



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 09:57 AM GMT

PDB ID : 3KAS  
Title : Machupo virus GP1 bound to human transferrin receptor 1  
Authors : Abraham, J.; Corbett, K.D.; Harrison, S.C.  
Deposited on : 2009-10-19  
Resolution : 2.40 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 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
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) : trunk26865

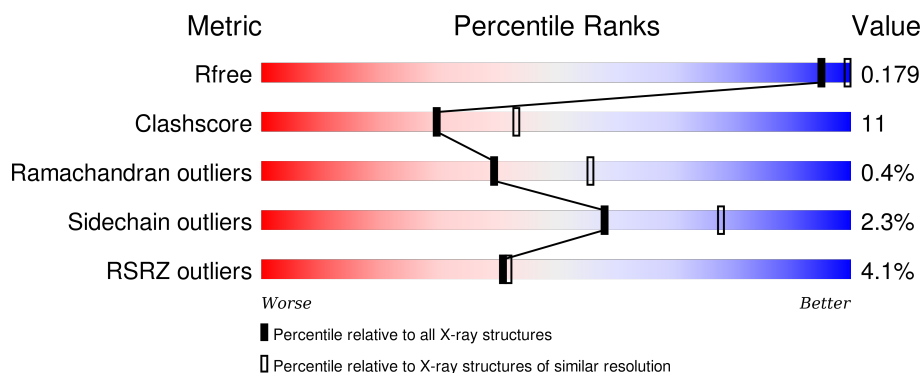
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.40 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))
$R_{free}$	91344	2919 (2.40-2.40)
Clashscore	102246	3407 (2.40-2.40)
Ramachandran outliers	100387	3351 (2.40-2.40)
Sidechain outliers	100360	3352 (2.40-2.40)
RSRZ outliers	91569	2928 (2.40-2.40)

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	A	640	<div> <div>4%</div> <div>76%</div> <div>22%</div> <div>..</div> </div>
2	B	162	<div> <div>3%</div> <div>72%</div> <div>23%</div> <div>..</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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	NAG	B	1001	-	-	-	X
4	PO4	A	4	-	-	-	X
4	PO4	B	1	-	-	-	X

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 6717 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 Transferrin receptor protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	633	Total	C	N	O	S	0	0	0
			5004	3210	843	937	14			

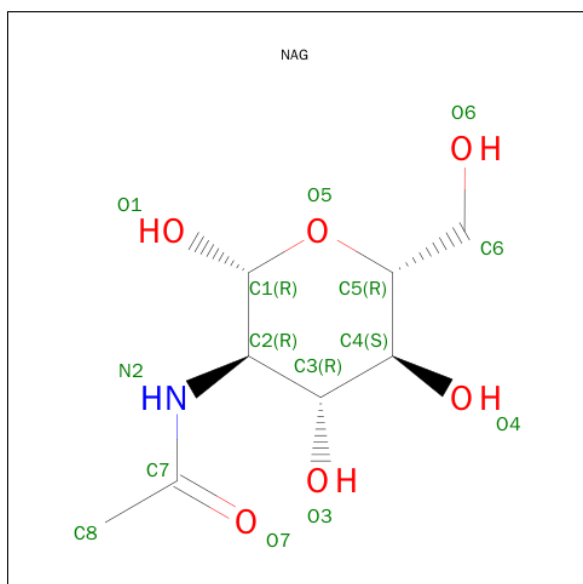
There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	142	SER	GLY	VARIANT	UNP P02786

- Molecule 2 is a protein called Glycoprotein.

