



# wwPDB X-ray Structure Validation Summary Report

Oct 18, 2023 – 11:25 AM EDT

PDB ID : 1YYL  
Title : crystal structure of CD4M33, a scorpion-toxin mimic of CD4, in complex with HIV-1 YU2 gp120 envelope glycoprotein and anti-HIV-1 antibody 17b  
Authors : Huang, C.C.; Stricher, F.; Martin, L.; Decker, J.M.; Majeed, S.; Barthe, P.; Hendrickson, W.A.; Robinson, J.; Roumestand, C.; Sodroski, J.; Wyatt, R.; Shaw, G.M.; Vita, C.; Kwong, P.D.  
Deposited on : 2005-02-25  
Resolution : 2.75 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 : 2.36  
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) : 2.36

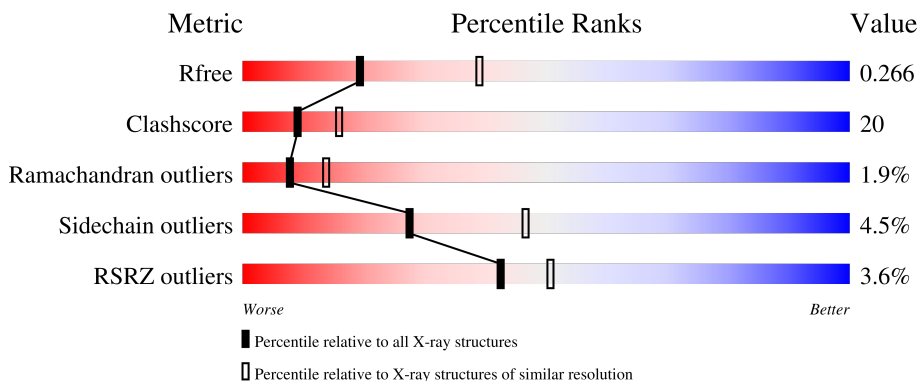
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.75 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	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)

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	313	 4% 64% 30% . .
1	P	313	 7% 55% 39% . .
2	L	214	 62% 36% .
2	Q	214	 61% 36% .

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	H	229	<p>4% 60% 36%</p>
3	R	229	<p>4% 64% 33%</p>
4	M	27	<p>63% 26% 11%</p>
4	S	27	<p>63% 26% 11%</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
5	NAG	G	734	X	-	-	-
5	NAG	P	1741	-	-	-	X

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 12370 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 Exterior membrane glycoprotein(GP120),Exterior membrane glycoprotein(GP120),Exterior membrane glycoprotein(GP120).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	G	301	Total	C	N	O	S	0	0	0
			2341	1468	407	446	20			
1	P	301	Total	C	N	O	S	0	0	0
			2341	1468	407	446	20			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
G	79	GLY	-	expression tag	UNP P35961
G	80	ALA	-	expression tag	UNP P35961
G	81	ARG	-	expression tag	UNP P35961
G	82	SER	-	expression tag	UNP P35961
G	128	GLY	-	linker	UNP P35961
G	129	ALA	-	linker	UNP P35961
G	194	GLY	-	linker	UNP P35961
G	298	GLY	-	linker	UNP P35961
G	299	ALA	-	linker	UNP P35961
G	329	GLY	-	linker	UNP P35961
P	1079	GLY	-	expression tag	UNP P35961
P	1080	ALA	-	expression tag	UNP P35961
P	1081	ARG	-	expression tag	UNP P35961
P	1082	SER	-	expression tag	UNP P35961
P	1128	GLY	-	linker	UNP P35961
P	1129	ALA	-	linker	UNP P35961
P	1194	GLY	-	linker	UNP P35961
P	1298	GLY	-	linker	UNP P35961
P	1299	ALA	-	linker	UNP P35961
P	1329	GLY	-	linker	UNP P35961

- Molecule 2 is a protein called antibody 17b light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	L	214	Total	C	N	O	S	0	0	0
			1647	1028	282	332	5			
2	Q	214	Total	C	N	O	S	0	0	0
			1647	1028	282	332	5			

