



wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 02:28 PM EDT

PDB ID : 2Y8T
Title : Co-structure of AMA1 with a surface exposed region of RON2 from Toxoplasma gondii
Authors : Tonkin, M.L.; Roques, M.; Lamarque, M.H.; Pugniere, M.; Douguet, D.; Crawford, J.; Lebrun, M.; Boulanger, M.J.
Deposited on : 2011-02-10
Resolution : 1.95 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	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
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)	:	2.37.1

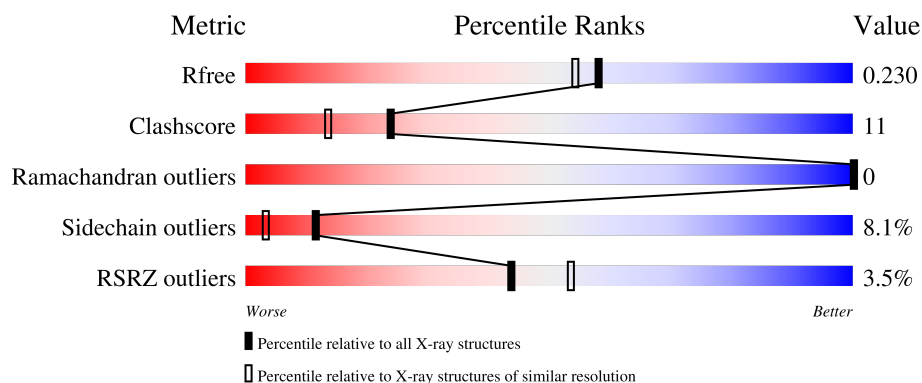
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


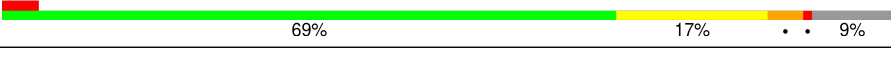
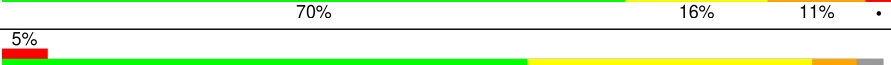

The reported resolution of this entry is 1.95 Å.

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	2580 (1.96-1.96)
Clashscore	141614	2705 (1.96-1.96)
Ramachandran outliers	138981	2678 (1.96-1.96)
Sidechain outliers	138945	2678 (1.96-1.96)
RSRZ outliers	127900	2539 (1.96-1.96)

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	432	 3% 73% 13% 10%
1	D	432	 4% 69% 17% 9%
2	B	37	 70% 16% 11% 3%
2	E	37	 5% 59% 32% 5%

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
2	9AT	B	1333	-	-	X	-
4	BO3	D	1475	-	X	-	-
4	BO3	D	1476	-	X	-	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 7336 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 APICAL MEMBRANE ANTIGEN, PUTATIVE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	388	Total	C	N	O	S	0	3	0
			3067	1923	528	594	22			
1	D	395	Total	C	N	O	S	0	0	0
			3117	1962	533	601	21			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	GLY	-	expression tag	UNP B9QC59
A	60	SER	-	expression tag	UNP B9QC59
A	61	ALA	-	expression tag	UNP B9QC59
A	62	MET	-	expression tag	UNP B9QC59
A	63	GLY	-	expression tag	UNP B9QC59
A	485	ALA	-	expression tag	UNP B9QC59
A	486	ALA	-	expression tag	UNP B9QC59
A	487	LEU	-	expression tag	UNP B9QC59
A	488	VAL	-	expression tag	UNP B9QC59
A	489	PRO	-	expression tag	UNP B9QC59
A	490	ARG	-	expression tag	UNP B9QC59
D	59	GLY	-	expression tag	UNP B9QC59
D	60	SER	-	expression tag	UNP B9QC59
D	61	ALA	-	expression tag	UNP B9QC59
D	62	MET	-	expression tag	UNP B9QC59
D	63	GLY	-	expression tag	UNP B9QC59
D	485	ALA	-	expression tag	UNP B9QC59
D	486	ALA	-	expression tag	UNP B9QC59
D	487	LEU	-	expression tag	UNP B9QC59
D	488	VAL	-	expression tag	UNP B9QC59
D	489	PRO	-	expression tag	UNP B9QC59
D	490	ARG	-	expression tag	UNP B9QC59

- Molecule 2 is a protein called RHOPTRY NECK PROTEIN 2.