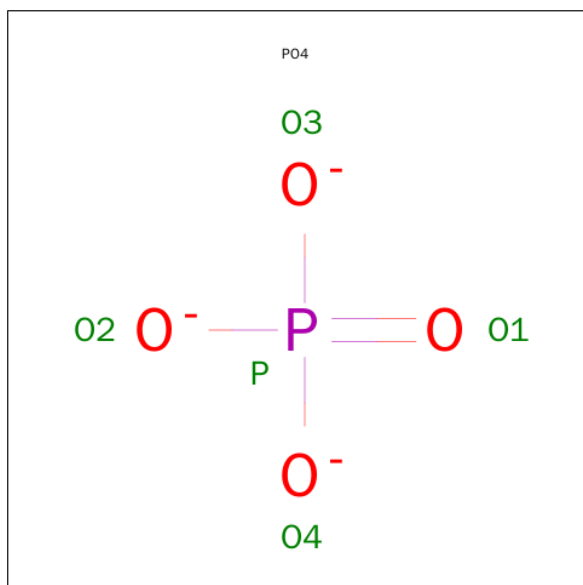
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	156	Total	C	N	O	S	0	0	0
			1255	789	218	234	14			

- Molecule 3 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		
3	B	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	O	P	0	0
			5	4	1		
4	A	1	Total	O	P	0	0
			5	4	1		
4	A	1	Total	O	P	0	0
			5	4	1		
4	A	1	Total	O	P	0	0
			5	4	1		
4	B	1	Total	O	P	0	0
			5	4	1		
4	B	1	Total	O	P	0	0
			5	4	1		

- Molecule 5 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total K 1 1	0	0

- Molecule 6 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	3	Total C N O 38 22 2 14	0	0

- Molecule 7 is a polymer of unknown type called SUGAR (7-MER).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	7	Total C N O 83 46 2 35	0	0

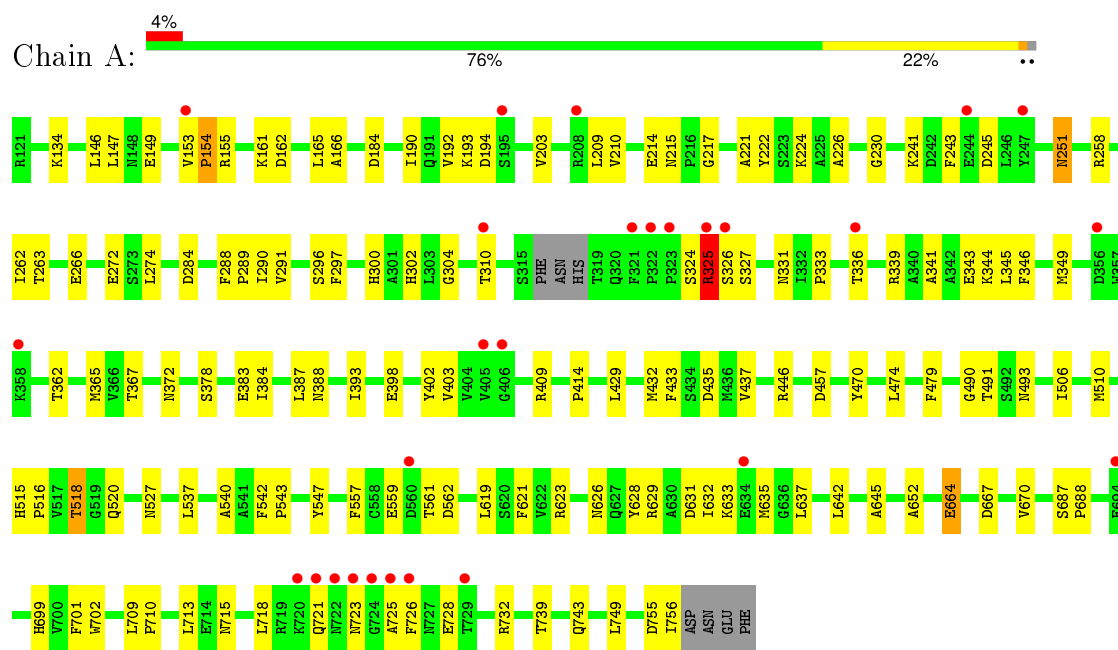
- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	200	Total O 200 200	0	0
8	B	50	Total O 50 50	0	0

