	1		

- Molecule 6 is (R,R)-2,3-BUTANEDIOL (three-letter code: BU3) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	H	1	Total	C	H	O	0	0
			16	4	10	2		

*Continued on next page...*

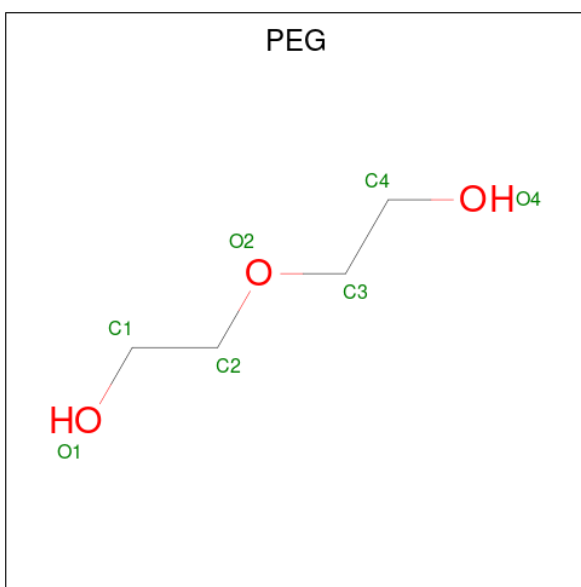
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total	C	H	O	0	0
			16	4	10	2		
6	B	1	Total	C	H	O	0	0
			16	4	10	2		
6	B	1	Total	C	H	O	0	0
			16	4	10	2		

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	L	1	Total	Na	0	0
			1	1		

- Molecule 8 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	A	1	Total	C	H	O	0	0
			17	4	10	3		
8	B	1	Total	C	H	O	0	0
			17	4	10	3		
8	B	1	Total	C	H	O	0	0
			17	4	10	3		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	B	1	Total Cl 1 1	0	0

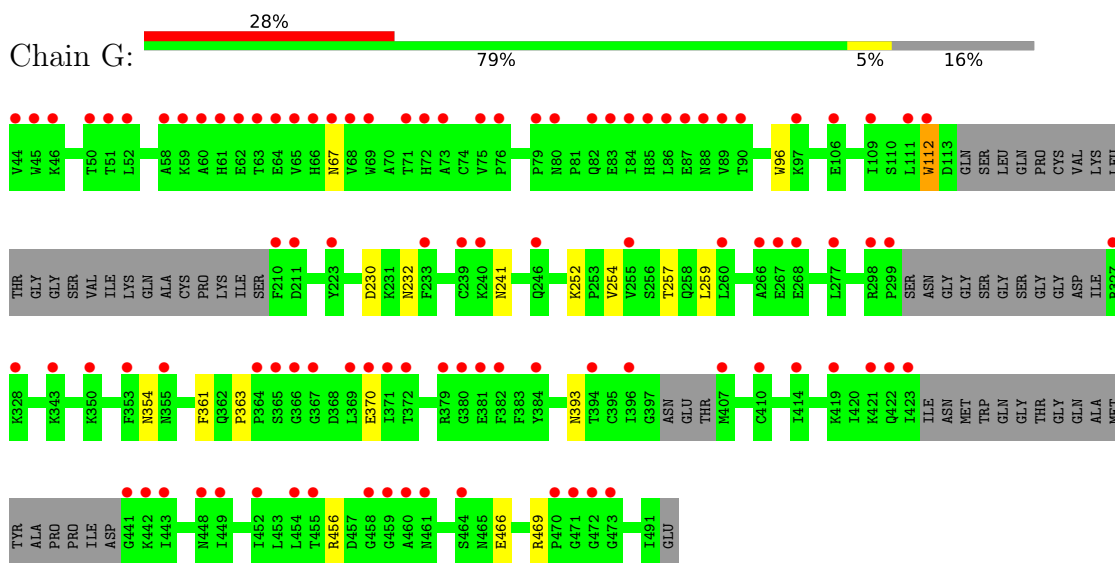
- Molecule 10 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
10	G	7	Total O 7 7	0	0
10	H	90	Total O 90 90	0	0
10	L	78	Total O 78 78	0	0
10	I	3	Total O 3 3	0	0
10	A	97	Total O 97 97	0	0
10	B	100	Total O 100 100	0	0