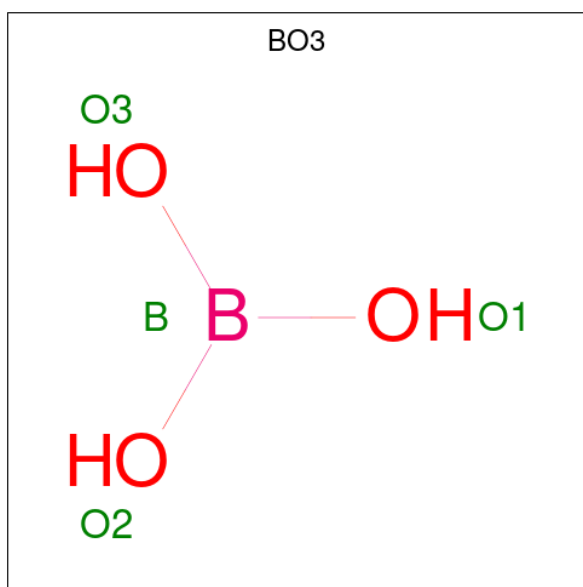
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	37	Total	C	N	O	S	0	0	0
			260	160	44	53	3			
2	E	36	Total	C	N	O	S	0	0	0
			252	156	42	51	3			

- 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
3	A	1	Total	C	N	O	0	0
			14	8	1	5		
3	D	1	Total	C	N	O	0	0
			14	8	1	5		

- Molecule 4 is BORIC ACID (three-letter code: BO3) (formula: BH_3O_3).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	1	Total B O 4 1 3	0	0
4	D	1	Total B O 4 1 3	0	0

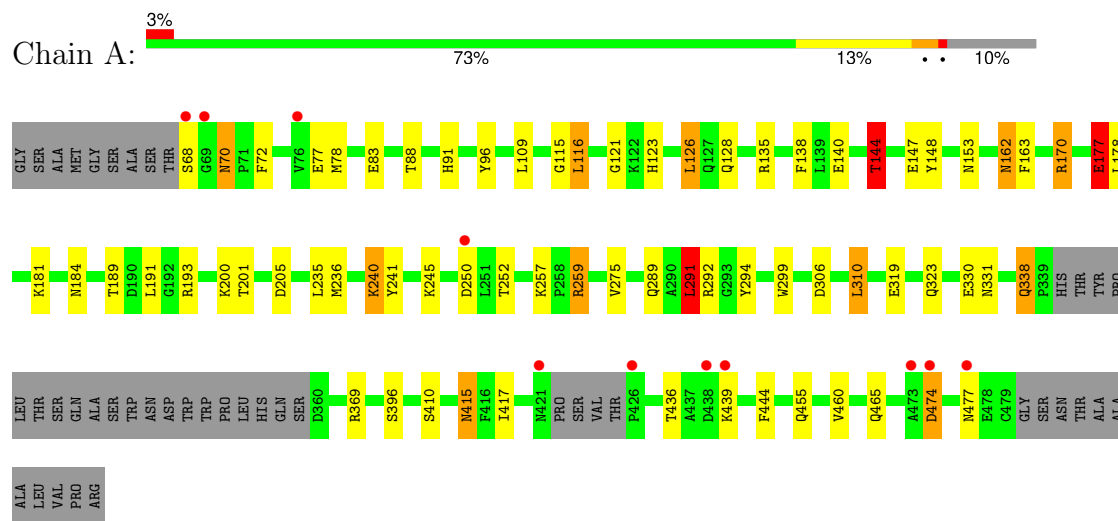
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	319	Total O 319 319	0	0
5	B	25	Total O 25 25	0	0
5	D	246	Total O 246 246	0	0
5	E	14	Total O 14 14	0	0