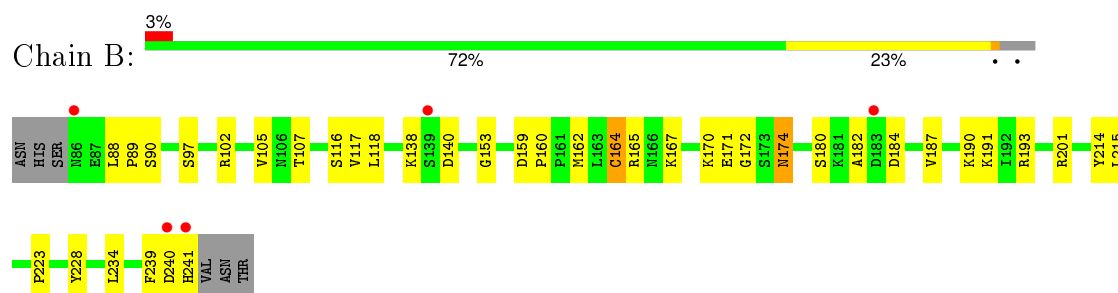
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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 ( $\text{RSRZ} > 2$ ). 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: Transferrin receptor protein 1



#### • Molecule 2: Glycoprotein



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	143.95Å 170.23Å 98.47Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	33.13 – 2.40 33.13 – 2.39	Depositor EDS
% Data completeness (in resolution range)	88.7 (33.13-2.40) 87.5 (33.13-2.39)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.80 (at 2.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.5_2)	Depositor
R, $R_{free}$	0.189 , 0.238 0.182 , 0.179	Depositor DCC
$R_{free}$ test set	2138 reflections (5.07%)	DCC
Wilson B-factor (Å <sup>2</sup> )	48.2	Xtriage
Anisotropy	0.008	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 52.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	1 of 42298 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	6717	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.83% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: BMA, NAG, K, FUC, PO4, MAN

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	A	0.35	0/5121	0.51	0/6942
2	B	0.38	0/1287	0.53	0/1735
All	All	0.36	0/6408	0.51	0/8677

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

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	A	5004	0	4956	114	0
2	B	1255	0	1181	32	0
3	A	28	0	26	3	0
3	B	28	0	26	1	0
4	A	20	0	0	0	0
4	B	10	0	0	1	0
5	A	1	0	0	0	0
6	B	38	0	34	0	0
7	B	83	0	70	2	0
8	A	200	0	0	8	0
8	B	50	0	0	2	0
All	All	6717	0	6293	143	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 11.

All (143) 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:A:230:GLY:O	1:A:372:ASN:HB2	1.81	0.80
1:A:154:PRO:HD2	1:A:161:LYS:HD2	1.63	0.80
1:A:518:THR:HG22	1:A:520:GLN:H	1.48	0.78
1:A:626:ASN:O	1:A:629:ARG:HG3	1.89	0.71
1:A:149:GLU:C	1:A:153:VAL:HG23	2.13	0.69
1:A:518:THR:CG2	1:A:520:GLN:H	2.06	0.68
1:A:153:VAL:HB	1:A:154:PRO:HD3	1.76	0.67
1:A:155:ARG:HD2	1:A:409:ARG:O	1.95	0.66
1:A:272:GLU:OE1	1:A:331:ASN:HB2	1.96	0.66
2:B:165:ARG:HD2	2:B:172:GLY:O	1.97	0.64
1:A:297:PHE:O	1:A:336:THR:HG21	1.97	0.64
1:A:664:GLU:CD	1:A:664:GLU:H	2.02	0.64
1:A:149:GLU:O	1:A:153:VAL:HG23	1.98	0.64
1:A:721:GLN:NE2	1:A:725:ALA:HB3	2.14	0.62
1:A:346:PHE:HD1	1:A:362:THR:HG22	1.65	0.61
2:B:187:VAL:HG12	8:B:270:HOH:O	2.00	0.61
1:A:345:LEU:O	1:A:349:MET:HG3	2.00	0.61
1:A:190:ILE:HD13	1:A:384:ILE:HD13	1.84	0.60
1:A:134:LYS:NZ	1:A:435:ASP:HB3	2.17	0.60
1:A:209:LEU:HD21	2:B:97:SER:HB3	1.84	0.60
1:A:297:PHE:CE2	1:A:336:THR:OG1	2.56	0.59
2:B:97:SER:OG	3:B:1001:NAG:H82	2.03	0.59
1:A:387:LEU:O	1:A:409:ARG:NH2	2.36	0.59
1:A:193:LYS:HE2	1:A:217:GLY:O	2.03	0.59
1:A:296:SER:HB2	8:A:766:HOH:O	2.02	0.58
1:A:300:HIS:HD2	1:A:302:HIS:H	1.52	0.58
1:A:203:VAL:HG22	1:A:210:VAL:HG12	1.86	0.58
2:B:153:GLY:O	2:B:191:LYS:HE2	2.04	0.57
2:B:102:ARG:NH2	2:B:239:PHE:CD1	2.72	0.57
1:A:290:ILE:O	1:A:339:ARG:NH2	2.37	0.57
1:A:506:ILE:O	1:A:510:MET:HG3	2.04	0.56
2:B:90:SER:HB2	2:B:240:ASP:O	2.05	0.56
1:A:230:GLY:C	1:A:372:ASN:HB2	2.26	0.56
1:A:540:ALA:O	1:A:543:PRO:HD2	2.06	0.56
1:A:251:ASN:C	1:A:251:ASN:OD1	2.45	0.55
1:A:258:ARG:HD2	1:A:284:ASP:OD1	2.07	0.55
1:A:718:LEU:O	1:A:726:PHE:HB2	2.08	0.54