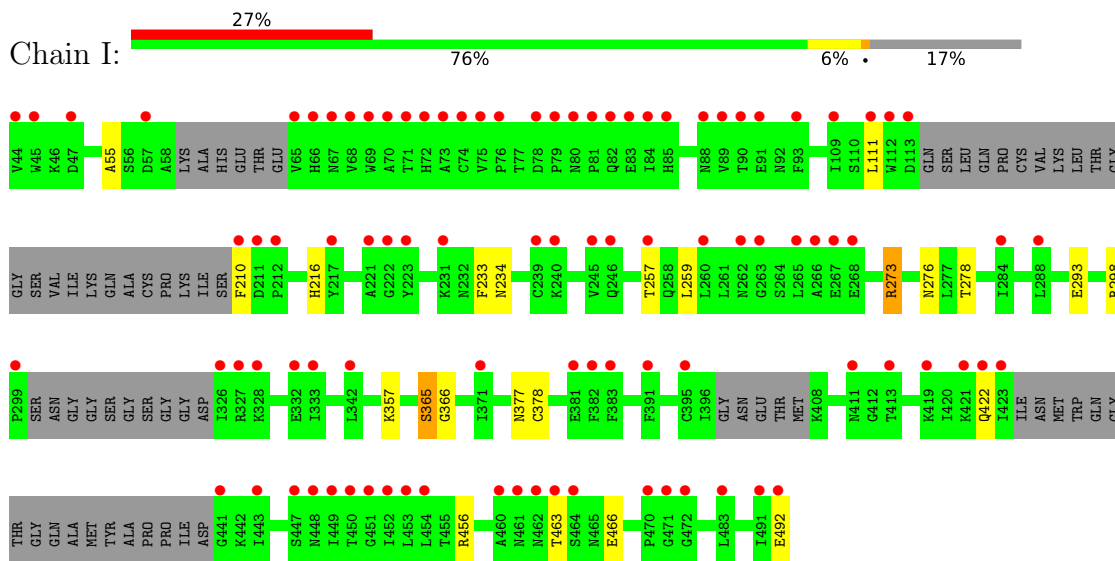
### 3 Residue-property plots i

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.

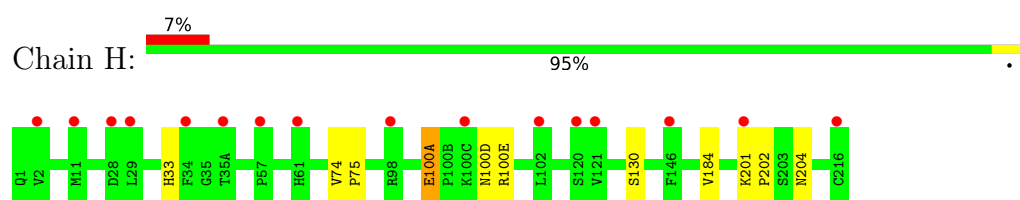
- Molecule 1: Envelope glycoprotein gp160,Envelope glycoprotein gp160



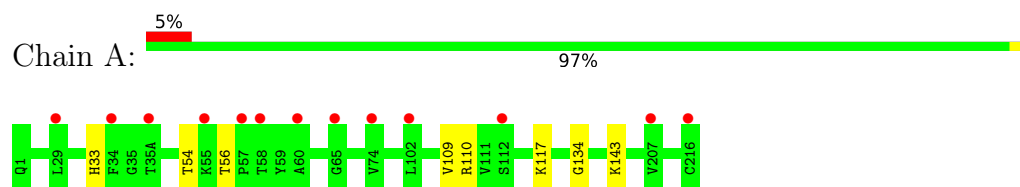
- Molecule 1: Envelope glycoprotein gp160,Envelope glycoprotein gp160



