



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 11, 2023 – 02:02 PM EDT

PDB ID : 7RI2
Title : Crystal structure of anti-HIV llama VHH antibody A12 in complex with HIV-1 C1086 gp120
Authors : Zhou, T.; Kwong, P.D.
Deposited on : 2021-07-19
Resolution : 2.80 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

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

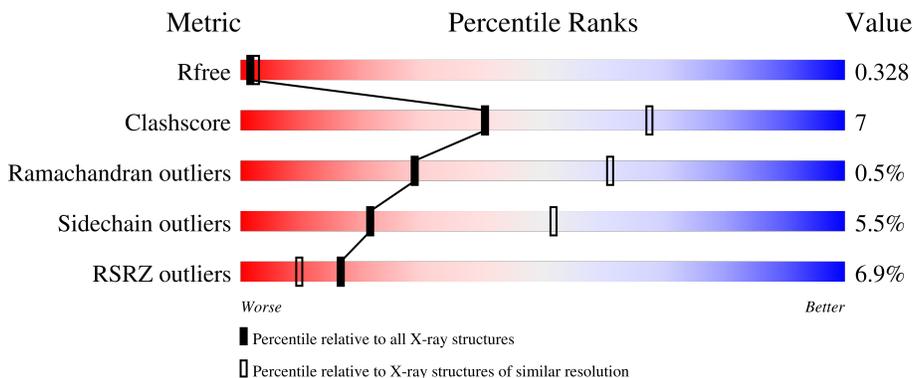
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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 ≥ 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	A	382	
2	B	143	

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	A	508	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 3399 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 Glycoprotein 120.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	293	2311	1449	398	447	17	0	0	0

There are 78 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	20	ASP	-	expression tag	UNP R4GRV3
A	21	ILE	-	expression tag	UNP R4GRV3
A	22	ALA	-	expression tag	UNP R4GRV3
A	23	THR	-	expression tag	UNP R4GRV3
A	24	MET	-	expression tag	UNP R4GRV3
A	25	TYR	-	expression tag	UNP R4GRV3
A	26	SER	-	expression tag	UNP R4GRV3
A	27	MET	-	expression tag	UNP R4GRV3
A	28	GLN	-	expression tag	UNP R4GRV3
A	29	LEU	-	expression tag	UNP R4GRV3
A	30	ALA	-	expression tag	UNP R4GRV3
A	31	SER	-	expression tag	UNP R4GRV3
A	32	CYS	-	expression tag	UNP R4GRV3
A	33	VAL	-	expression tag	UNP R4GRV3
A	34	THR	-	expression tag	UNP R4GRV3
A	35	LEU	-	expression tag	UNP R4GRV3
A	36	THR	-	expression tag	UNP R4GRV3
A	37	LEU	-	expression tag	UNP R4GRV3
A	38	VAL	-	expression tag	UNP R4GRV3
A	39	LEU	-	expression tag	UNP R4GRV3
A	40	LEU	-	expression tag	UNP R4GRV3
A	41	VAL	-	expression tag	UNP R4GRV3
A	42	ASN	-	expression tag	UNP R4GRV3
A	43	SER	-	expression tag	UNP R4GRV3
A	63	LYS	ARG	conflict	UNP R4GRV3
A	85	VAL	ALA	conflict	UNP R4GRV3
A	87	ALA	GLY	conflict	UNP R4GRV3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	102	GLU	ASP	conflict	UNP R4GRV3
A	114	GLU	GLN	conflict	UNP R4GRV3
A	200	ALA	THR	conflict	UNP R4GRV3
A	201	ILE	VAL	conflict	UNP R4GRV3
A	215	LEU	ILE	conflict	UNP R4GRV3
A	223	PHE	TYR	conflict	UNP R4GRV3
A	231	LYS	GLU	conflict	UNP R4GRV3
A	234	ASN	GLU	conflict	UNP R4GRV3
A	240	ARG	ASN	conflict	UNP R4GRV3
A	267	GLU	LYS	conflict	UNP R4GRV3
A	271	ILE	VAL	conflict	UNP R4GRV3
A	279	ASN	ASP	conflict	UNP R4GRV3
A	289	ASN	LYS	conflict	UNP R4GRV3
A	293	ASN	GLU	conflict	UNP R4GRV3
A	316	ASN	GLY	conflict	UNP R4GRV3
A	325	ASN	ASP	conflict	UNP R4GRV3
A	326	ILE	THR	conflict	UNP R4GRV3
A	334	ASN	SER	conflict	UNP R4GRV3
A	336	SER	GLU	conflict	UNP R4GRV3
A	340	ASN	LYS	conflict	UNP R4GRV3
A	346	GLY	SER	conflict	UNP R4GRV3
A	347	GLU	LYS	conflict	UNP R4GRV3
A	348	GLU	ILE	conflict	UNP R4GRV3
A	350	ALA	GLN	conflict	UNP R4GRV3
A	351	LYS	GLU	conflict	UNP R4GRV3
A	356	SER	ASN	conflict	UNP R4GRV3
A	358	THR	ALA	conflict	UNP R4GRV3
A	364	SER	HIS	conflict	UNP R4GRV3
A	388	SER	THR	conflict	UNP R4GRV3
A	389	ASP	LYS	conflict	UNP R4GRV3
A	396	ARG	-	insertion	UNP R4GRV3
A	398	GLY	SER	conflict	UNP R4GRV3
A	400	TYR	ALA	conflict	UNP R4GRV3
A	402	HIS	SER	conflict	UNP R4GRV3
A	408	ARG	SER	conflict	UNP R4GRV3
A	409	SER	VAL	conflict	UNP R4GRV3
A	412	GLY	THR	conflict	UNP R4GRV3
A	417	GLN	PRO	conflict	UNP R4GRV3
A	419	LYS	ARG	conflict	UNP R4GRV3
A	424	ILE	VAL	conflict	UNP R4GRV3
A	429	GLU	GLY	conflict	UNP R4GRV3
A	434	ILE	MET	conflict	UNP R4GRV3