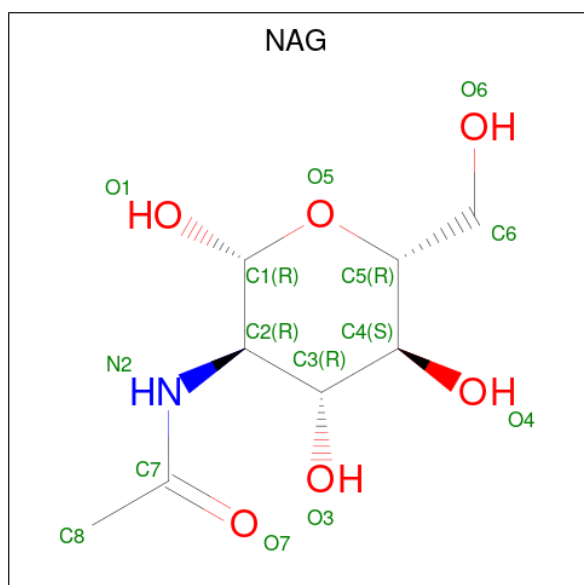
- Molecule 3 is a protein called antibody 17b heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	H	229	Total	C	N	O	S	0	0	0
			1722	1086	289	342	5			
3	R	229	Total	C	N	O	S	0	0	0
			1722	1086	289	342	5			

- Molecule 4 is a protein called CD4M33, scorpion-toxin mimic of CD4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	M	27	Total	C	N	O	S	0	0	0
			201	128	36	31	6			
4	S	27	Total	C	N	O	S	0	0	0
			201	128	36	31	6			

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



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

Continued on next page...

*Continued from previous page...*

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

- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	G	66	Total O 66 66	0	0
6	L	44	Total O 44 44	0	0
6	H	76	Total O 76 76	0	0
6	M	6	Total O 6 6	0	0

*Continued on next page...*

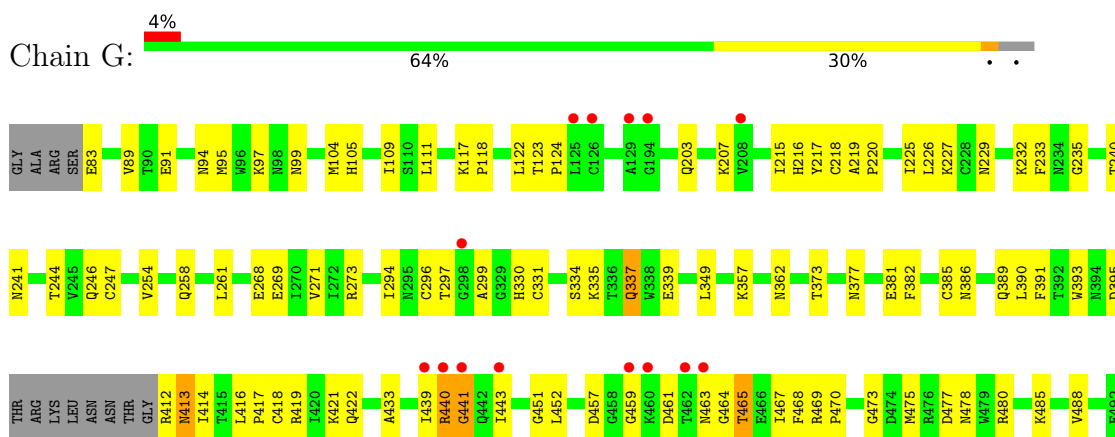
*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	P	45	Total 45	O 45	0	0
6	Q	35	Total 35	O 35	0	0
6	R	47	Total 47	O 47	0	0
6	S	5	Total 5	O 5	0	0