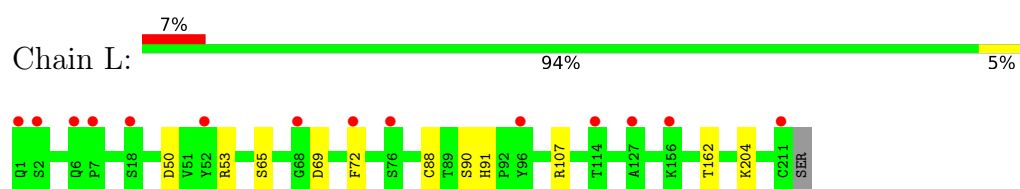
- Molecule 2: HEAVY CHAIN OF ANTIBODY 44-VRC13.01



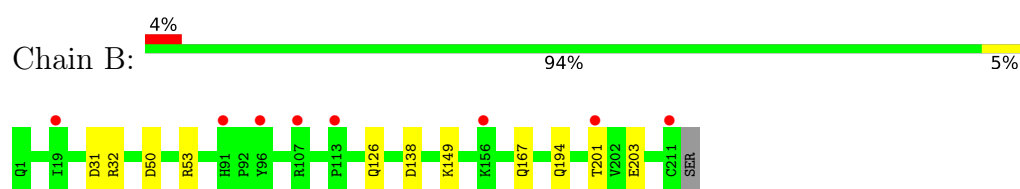
- Molecule 2: HEAVY CHAIN OF ANTIBODY 44-VRC13.01



- Molecule 3: LIGHT CHAIN OF ANTIBODY 44-VRC13.01



- Molecule 3: LIGHT CHAIN OF ANTIBODY 44-VRC13.01



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	293.80Å 67.05Å 93.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	38.30 – 2.31 38.30 – 2.31	Depositor EDS
% Data completeness (in resolution range)	98.9 (38.30-2.31) 92.8 (38.30-2.31)	Depositor EDS
$R_{merge}$	0.14	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.96 (at 2.31Å)	Xtrriage
Refinement program	PHENIX dev_1702	Depositor
R, $R_{free}$	0.223 , 0.249 0.227 , 0.252	Depositor DCC
$R_{free}$ test set	4079 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	49.9	Xtrriage
Anisotropy	0.361	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 51.9	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	23589	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	98.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.55 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.8071e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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: NAG, BU3, CL, NA, IPA, PEG

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	0.22	0/2398	0.37	0/3249
1	I	0.21	0/2352	0.38	0/3187
2	A	0.24	0/1844	0.42	0/2510
2	H	0.24	0/1844	0.42	0/2510
3	B	0.24	0/1591	0.41	0/2170
3	L	0.25	0/1591	0.41	0/2170
All	All	0.23	0/11620	0.40	0/15796

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [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	2351	2295	2285	11	0
1	I	2307	2248	2245	11	0
2	A	1796	1786	1781	5	0
2	H	1796	1787	1782	7	0
3	B	1552	1497	1493	9	0
3	L	1552	1500	1494	7	0
4	A	14	13	13	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	14	13	13	0	0
4	G	126	117	117	1	0
4	H	14	13	13	0	0
4	I	126	117	117	0	0
4	L	14	13	13	0	0
5	A	4	8	8	1	0
5	H	4	8	8	0	0
5	L	4	8	8	1	0
6	A	6	10	10	0	0
6	B	12	20	20	4	0
6	H	6	10	10	0	0
7	L	1	0	0	0	0
8	A	7	10	10	0	0
8	B	14	20	20	0	0
9	B	1	0	0	0	0
10	A	97	0	0	3	0
10	B	100	0	0	1	0
10	G	7	0	0	0	0
10	H	90	0	0	1	1
10	I	3	0	0	0	0
10	L	78	0	0	1	1
All	All	12096	11493	11460	48	1

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 2.