Continued on next page...

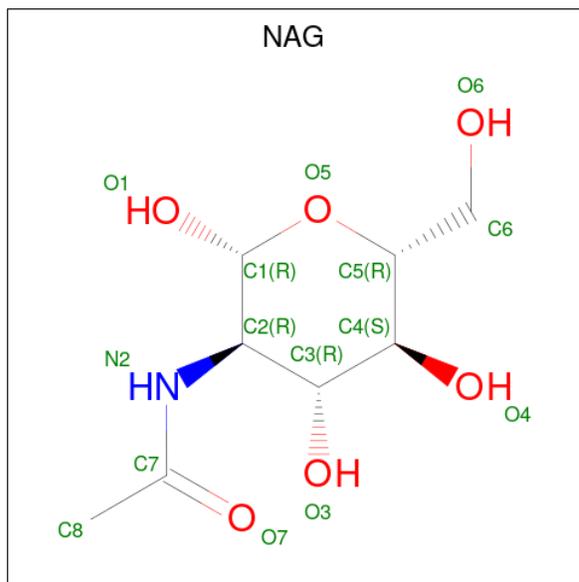
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	440	GLU	LYS	conflict	UNP R4GRV3
A	442	GLU	ASN	conflict	UNP R4GRV3
A	446	ASN	LYS	conflict	UNP R4GRV3
A	448	ASN	ILE	conflict	UNP R4GRV3
A	455	LEU	THR	conflict	UNP R4GRV3
A	458B	GLN	-	insertion	UNP R4GRV3
A	458C	SER	-	insertion	UNP R4GRV3
A	458E	GLU	ASP	conflict	UNP R4GRV3
A	458F	THR	ASP	conflict	UNP R4GRV3

- Molecule 2 is a protein called anti-HIV llama VHH antibody A12.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace	
			Total	C	N	O				S
2	B	127	976	609	169	194	4	2	0	0

- Molecule 3 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



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

Continued on next page...