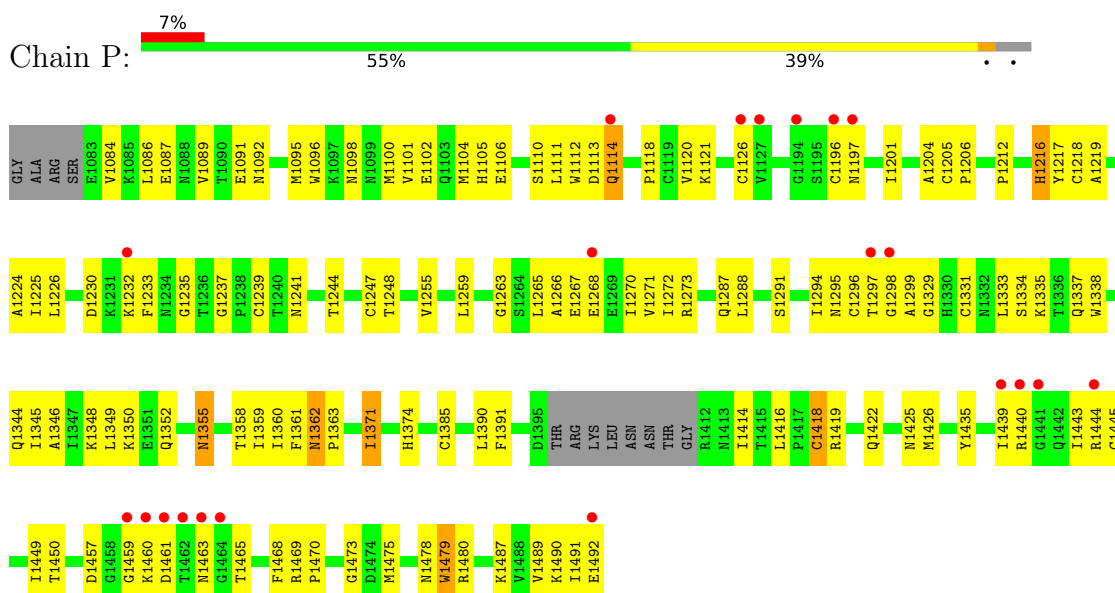
### 3 Residue-property plots [i](#)

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: Exterior membrane glycoprotein(GP120),Exterior membrane glycoprotein(GP120), Exterior membrane glycoprotein(GP120)



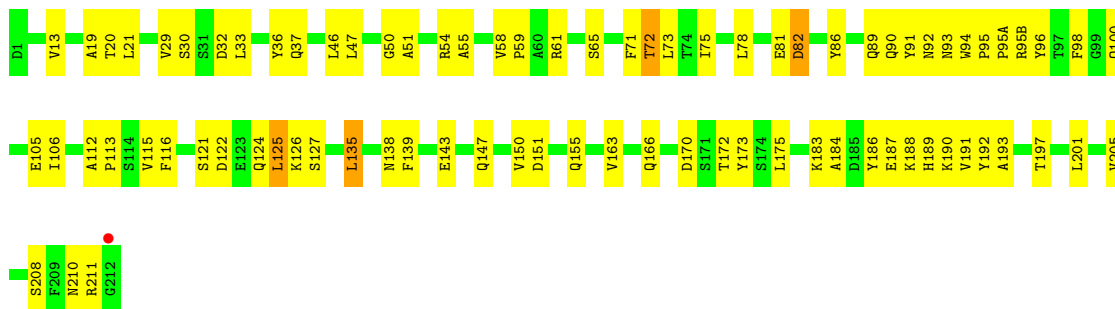
- Molecule 1: Exterior membrane glycoprotein(GP120),Exterior membrane glycoprotein(GP120), Exterior membrane glycoprotein(GP120)