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:721:GLN:NE2	1:A:723:ASN:HB3	2.23	0.54
1:A:637:LEU:HD11	1:A:732:ARG:HG2	1.90	0.53
1:A:221:ALA:O	1:A:222:TYR:HB2	2.08	0.52
1:A:241:LYS:HE2	1:A:245:ASP:OD2	2.09	0.52
1:A:184:ASP:HB3	1:A:388:ASN:HB2	1.92	0.52
1:A:224:LYS:HD3	1:A:331:ASN:O	2.10	0.52
1:A:209:LEU:N	1:A:209:LEU:HD12	2.25	0.51
2:B:240:ASP:O	2:B:241:HIS:HB2	2.10	0.51
1:A:154:PRO:HD2	1:A:161:LYS:CD	2.36	0.51
1:A:243:PHE:HB3	1:A:274:LEU:HD11	1.91	0.51
1:A:153:VAL:CB	1:A:154:PRO:HD3	2.40	0.51
1:A:709:LEU:HB2	1:A:710:PRO:HD3	1.93	0.51
1:A:642:LEU:HA	1:A:739:THR:HG23	1.92	0.51
3:A:1011:NAG:H3	8:A:802:HOH:O	2.09	0.51
1:A:162:ASP:OD2	1:A:409:ARG:NH1	2.44	0.50
1:A:251:ASN:O	1:A:251:ASN:OD1	2.28	0.50
2:B:116:SER:HB3	2:B:214:TYR:CE1	2.46	0.50
1:A:470:TYR:HB2	1:A:474:LEU:HB2	1.93	0.50
2:B:171:GLU:O	2:B:223:PRO:HB2	2.11	0.50
1:A:226:ALA:HB3	1:A:333:PRO:HG3	1.93	0.50
1:A:365:MET:HE3	8:A:807:HOH:O	2.11	0.50
2:B:164:CYS:HA	2:B:174:ASN:O	2.12	0.50
1:A:343:GLU:CG	1:A:362:THR:HG21	2.42	0.49
1:A:304:GLY:O	1:A:537:LEU:HD21	2.12	0.49
1:A:493:ASN:O	1:A:557:PHE:HA	2.12	0.49
1:A:635:MET:CE	1:A:728:GLU:HG3	2.42	0.49
1:A:543:PRO:O	1:A:547:TYR:HB3	2.13	0.49
1:A:699:HIS:CE1	1:A:701:PHE:HB2	2.47	0.49
1:A:344:LYS:HD3	2:B:117:VAL:O	2.11	0.49
1:A:403:VAL:HG22	1:A:479:PHE:CE1	2.48	0.48
1:A:251:ASN:OD1	3:A:1001:NAG:N2	2.46	0.48
1:A:667:ASP:HB3	1:A:670:VAL:HG12	1.93	0.48
1:A:192:VAL:HG12	1:A:193:LYS:N	2.28	0.48
1:A:343:GLU:CD	1:A:362:THR:HG21	2.34	0.48
2:B:153:GLY:O	2:B:191:LYS:CE	2.62	0.48
1:A:561:THR:HG22	1:A:562:ASP:O	2.14	0.48
1:A:346:PHE:CD1	1:A:362:THR:HG22	2.46	0.47
1:A:537:LEU:HD23	8:A:867:HOH:O	2.14	0.47
1:A:153:VAL:O	1:A:414:PRO:HA	2.14	0.47
1:A:619:LEU:HD21	1:A:623:ARG:HH21	1.80	0.47
1:A:310:THR:HG22	1:A:310:THR:O	2.14	0.47