The worst 5 of 48 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:I:233:PHE:O	1:I:273:ARG:NH2	2.25	0.69
2:H:204:ASN:ND2	10:H:401:HOH:O	2.29	0.66
2:A:117:LYS:NZ	10:A:401:HOH:O	2.30	0.65
2:A:134:GLY:H	5:A:302:IPA:H12	1.67	0.60
1:I:276:ASN:OD1	1:I:278:THR:OG1	2.20	0.60

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
10:H:425:HOH:O	10:L:412:HOH:O[3_453]	2.16	0.04

## 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	288/353 (82%)	266 (92%)	21 (7%)	1 (0%)	41	50
1	I	280/353 (79%)	253 (90%)	25 (9%)	2 (1%)	22	26
2	A	236/238 (99%)	229 (97%)	7 (3%)	0	100	100
2	H	236/238 (99%)	230 (98%)	6 (2%)	0	100	100
3	B	203/206 (98%)	191 (94%)	12 (6%)	0	100	100
3	L	203/206 (98%)	192 (95%)	11 (5%)	0	100	100
All	All	1446/1594 (91%)	1361 (94%)	82 (6%)	3 (0%)	47	58

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	354	ASN
1	I	422	GLN
1	I	365	SER

### 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	268/311 (86%)	267 (100%)	1 (0%)	91	96
1	I	264/311 (85%)	260 (98%)	4 (2%)	65	79
2	A	202/202 (100%)	200 (99%)	2 (1%)	76	87
2	H	202/202 (100%)	199 (98%)	3 (2%)	65	79

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
3	B	175/176 (99%)	175 (100%)	0	100	100
3	L	175/176 (99%)	174 (99%)	1 (1%)	86	94
All	All	1286/1378 (93%)	1275 (99%)	11 (1%)	78	89

5 of 11 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	I	273	ARG
1	I	298	ARG
2	A	109	VAL
2	A	33	HIS
3	L	88	CYS