- Molecule 2: antibody 17b light chain

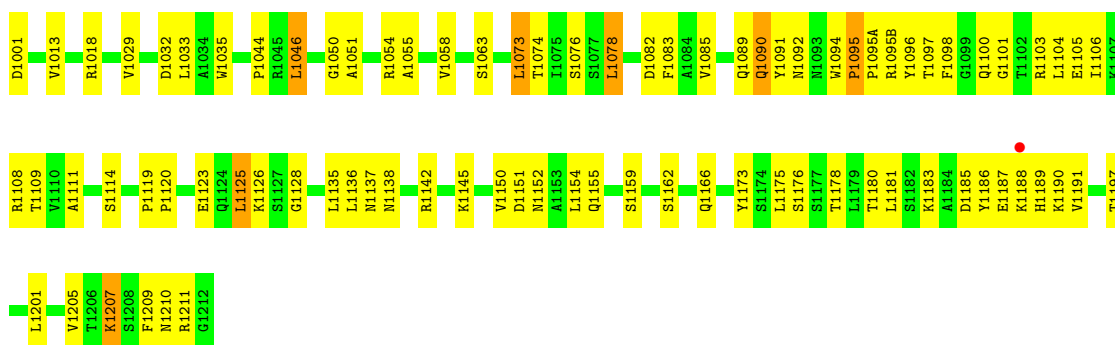






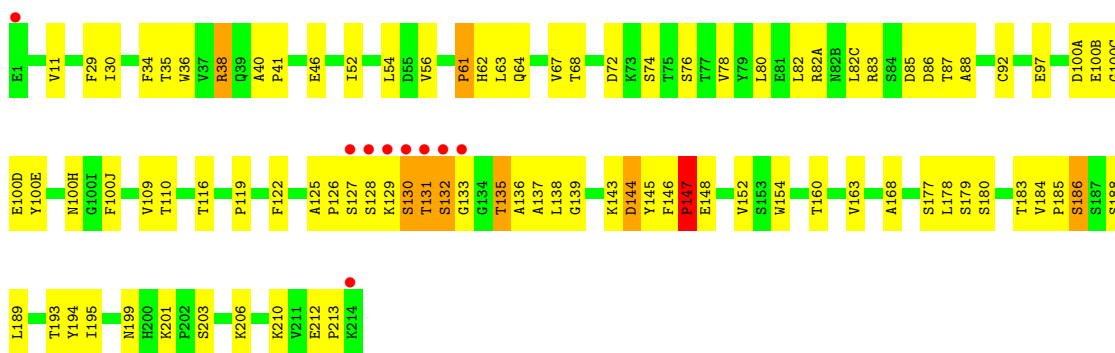
- Molecule 2: antibody 17b light chain

Chain Q: 61% 36%



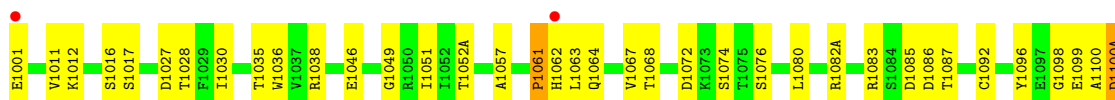
- Molecule 3: antibody 17b heavy chain

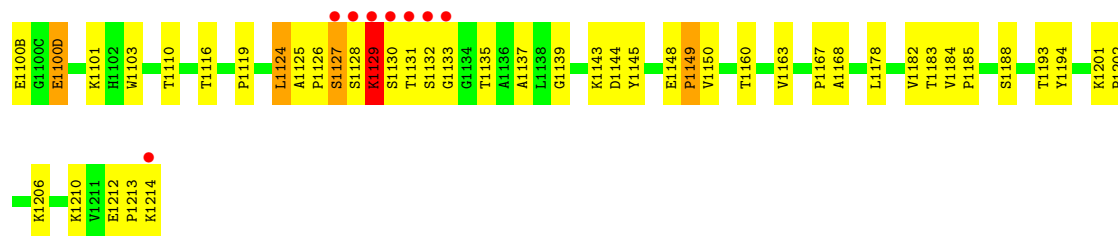
Chain H: 4% 60% 36%



- Molecule 3: antibody 17b heavy chain

Chain R: 4% 64% 33%





- Molecule 4: CD4M33, scorpion-toxin mimic of CD4

Chain M:    63% 26% 11%



- Molecule 4: CD4M33, scorpion-toxin mimic of CD4

Chain S:    63% 26% 11%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	51.52Å 157.84Å 109.92Å 90.00° 93.55° 90.00°	Depositor
Resolution (Å)	19.94 – 2.75 19.93 – 2.74	Depositor EDS
% Data completeness (in resolution range)	88.0 (19.94-2.75) 87.4 (19.93-2.74)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.89 (at 2.75Å)	Xtrriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.202 , 0.275 0.191 , 0.266	Depositor DCC
$R_{free}$ test set	4201 reflections (10.09%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	27.0	Xtrriage
Anisotropy	0.641	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	12370	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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: BIF, NAG, VLM, MPT