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:429:LEU:HD23	1:A:432:MET:HE3	1.96	0.47
1:A:146:LEU:O	1:A:146:LEU:HD23	2.15	0.47
1:A:542:PHE:HB3	1:A:543:PRO:HD3	1.96	0.47
2:B:165:ARG:NH1	2:B:171:GLU:HB3	2.30	0.46
1:A:134:LYS:HZ2	1:A:435:ASP:HB3	1.80	0.46
1:A:147:LEU:CD2	1:A:165:LEU:HD11	2.45	0.46
2:B:215:LEU:C	2:B:215:LEU:HD23	2.37	0.46
1:A:715:ASN:O	1:A:726:PHE:HD1	1.99	0.45
1:A:628:TYR:O	1:A:632:ILE:HG12	2.16	0.45
1:A:243:PHE:HB3	1:A:274:LEU:CD1	2.45	0.45
1:A:194:ASP:HB2	1:A:378:SER:O	2.15	0.45
1:A:288:PHE:O	1:A:290:ILE:N	2.45	0.45
1:A:667:ASP:HB3	1:A:670:VAL:CG1	2.47	0.45
1:A:652:ALA:HB1	1:A:749:LEU:HB3	1.99	0.45
1:A:629:ARG:O	1:A:633:LYS:HG2	2.16	0.45
1:A:490:GLY:HA3	1:A:559:GLU:OE1	2.17	0.45
2:B:182:ALA:CB	7:B:1032:NAG:H82	2.46	0.45
1:A:518:THR:HG23	1:A:520:GLN:CG	2.47	0.44
1:A:296:SER:HB3	8:A:763:HOH:O	2.17	0.44
2:B:184:ASP:HB2	8:B:275:HOH:O	2.18	0.44
1:A:632:ILE:HG22	1:A:637:LEU:O	2.16	0.44
2:B:165:ARG:CD	2:B:171:GLU:HB2	2.48	0.44
1:A:324:SER:HB3	1:A:327:SER:HB3	1.99	0.44
1:A:291:VAL:HG23	8:A:71:HOH:O	2.18	0.44
1:A:134:LYS:HZ1	1:A:435:ASP:HB3	1.83	0.44
1:A:403:VAL:HG22	1:A:479:PHE:CZ	2.53	0.44
1:A:325:ARG:HB3	1:A:326:SER:H	1.47	0.44
1:A:251:ASN:OD1	3:A:1001:NAG:C7	2.66	0.43
1:A:635:MET:HE3	1:A:728:GLU:HG3	2.00	0.43
1:A:645:ALA:HB2	1:A:743:GLN:HG2	2.00	0.43
1:A:621:PHE:CZ	1:A:709:LEU:HD22	2.53	0.43
1:A:721:GLN:O	1:A:721:GLN:HG3	2.16	0.43
1:A:304:GLY:HA2	1:A:537:LEU:HD11	2.01	0.43
1:A:166:ALA:HB2	1:A:409:ARG:NH1	2.33	0.43
2:B:193:ARG:NE	4:B:5:PO4:O3	2.38	0.43
1:A:192:VAL:HG12	1:A:193:LYS:H	1.84	0.42
2:B:159:ASP:HB3	2:B:160:PRO:HD2	2.00	0.42
1:A:210:VAL:HG22	2:B:228:TYR:OH	2.19	0.42
1:A:699:HIS:CE1	1:A:702:TRP:CD1	3.07	0.42
2:B:138:LYS:HB3	2:B:140:ASP:OD1	2.19	0.42
1:A:214:GLU:OE2	1:A:341:ALA:HB2	2.20	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:262:ILE:HB	1:A:266:GLU:OE1	2.19	0.42
2:B:190:LYS:HE3	2:B:190:LYS:HB3	1.83	0.42
2:B:107:THR:HG21	2:B:234:LEU:HD21	2.01	0.42
2:B:105:VAL:HG22	2:B:105:VAL:O	2.19	0.42
2:B:182:ALA:HB2	7:B:1032:NAG:H82	2.00	0.42
1:A:755:ASP:O	1:A:756:ILE:HG13	2.20	0.42
1:A:215:ASN:O	2:B:170:LYS:HE3	2.20	0.42
1:A:515:HIS:HA	1:A:516:PRO:HD3	1.82	0.41
1:A:263:THR:HB	8:A:767:HOH:O	2.18	0.41
2:B:88:LEU:HB3	2:B:89:PRO:CD	2.51	0.41
1:A:146:LEU:C	1:A:146:LEU:HD23	2.41	0.41
1:A:713:LEU:HD23	1:A:713:LEU:HA	1.76	0.41
1:A:433:PHE:O	1:A:437:VAL:HG23	2.19	0.41
2:B:118:LEU:HD11	2:B:174:ASN:HD22	1.85	0.41
1:A:258:ARG:HH11	1:A:284:ASP:CG	2.24	0.41
1:A:559:GLU:C	1:A:561:THR:H	2.24	0.41
2:B:167:LYS:HG2	2:B:167:LYS:O	2.20	0.41
1:A:297:PHE:CD2	1:A:336:THR:OG1	2.74	0.41
1:A:687:SER:HA	1:A:688:PRO:HD3	1.72	0.41
1:A:339:ARG:O	1:A:343:GLU:HG3	2.21	0.40
2:B:180:SER:C	2:B:182:ALA:H	2.25	0.40
1:A:393:ILE:HG21	1:A:437:VAL:HG21	2.03	0.40
2:B:201:ARG:HE	2:B:201:ARG:HB2	1.74	0.40
1:A:362:THR:HG23	8:A:86:HOH:O	2.20	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	A	629/640 (98%)	604 (96%)	22 (4%)	3 (0%)	34 48