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

Mol	Chain	Res	Type
3	L	66	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 34 ligands modelled in this entry, 2 are monoatomic - leaving 32 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
8	PEG	A	304	-	6,6,6	0.54	0	5,5,5	0.28	0
4	NAG	I	504	1	14,14,15	0.28	0	17,19,21	0.40	0
4	NAG	G	501	1	14,14,15	0.20	0	17,19,21	0.45	0
4	NAG	B	301	3	14,14,15	0.27	0	17,19,21	0.40	0
4	NAG	G	505	1	14,14,15	0.28	0	17,19,21	0.49	0
4	NAG	G	504	1	14,14,15	0.25	0	17,19,21	0.42	0
6	BU3	B	302	-	4,5,5	1.36	0	6,6,6	0.31	0
4	NAG	G	509	1	14,14,15	0.25	0	17,19,21	0.52	0
4	NAG	I	509	1	14,14,15	0.26	0	17,19,21	0.46	0
4	NAG	A	301	2	14,14,15	0.19	0	17,19,21	0.51	0
8	PEG	B	305	-	6,6,6	0.53	0	5,5,5	0.48	0
6	BU3	H	303	-	4,5,5	1.37	0	6,6,6	0.35	0
4	NAG	G	508	1	14,14,15	0.22	0	17,19,21	0.38	0
4	NAG	G	506	1	14,14,15	0.25	0	17,19,21	0.46	0
6	BU3	B	303	-	4,5,5	1.42	0	6,6,6	0.39	0
4	NAG	I	508	1	14,14,15	0.26	0	17,19,21	0.46	0
5	IPA	A	302	-	3,3,3	0.55	0	3,3,3	0.18	0
4	NAG	I	503	1	14,14,15	0.23	0	17,19,21	0.41	0
4	NAG	I	507	1	14,14,15	0.21	0	17,19,21	0.53	0
4	NAG	I	505	1	14,14,15	0.23	0	17,19,21	0.47	0
4	NAG	I	501	1	14,14,15	0.21	0	17,19,21	0.48	0
5	IPA	L	302	-	3,3,3	0.53	0	3,3,3	0.14	0
4	NAG	G	503	1	14,14,15	0.23	0	17,19,21	0.43	0
4	NAG	G	507	1	14,14,15	0.20	0	17,19,21	0.52	0
5	IPA	H	302	-	3,3,3	0.57	0	3,3,3	0.18	0
6	BU3	A	303	-	4,5,5	1.40	0	6,6,6	0.37	0
4	NAG	L	301	3	14,14,15	0.23	0	17,19,21	0.47	0
4	NAG	G	502	1	14,14,15	0.51	0	17,19,21	0.70	1 (5%)
8	PEG	B	304	-	6,6,6	0.54	0	5,5,5	0.39	0
4	NAG	I	502	1	14,14,15	0.24	0	17,19,21	0.41	0
4	NAG	I	506	1	14,14,15	0.25	0	17,19,21	0.41	0
4	NAG	H	301	2	14,14,15	0.24	0	17,19,21	0.49	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
8	PEG	A	304	-	-	2/4/4/4	-
4	NAG	I	504	1	-	0/6/23/26	0/1/1/1
4	NAG	G	501	1	-	2/6/23/26	0/1/1/1
4	NAG	B	301	3	-	1/6/23/26	0/1/1/1
4	NAG	G	505	1	-	0/6/23/26	0/1/1/1
4	NAG	G	504	1	-	1/6/23/26	0/1/1/1
6	BU3	B	302	-	-	1/4/4/4	-
4	NAG	G	509	1	-	2/6/23/26	0/1/1/1
4	NAG	I	509	1	-	3/6/23/26	0/1/1/1
4	NAG	A	301	2	-	4/6/23/26	0/1/1/1
8	PEG	B	305	-	-	3/4/4/4	-
6	BU3	H	303	-	-	0/4/4/4	-
4	NAG	G	508	1	-	0/6/23/26	0/1/1/1
4	NAG	G	506	1	-	0/6/23/26	0/1/1/1
6	BU3	B	303	-	-	0/4/4/4	-
4	NAG	I	508	1	-	2/6/23/26	0/1/1/1
4	NAG	I	503	1	-	2/6/23/26	0/1/1/1
4	NAG	I	507	1	-	3/6/23/26	0/1/1/1
4	NAG	I	505	1	-	0/6/23/26	0/1/1/1
4	NAG	I	501	1	-	3/6/23/26	0/1/1/1
4	NAG	G	503	1	-	2/6/23/26	0/1/1/1
4	NAG	G	507	1	-	2/6/23/26	0/1/1/1
6	BU3	A	303	-	-	0/4/4/4	-
4	NAG	L	301	3	-	4/6/23/26	0/1/1/1
4	NAG	G	502	1	-	3/6/23/26	0/1/1/1
8	PEG	B	304	-	-	1/4/4/4	-
4	NAG	I	502	1	-	2/6/23/26	0/1/1/1
4	NAG	I	506	1	-	2/6/23/26	0/1/1/1
4	NAG	H	301	2	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	G	502	NAG	C1-O5-C5	2.27	115.27	112.19

There are no chirality outliers.

5 of 47 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	I	503	NAG	O5-C5-C6-O6
4	G	502	NAG	O5-C5-C6-O6
4	I	508	NAG	O5-C5-C6-O6
4	G	509	NAG	O5-C5-C6-O6
4	I	503	NAG	C4-C5-C6-O6

There are no ring outliers.