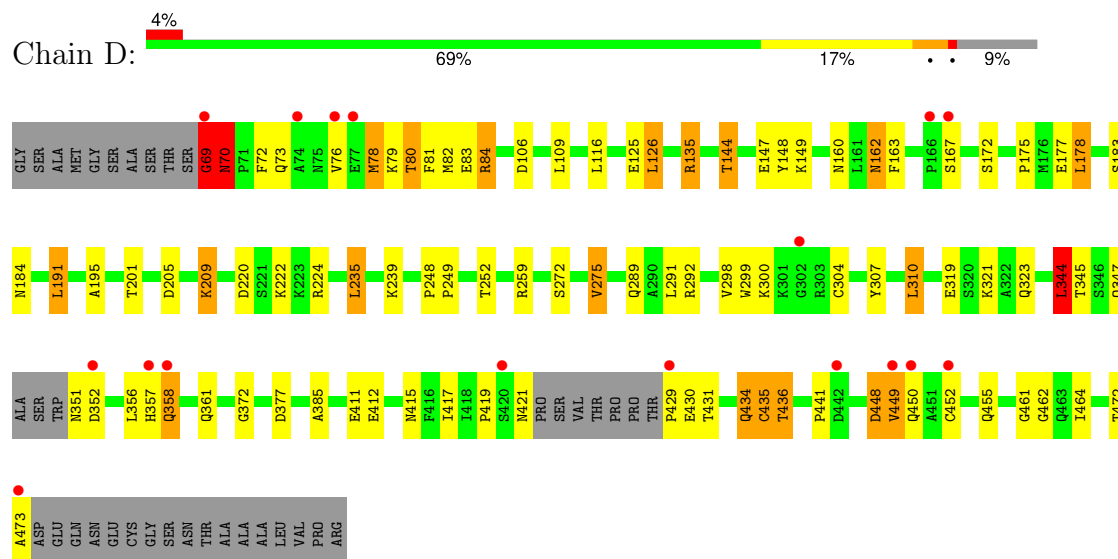
3 Residue-property plots [i](#)

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: APICAL MEMBRANE ANTIGEN, PUTATIVE

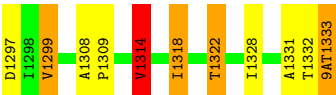


• Molecule 1: APICAL MEMBRANE ANTIGEN, PUTATIVE



• Molecule 2: RHOPTRY NECK PROTEIN 2





● Molecule 2: RHOPTRY NECK PROTEIN 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	69.86Å 96.38Å 78.35Å 90.00° 115.64° 90.00°	Depositor
Resolution (Å)	39.81 – 1.95 39.81 – 1.95	Depositor EDS
% Data completeness (in resolution range)	100.0 (39.81-1.95) 100.0 (39.81-1.95)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.05 (at 1.95Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.170 , 0.228 0.171 , 0.230	Depositor DCC
R_{free} test set	3453 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	19.6	Xtriage
Anisotropy	0.060	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 52.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.023 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7336	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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: 9AT, BO3, 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	0.97	1/3152 (0.0%)	0.95	12/4267 (0.3%)
1	D	0.89	2/3200 (0.1%)	0.94	11/4339 (0.3%)
2	B	0.95	0/255	1.14	1/349 (0.3%)
2	E	0.78	0/255	0.92	1/349 (0.3%)
All	All	0.93	3/6862 (0.0%)	0.95	25/9304 (0.3%)

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
1	D	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	435	CYS	CB-SG	-10.96	1.63	1.82
1	A	177	GLU	CB-CG	6.25	1.64	1.52
1	D	125	GLU	CG-CD	5.05	1.59	1.51

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	344	LEU	CA-CB-CG	8.60	135.09	115.30
1	A	236	MET	CA-CB-CG	-7.61	100.37	113.30
2	B	1314	VAL	CG1-CB-CG2	7.51	122.92	110.90
1	D	275	VAL	CG1-CB-CG2	7.38	122.71	110.90
1	A	116	LEU	CB-CG-CD1	6.81	122.57	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	D	434	GLN	Peptide
1	D	69	GLY	Peptide

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	3067	0	2928	52	0
1	D	3117	0	2959	84	0
2	B	260	0	260	14	0
2	E	252	0	250	15	0
3	A	14	0	13	1	0
3	D	14	0	13	1	0
4	D	8	0	6	0	0
5	A	319	0	0	8	1
5	B	25	0	0	1	0
5	D	246	0	0	15	0
5	E	14	0	0	0	0
All	All	7336	0	6429	146	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 11.