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	154/162 (95%)	143 (93%)	11 (7%)	0	100	100
All	All	783/802 (98%)	747 (95%)	33 (4%)	3 (0%)	39	56

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	325	ARG
1	A	154	PRO
1	A	289	PRO

### 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	A	542/549 (99%)	529 (98%)	13 (2%)	57	76
2	B	141/147 (96%)	138 (98%)	3 (2%)	61	80
All	All	683/696 (98%)	667 (98%)	16 (2%)	58	78

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	251	ASN
1	A	325	ARG
1	A	367	THR
1	A	383	GLU
1	A	398	GLU
1	A	402	TYR
1	A	446	ARG
1	A	457	ASP
1	A	491	THR
1	A	518	THR
1	A	527	ASN
1	A	631	ASP
1	A	664	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	B	162	MET
2	B	164	CYS
2	B	174	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	215	ASN
1	A	300	HIS
1	A	626	ASN
1	A	721	GLN
2	B	106	ASN
2	B	174	ASN
2	B	241	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

10 carbohydrates 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	$\# Z  > 2$
6	NAG	B	1021	2,6	14,14,15	0.56	0	15,19,21	0.76	1 (6%)
6	NAG	B	1022	6	14,14,15	0.51	0	15,19,21	0.58	0
6	FUC	B	1023	6	10,10,11	0.53	0	14,14,16	1.82	5 (35%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
7	NAG	B	1031	2,7	14,14,15	0.47	0	15,19,21	0.82	0
7	NAG	B	1032	7	14,14,15	0.52	0	15,19,21	0.85	1 (6%)
7	BMA	B	1033	7	11,11,12	0.67	0	14,15,17	1.58	2 (14%)
7	MAN	B	1034	7	11,11,12	0.63	0	14,15,17	0.76	0
7	MAN	B	1035	7	11,11,12	0.57	0	14,15,17	1.64	1 (7%)
7	MAN	B	1036	7	11,11,12	0.66	0	14,15,17	2.16	3 (21%)
7	MAN	B	1037	7	11,11,12	0.59	0	14,15,17	1.53	2 (14%)