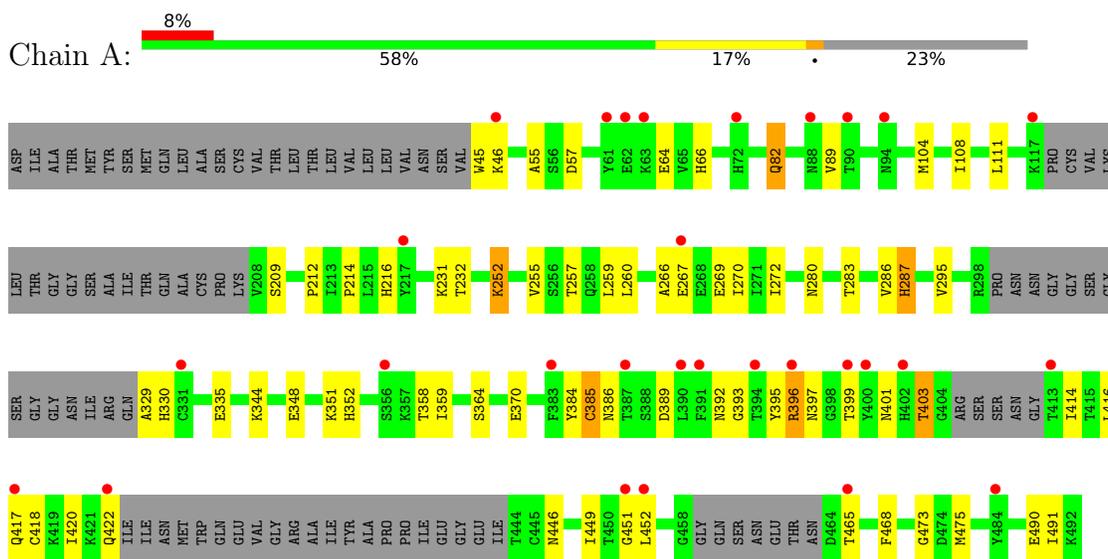
Continued from previous page...

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	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	A	1	Total	C	N	O	0	0
			14	8	1	5		

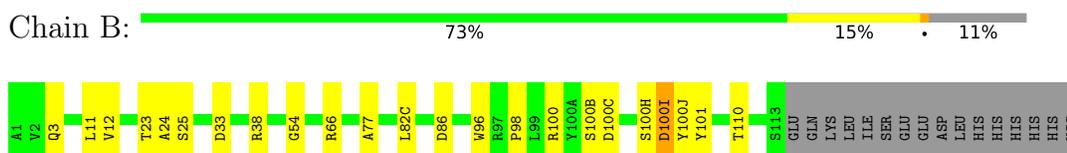
3 Residue-property plots [i](#)

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: Glycoprotein 120



- Molecule 2: anti-HIV llama VHH antibody A12



4 Data and refinement statistics

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α , β , γ	66.60Å 66.60Å 266.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	23.55 – 2.80 23.55 – 2.80	Depositor EDS
% Data completeness (in resolution range)	85.2 (23.55-2.80) 85.2 (23.55-2.80)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.05 (at 2.80Å)	Xtrriage
Refinement program	PHENIX 1.19-4092	Depositor
R, R_{free}	0.273 , 0.330 0.273 , 0.328	Depositor DCC
R_{free} test set	623 reflections (4.67%)	wwPDB-VP
Wilson B-factor (Å ²)	79.9	Xtrriage
Anisotropy	0.746	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.24 , 65.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	3399	wwPDB-VP
Average B, all atoms (Å ²)	151.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.76% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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

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	0.26	0/2356	0.48	0/3189
2	B	0.29	0/999	0.62	0/1353
All	All	0.27	0/3355	0.52	0/4542

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
2	B	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
2	B	24	ALA	Peptide

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	A	2311	0	2214	36	1
2	B	976	0	922	13	1
3	A	112	0	104	0	0
All	All	3399	0	3240	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 7.