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	0.52	0/2386	0.64	1/3234 (0.0%)
1	P	0.54	0/2386	0.62	0/3234
2	L	0.53	0/1684	0.65	0/2288
2	Q	0.56	0/1684	0.66	0/2288
3	H	0.54	0/1762	0.67	0/2399
3	R	0.53	0/1762	0.65	0/2399
4	M	0.54	0/171	0.75	0/224
4	S	0.52	0/171	0.66	0/224
All	All	0.53	0/12006	0.65	1/16290 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	G	441	GLY	N-CA-C	-5.43	99.52	113.10

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	2341	0	2295	84	0
1	P	2341	0	2293	112	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	L	1647	0	1593	69	0
2	Q	1647	0	1590	77	0
3	H	1722	0	1691	75	0
3	R	1722	0	1688	67	0
4	M	201	0	198	7	0
4	S	201	0	198	7	0
5	G	112	0	104	5	0
5	P	112	0	104	6	0
6	G	66	0	0	4	0
6	H	76	0	0	0	0
6	L	44	0	0	1	0
6	M	6	0	0	0	0
6	P	45	0	0	2	0
6	Q	35	0	0	2	0
6	R	47	0	0	0	0
6	S	5	0	0	1	0
All	All	12370	0	11754	478	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 20.

The worst 5 of 478 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)
3:H:185:PRO:HG2	3:H:188:SER:HB2	1.30	1.13
1:G:439:ILE:HG13	1:G:440:ARG:H	1.14	1.06
3:H:126:PRO:HA	3:H:129:LYS:HE2	1.36	1.03
3:R:1212:GLU:HG2	3:R:1213:PRO:HD2	1.47	0.95
1:G:389:GLN:HG3	5:G:886:NAG:H81	1.50	0.94

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	297/313 (95%)	271 (91%)	24 (8%)	2 (1%)	22	39
1	P	297/313 (95%)	262 (88%)	31 (10%)	4 (1%)	12	21
2	L	212/214 (99%)	193 (91%)	17 (8%)	2 (1%)	17	31
2	Q	212/214 (99%)	193 (91%)	17 (8%)	2 (1%)	17	31
3	H	227/229 (99%)	204 (90%)	13 (6%)	10 (4%)	2	2
3	R	227/229 (99%)	204 (90%)	15 (7%)	8 (4%)	3	5
4	M	24/27 (89%)	23 (96%)	0	1 (4%)	3	3
4	S	24/27 (89%)	22 (92%)	2 (8%)	0	100	100
All	All	1520/1566 (97%)	1372 (90%)	119 (8%)	29 (2%)	8	14

5 of 29 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	G	268	GLU
1	G	440	ARG
3	H	64	GLN
3	H	131	THR
1	P	1460	LYS

### 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	266/276 (96%)	259 (97%)	7 (3%)	46	66
1	P	266/276 (96%)	251 (94%)	15 (6%)	21	36
2	L	184/184 (100%)	176 (96%)	8 (4%)	29	48
2	Q	184/184 (100%)	176 (96%)	8 (4%)	29	48
3	H	193/193 (100%)	184 (95%)	9 (5%)	26	45
3	R	193/193 (100%)	185 (96%)	8 (4%)	30	50

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
4	M	19/19 (100%)	17 (90%)	2 (10%)	7 11
4	S	19/19 (100%)	16 (84%)	3 (16%)	2 3
All	All	1324/1344 (98%)	1264 (96%)	60 (4%)	27 46

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	P	1111	LEU
3	R	1206	LYS
1	P	1267	GLU
3	R	1149	PRO
4	S	1018	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 34 such sidechains are listed below:

Mol	Chain	Res	Type
2	Q	1090	GLN
2	Q	1124	GLN
2	Q	1210	ASN
3	H	64	GLN
2	L	166	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](#)