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
6	NAG	B	1021	2,6	-	0/6/23/26	0/1/1/1
6	NAG	B	1022	6	-	0/6/23/26	0/1/1/1
6	FUC	B	1023	6	-	0/0/17/20	0/1/1/1
7	NAG	B	1031	2,7	-	0/6/23/26	0/1/1/1
7	NAG	B	1032	7	-	0/6/23/26	0/1/1/1
7	BMA	B	1033	7	-	0/2/19/22	0/1/1/1
7	MAN	B	1034	7	-	0/2/19/22	0/1/1/1
7	MAN	B	1035	7	-	0/2/19/22	0/1/1/1
7	MAN	B	1036	7	-	0/2/19/22	0/1/1/1
7	MAN	B	1037	7	-	0/2/19/22	1/1/1/1

There are no bond length outliers.

All (15) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1033	BMA	C1-O5-C5	-3.54	107.76	112.25
6	B	1023	FUC	O4-C4-C3	-2.98	103.63	110.34
6	B	1023	FUC	O2-C2-C3	-2.84	104.40	110.12
6	B	1023	FUC	O3-C3-C2	-2.50	105.48	110.00
7	B	1037	MAN	C2-C3-C4	-2.46	106.86	111.04
7	B	1033	BMA	C3-C4-C5	-2.41	105.99	110.20
6	B	1023	FUC	C6-C5-C4	-2.20	108.75	113.08
7	B	1032	NAG	C2-N2-C7	-2.16	120.26	123.04
7	B	1036	MAN	C1-C2-C3	2.21	112.15	109.54
6	B	1021	NAG	C1-O5-C5	2.31	115.18	112.25
7	B	1036	MAN	O2-C2-C1	2.61	114.44	109.21
6	B	1023	FUC	C1-C2-C3	2.62	112.64	109.54

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
7	B	1037	MAN	C1-O5-C5	4.31	117.72	112.25
7	B	1035	MAN	C1-O5-C5	4.75	118.28	112.25
7	B	1036	MAN	C1-O5-C5	6.65	120.69	112.25

There are no chirality outliers.

There are no torsion outliers.

All (1) ring outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	1037	MAN	C1-C2-C3-C4-C5-O5

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	B	1032	NAG	2	0

## 5.6 Ligand geometry

Of 11 ligands modelled in this entry, 1 is monoatomic - leaving 10 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	$\# Z  > 2$
3	NAG	A	1001	1	14,14,15	0.55	0	15,19,21	2.06	1 (6%)
3	NAG	A	1011	1	14,14,15	0.50	0	15,19,21	0.68	0
4	PO4	A	2	-	4,4,4	0.47	0	6,6,6	0.27	0
4	PO4	A	3	-	4,4,4	0.45	0	6,6,6	0.27	0
4	PO4	A	4	-	4,4,4	0.44	0	6,6,6	0.28	0
4	PO4	A	6	-	4,4,4	0.42	0	6,6,6	0.27	0
4	PO4	B	1	-	4,4,4	0.43	0	6,6,6	0.27	0
3	NAG	B	1001	2	14,14,15	0.46	0	15,19,21	1.33	2 (13%)
3	NAG	B	1011	2	14,14,15	0.48	0	15,19,21	0.75	1 (6%)
4	PO4	B	5	-	4,4,4	0.52	0	6,6,6	0.27	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
3	NAG	A	1001	1	-	0/6/23/26	0/1/1/1
3	NAG	A	1011	1	-	0/6/23/26	0/1/1/1
4	PO4	A	2	-	-	0/0/0/0	0/0/0/0
4	PO4	A	3	-	-	0/0/0/0	0/0/0/0
4	PO4	A	4	-	-	0/0/0/0	0/0/0/0
4	PO4	A	6	-	-	0/0/0/0	0/0/0/0
4	PO4	B	1	-	-	0/0/0/0	0/0/0/0
3	NAG	B	1001	2	-	0/6/23/26	0/1/1/1
3	NAG	B	1011	2	-	0/6/23/26	0/1/1/1
4	PO4	B	5	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	1001	NAG	C6-C5-C4	-2.21	107.55	113.02
3	B	1011	NAG	C1-O5-C5	2.14	114.96	112.25
3	B	1001	NAG	C1-O5-C5	3.50	116.69	112.25
3	A	1001	NAG	C1-O5-C5	7.57	121.85	112.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