5 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	B	302	BU3	3	0
6	B	303	BU3	1	0
5	A	302	IPA	1	0
5	L	302	IPA	1	0
4	G	502	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 [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, 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	298/353 (84%)	1.72	99 (33%) 0 0	66, 119, 181, 232	0
1	I	292/353 (82%)	1.67	95 (32%) 0 0	65, 130, 194, 260	0
2	A	238/238 (100%)	0.41	13 (5%) 25 31	36, 57, 105, 153	0
2	H	238/238 (100%)	0.45	16 (6%) 17 23	37, 63, 109, 144	0
3	B	205/206 (99%)	0.31	8 (3%) 39 46	35, 56, 100, 163	0
3	L	205/206 (99%)	0.52	14 (6%) 17 22	41, 65, 132, 202	0
All	All	1476/1594 (92%)	0.93	245 (16%) 1 2	35, 78, 167, 260	0

The worst 5 of 245 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	G	72	HIS	17.6
1	G	63	THR	13.7
1	I	66	HIS	10.2
1	G	210	PHE	9.3
1	I	112	TRP	9.0

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

There are no non-standard protein/DNA/RNA residues in this entry.

### 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands i

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<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
8	PEG	B	304	7/7	0.45	0.45	111,133,134,134	0
8	PEG	A	304	7/7	0.53	0.32	89,107,109,109	0
4	NAG	G	502	14/15	0.70	0.69	86,97,115,115	0
8	PEG	B	305	7/7	0.70	0.54	93,112,117,117	0
4	NAG	G	509	14/15	0.73	0.32	81,91,109,109	0
4	NAG	L	301	14/15	0.74	0.48	94,103,124,125	0
4	NAG	B	301	14/15	0.75	0.35	79,92,109,112	0
4	NAG	I	506	14/15	0.75	0.41	104,110,132,133	0
5	IPA	L	302	4/4	0.77	0.21	78,96,97,97	0
4	NAG	A	301	14/15	0.77	0.34	71,82,99,99	0
9	CL	B	306	1/1	0.78	0.11	106,106,106,106	0
6	BU3	H	303	6/6	0.79	0.15	96,116,117,117	0
4	NAG	G	506	14/15	0.80	0.21	81,86,103,106	0
4	NAG	G	504	14/15	0.80	0.28	88,97,116,119	0
4	NAG	I	508	14/15	0.81	0.27	72,79,94,96	0
6	BU3	B	302	6/6	0.81	0.23	80,97,98,98	0
4	NAG	I	501	14/15	0.81	0.27	84,92,110,111	0
4	NAG	G	507	14/15	0.81	0.33	79,90,107,107	0
4	NAG	I	507	14/15	0.81	0.40	96,107,128,129	0
5	IPA	A	302	4/4	0.81	0.15	64,77,79,81	0
4	NAG	I	509	14/15	0.82	0.36	81,88,105,107	0
4	NAG	I	502	14/15	0.86	0.51	93,100,120,120	0
4	NAG	I	503	14/15	0.86	0.24	80,93,112,112	0
6	BU3	A	303	6/6	0.86	0.17	105,126,127,127	0
4	NAG	G	505	14/15	0.86	0.14	63,69,82,84	0
4	NAG	G	501	14/15	0.87	0.21	77,83,99,101	0
6	BU3	B	303	6/6	0.87	0.23	104,125,126,126	0
4	NAG	I	504	14/15	0.88	0.34	83,93,113,113	0
4	NAG	G	508	14/15	0.89	0.17	70,78,93,95	0
4	NAG	G	503	14/15	0.89	0.16	75,82,99,100	0
4	NAG	H	301	14/15	0.89	0.36	80,91,109,110	0
5	IPA	H	302	4/4	0.90	0.15	80,96,96,96	0
4	NAG	I	505	14/15	0.91	0.16	75,79,94,96	0
7	NA	L	303	1/1	0.97	0.16	59,59,59,59	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.