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
4	BIF	M	23	4	16,18,19	1.47	4 (25%)	16,23,25	1.36	3 (18%)
4	BIF	S	1023	4	16,18,19	1.46	4 (25%)	16,23,25	1.29	3 (18%)
4	VLM	M	27	4	6,7,7	1.57	2 (33%)	7,9,9	0.58	0
4	VLM	S	1027	4	6,7,7	1.32	1 (16%)	7,9,9	0.73	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
4	BIF	M	23	4	-	2/5/10/12	0/2/2/2
4	BIF	S	1023	4	-	0/5/10/12	0/2/2/2
4	VLM	M	27	4	-	0/7/8/8	-
4	VLM	S	1027	4	-	0/7/8/8	-

The worst 5 of 11 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	M	27	VLM	C-NT	-2.96	1.25	1.32
4	S	1027	VLM	C-NT	-2.82	1.25	1.32
4	M	23	BIF	CE1-CD1	2.59	1.42	1.36
4	S	1023	BIF	CE1-CD1	2.56	1.42	1.36
4	M	23	BIF	C10-C11	2.18	1.41	1.36

The worst 5 of 6 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	M	23	BIF	CE2-CZ-CE1	-2.83	116.14	118.65
4	S	1023	BIF	C13-C12-C11	-2.76	116.20	118.65
4	M	23	BIF	CG-CB-CA	2.68	119.53	114.10
4	S	1023	BIF	CE2-CZ-CE1	-2.65	116.30	118.65
4	M	23	BIF	C13-C12-C11	-2.64	116.30	118.65

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	M	23	BIF	CA-CB-CG-CD2
4	M	23	BIF	CA-CB-CG-CD1



There are no ring outliers.

4 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	M	23	BIF	4	0
4	S	1023	BIF	1	0
4	M	27	VLM	1	0
4	S	1027	VLM	2	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

16 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	776	1	14,14,15	0.45	0	17,19,21	0.86	1 (5%)
5	NAG	P	1789	1	14,14,15	0.57	0	17,19,21	0.77	1 (5%)
5	NAG	P	1795	1	14,14,15	0.52	0	17,19,21	0.71	1 (5%)
5	NAG	P	1886	1	14,14,15	0.55	0	17,19,21	0.83	1 (5%)
5	NAG	G	795	1	14,14,15	0.55	0	17,19,21	0.68	1 (5%)
5	NAG	G	762	1	14,14,15	0.54	0	17,19,21	0.70	0
5	NAG	P	1588	1	14,14,15	0.52	0	17,19,21	0.67	1 (5%)
5	NAG	P	1762	1	14,14,15	0.58	0	17,19,21	0.79	0
5	NAG	G	588	1	14,14,15	0.63	0	17,19,21	0.63	0
5	NAG	G	886	1	14,14,15	0.63	0	17,19,21	0.85	1 (5%)
5	NAG	G	734	1	14,14,15	0.55	0	17,19,21	0.56	0
5	NAG	G	741	1	14,14,15	0.51	0	17,19,21	0.81	0
5	NAG	P	1734	1	14,14,15	0.56	0	17,19,21	0.59	0
5	NAG	P	1776	1	14,14,15	0.63	0	17,19,21	0.70	1 (5%)
5	NAG	G	789	1	14,14,15	0.49	0	17,19,21	0.72	1 (5%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	NAG	P	1741	1	14,14,15	0.67	0	17,19,21	0.65	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	776	1	-	4/6/23/26	0/1/1/1
5	NAG	P	1789	1	-	4/6/23/26	0/1/1/1
5	NAG	P	1795	1	-	4/6/23/26	0/1/1/1
5	NAG	P	1886	1	-	2/6/23/26	0/1/1/1
5	NAG	G	795	1	-	4/6/23/26	0/1/1/1
5	NAG	G	762	1	-	4/6/23/26	0/1/1/1
5	NAG	P	1588	1	-	2/6/23/26	0/1/1/1
5	NAG	P	1762	1	-	2/6/23/26	0/1/1/1
5	NAG	G	588	1	-	2/6/23/26	0/1/1/1
5	NAG	G	886	1	-	4/6/23/26	0/1/1/1
5	NAG	G	734	1	1/1/5/7	5/6/23/26	0/1/1/1
5	NAG	G	741	1	-	2/6/23/26	0/1/1/1
5	NAG	P	1734	1	-	5/6/23/26	0/1/1/1
5	NAG	P	1776	1	-	2/6/23/26	0/1/1/1
5	NAG	G	789	1	-	4/6/23/26	0/1/1/1
5	NAG	P	1741	1	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	G	776	NAG	C2-N2-C7	-2.66	119.12	122.90
5	P	1886	NAG	C2-N2-C7	-2.35	119.56	122.90
5	P	1789	NAG	C2-N2-C7	-2.31	119.61	122.90
5	P	1776	NAG	C2-N2-C7	-2.26	119.68	122.90
5	G	795	NAG	C2-N2-C7	-2.14	119.85	122.90