4 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1001	NAG	2	0
3	A	1011	NAG	1	0
3	B	1001	NAG	1	0
4	B	5	PO4	1	0

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues ⓘ

There are no chain breaks in this entry.

## 6 Fit of model and data ⓘ

### 6.1 Protein, DNA and RNA chains ⓘ

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<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	A	633/640 (98%)	-0.02	27 (4%) 39 40	27, 49, 84, 146	0
2	B	156/162 (96%)	-0.29	5 (3%) 51 51	24, 47, 75, 110	0
All	All	789/802 (98%)	-0.08	32 (4%) 41 42	24, 48, 84, 146	0

All (32) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	722	ASN	7.3
1	A	323	PRO	5.8
1	A	322	PRO	5.1
1	A	321	PHE	4.8
1	A	325	ARG	4.8
2	B	139	SER	4.7
2	B	86	ASN	4.6
1	A	723	ASN	4.5
1	A	724	GLY	4.5
1	A	247	TYR	4.4
1	A	356	ASP	3.7
2	B	241	HIS	3.5
1	A	153	VAL	3.4
1	A	634	GLU	3.3
1	A	720	LYS	3.1
1	A	729	THR	2.9
1	A	208	ARG	2.8
1	A	726	PHE	2.8
1	A	195	SER	2.7
1	A	244	GLU	2.6
1	A	336	THR	2.6
1	A	405	VAL	2.6
1	A	725	ALA	2.6
2	B	240	ASP	2.4

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	310	THR	2.4
1	A	358	LYS	2.4
1	A	326	SER	2.3
1	A	560	ASP	2.3
1	A	721	GLN	2.2
1	A	694	GLU	2.1
2	B	183	ASP	2.1
1	A	406	GLY	2.0

## 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](#)

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
7	NAG	B	1031	14/15	0.95	0.11	-0.94	52,58,67,70	0
7	MAN	B	1034	11/12	0.73	0.32	-	120,124,127,130	0
7	MAN	B	1036	11/12	0.72	0.45	-	130,141,145,148	0
7	NAG	B	1032	14/15	0.92	0.15	-	50,71,79,88	0
7	MAN	B	1037	11/12	0.82	0.43	-	125,127,128,129	0
6	NAG	B	1021	14/15	0.79	0.31	-	72,88,108,111	0
7	BMA	B	1033	11/12	0.74	0.20	-	98,104,107,114	0
6	FUC	B	1023	10/11	0.76	0.44	-	114,117,122,125	0
6	NAG	B	1022	14/15	0.79	0.46	-	109,117,125,128	0
7	MAN	B	1035	11/12	0.75	0.30	-	113,119,125,131	0

## 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( $\text{\AA}^2$ )	Q<0.9
3	NAG	B	1001	14/15	0.71	0.42	5.85	102,110,112,115	0
4	PO4	B	1	5/5	0.79	0.27	3.58	125,126,127,128	0
4	PO4	A	4	5/5	0.91	0.19	2.05	102,103,108,108	0
4	PO4	A	2	5/5	0.98	0.12	0.42	61,66,70,72	0
4	PO4	A	3	5/5	0.95	0.14	-0.03	103,107,108,109	0
5	K	A	1	1/1	0.99	0.09	-1.44	54,54,54,54	0
3	NAG	A	1011	14/15	0.82	0.32	-	82,94,101,102	0
4	PO4	A	6	5/5	0.93	0.12	-	103,106,106,108	0
3	NAG	B	1011	14/15	0.80	0.47	-	104,114,122,123	0
3	NAG	A	1001	14/15	0.77	0.23	-	68,86,91,91	0
4	PO4	B	5	5/5	0.83	0.26	-	109,110,112,115	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.