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:A:389:ASP:HA	1:A:403:THR:HG21	1.58	0.83
1:A:392:ASN:HD22	1:A:403:THR:HG23	1.44	0.81
2:B:66:ARG:NH2	2:B:86:ASP:OD2	2.25	0.69
1:A:295:VAL:HG22	1:A:446:ASN:HA	1.82	0.62
1:A:358:THR:HG23	1:A:396:ARG:HB3	1.83	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)
1:A:491:ILE:CD1	2:B:110:THR:CG2[5_554]	1.78	0.42

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	279/382 (73%)	263 (94%)	15 (5%)	1 (0%)	34 66
2	B	125/143 (87%)	124 (99%)	0	1 (1%)	19 49
All	All	404/525 (77%)	387 (96%)	15 (4%)	2 (0%)	29 61

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	100(I)	ASP
1	A	82	GLN

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	A	261/334 (78%)	242 (93%)	19 (7%)	14	38
2	B	101/117 (86%)	100 (99%)	1 (1%)	76	93
All	All	362/451 (80%)	342 (94%)	20 (6%)	21	52

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	396	ARG
1	A	420	ILE
2	B	100(I)	ASP
1	A	422	GLN
1	A	269	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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

8 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
3	NAG	A	504	1	14,14,15	1.05	1 (7%)	17,19,21	1.22	1 (5%)
3	NAG	A	506	1	14,14,15	0.25	0	17,19,21	0.49	0
3	NAG	A	507	1	14,14,15	0.33	0	17,19,21	0.70	1 (5%)
3	NAG	A	503	1	14,14,15	0.33	0	17,19,21	0.50	0
3	NAG	A	508	1	14,14,15	0.50	0	17,19,21	0.56	0
3	NAG	A	505	1	14,14,15	0.32	0	17,19,21	0.56	0
3	NAG	A	501	1	14,14,15	0.23	0	17,19,21	0.34	0
3	NAG	A	502	1	14,14,15	0.32	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
3	NAG	A	504	1	-	1/6/23/26	0/1/1/1
3	NAG	A	506	1	-	0/6/23/26	0/1/1/1
3	NAG	A	507	1	-	4/6/23/26	0/1/1/1
3	NAG	A	503	1	-	2/6/23/26	0/1/1/1
3	NAG	A	508	1	-	0/6/23/26	0/1/1/1
3	NAG	A	505	1	-	4/6/23/26	0/1/1/1
3	NAG	A	501	1	-	0/6/23/26	0/1/1/1
3	NAG	A	502	1	-	2/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	504	NAG	O5-C1	3.69	1.49	1.43

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
3	A	504	NAG	C1-O5-C5	4.79	118.68	112.19
3	A	507	NAG	C1-O5-C5	2.36	115.38	112.19

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	503	NAG	O5-C5-C6-O6
3	A	502	NAG	C8-C7-N2-C2
3	A	502	NAG	O7-C7-N2-C2
3	A	505	NAG	C8-C7-N2-C2
3	A	505	NAG	O7-C7-N2-C2

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	A	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	A	354:PRO	C	356:SER	N	3.86

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, 95th 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 ²)	Q<0.9
1	A	293/382 (76%)	0.59	29 (9%) 7 4	105, 164, 228, 293	0
2	B	127/143 (88%)	0.03	0 100 100	74, 109, 137, 175	1 (0%)
All	All	420/525 (80%)	0.42	29 (6%) 16 10	74, 144, 217, 293	1 (0%)

The worst 5 of 29 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	413	THR	7.0
1	A	396	ARG	5.7
1	A	400	TYR	5.4
1	A	90	THR	4.7
1	A	390	LEU	4.5

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, 95th 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(\AA^2)	Q<0.9
3	NAG	A	507	14/15	0.73	0.32	174,187,197,201	0
3	NAG	A	508	14/15	0.74	0.41	147,179,208,216	0
3	NAG	A	505	14/15	0.80	0.20	185,195,199,201	0
3	NAG	A	502	14/15	0.82	0.32	182,187,191,193	0
3	NAG	A	506	14/15	0.85	0.22	139,154,160,163	0
3	NAG	A	503	14/15	0.87	0.22	158,177,180,181	0
3	NAG	A	501	14/15	0.88	0.25	153,159,164,168	0
3	NAG	A	504	14/15	0.91	0.19	144,150,152,153	0

6.5 Other polymers [i](#)

There are no such residues in this entry.