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
5	G	734	NAG	C1

5 of 54 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	G	588	NAG	C8-C7-N2-C2
5	G	588	NAG	O7-C7-N2-C2
5	G	734	NAG	C3-C2-N2-C7
5	G	734	NAG	C8-C7-N2-C2
5	G	734	NAG	O7-C7-N2-C2

There are no ring outliers.

8 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	P	1795	NAG	2	0
5	G	762	NAG	1	0
5	P	1588	NAG	1	0
5	P	1762	NAG	2	0
5	G	886	NAG	2	0
5	G	741	NAG	1	0
5	G	789	NAG	1	0
5	P	1741	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	301/313 (96%)	-0.01	14 (4%) 31 37	7, 23, 50, 69	0
1	P	301/313 (96%)	0.20	21 (6%) 16 19	14, 33, 56, 84	0
2	L	214/214 (100%)	-0.27	1 (0%) 91 94	6, 20, 33, 41	0
2	Q	214/214 (100%)	0.00	1 (0%) 91 94	11, 27, 50, 56	0
3	H	229/229 (100%)	-0.13	9 (3%) 39 46	5, 18, 48, 80	0
3	R	229/229 (100%)	-0.07	10 (4%) 34 41	12, 24, 49, 83	0
4	M	24/27 (88%)	-0.02	0 100 100	17, 29, 39, 44	0
4	S	24/27 (88%)	0.12	0 100 100	21, 30, 42, 47	0
All	All	1536/1566 (98%)	-0.03	56 (3%) 42 51	5, 25, 51, 84	0

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	R	1130	SER	8.0
3	H	131	THR	6.9
1	P	1459	GLY	6.4
3	R	1132	SER	6.3
1	P	1462	THR	6.1

### 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
4	VLM	M	27	8/8	0.91	0.18	24,25,27,28	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	BIF	S	1023	17/18	0.93	0.18	14,17,28,30	0
4	VLM	S	1027	8/8	0.93	0.19	29,30,31,31	0
4	BIF	M	23	17/18	0.96	0.15	10,12,17,17	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( $\text{\AA}^2$ )	Q<0.9
5	NAG	P	1734	14/15	0.65	0.40	61,63,66,66	0
5	NAG	G	741	14/15	0.71	0.38	58,60,63,64	0
5	NAG	P	1741	14/15	0.76	0.46	60,64,67,68	0
5	NAG	P	1789	14/15	0.79	0.24	57,59,62,63	0
5	NAG	P	1776	14/15	0.81	0.34	35,41,42,43	0
5	NAG	G	734	14/15	0.82	0.36	56,58,60,61	0
5	NAG	G	886	14/15	0.82	0.26	33,39,47,48	0
5	NAG	P	1588	14/15	0.85	0.22	48,49,50,50	0
5	NAG	P	1886	14/15	0.85	0.25	47,50,52,54	0
5	NAG	P	1762	14/15	0.89	0.19	24,28,32,32	0
5	NAG	G	776	14/15	0.89	0.31	36,38,41,41	0
5	NAG	G	789	14/15	0.89	0.19	26,30,36,38	0
5	NAG	P	1795	14/15	0.89	0.29	45,48,52,53	0
5	NAG	G	588	14/15	0.89	0.25	47,49,50,51	0
5	NAG	G	795	14/15	0.92	0.20	37,39,44,46	0
5	NAG	G	762	14/15	0.94	0.13	15,16,16,18	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.