The worst 5 of 146 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)
2:B:1332:THR:C	2:B:1333:9AT:HN2	1.29	1.35
2:B:1332:THR:C	2:B:1333:9AT:N	1.84	1.29
1:A:259[A]:ARG:HD2	5:A:2059:HOH:O	1.43	1.16
1:A:144:THR:HG21	5:B:2022:HOH:O	1.51	1.10
1:D:292:ARG:HD3	1:D:411:GLU:OE1	1.54	1.05

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)
5:A:2131:HOH:O	5:A:2249:HOH:O[2_555]	2.02	0.18

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	385/432 (89%)	377 (98%)	8 (2%)	0	100	100
1	D	389/432 (90%)	376 (97%)	13 (3%)	0	100	100
2	B	35/37 (95%)	34 (97%)	1 (3%)	0	100	100
2	E	34/37 (92%)	33 (97%)	1 (3%)	0	100	100
All	All	843/938 (90%)	820 (97%)	23 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	340/372 (91%)	319 (94%)	21 (6%)	18	7
1	D	344/372 (92%)	314 (91%)	30 (9%)	10	2
2	B	28/28 (100%)	24 (86%)	4 (14%)	3	0
2	E	28/28 (100%)	22 (79%)	6 (21%)	1	0
All	All	740/800 (92%)	679 (92%)	61 (8%)	11	3

5 of 61 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	84	ARG
2	E	1297	ASP
1	D	162	ASN
1	D	449	VAL
2	E	1322	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	D	101	GLN
1	D	162	ASN
1	D	160	ASN
1	D	207	ASN
1	A	162	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is 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	$\# Z > 2$
2	9AT	B	1333	-	7,7,7	0.67	0	9,9,9	1.54	2 (22%)

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
2	9AT	B	1333	-	-	0/7/8/8	-

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1333	9AT	CA-C-NXT	-2.55	113.01	117.02
2	B	1333	9AT	O-C-CA	2.16	123.17	119.94

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	1333	9AT	4	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

4 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	1480	1	14,14,15	0.72	0	17,19,21	2.14	7 (41%)
3	NAG	D	1474	1	14,14,15	0.48	0	17,19,21	1.32	2 (11%)
4	BO3	D	1476	-	3,3,3	5.21	3 (100%)	3,3,3	0.22	0
4	BO3	D	1475	-	3,3,3	5.25	3 (100%)	3,3,3	0.77	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	1480	1	-	4/6/23/26	0/1/1/1
3	NAG	D	1474	1	-	4/6/23/26	0/1/1/1

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	1475	BO3	B-O3	5.57	1.54	1.36
4	D	1476	BO3	B-O3	5.38	1.54	1.36
4	D	1475	BO3	B-O1	5.30	1.53	1.36
4	D	1476	BO3	B-O1	5.13	1.53	1.36
4	D	1476	BO3	B-O2	5.11	1.53	1.36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	1480	NAG	C8-C7-N2	4.29	123.23	116.12
3	A	1480	NAG	O5-C5-C4	-3.33	102.73	110.83
3	A	1480	NAG	O5-C1-C2	-3.21	106.32	111.29
3	A	1480	NAG	O7-C7-C8	-2.71	117.22	122.05
3	D	1474	NAG	C1-C2-N2	2.66	114.63	110.43

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	A	1480	NAG	O5-C5-C6-O6
3	A	1480	NAG	C8-C7-N2-C2
3	A	1480	NAG	O7-C7-N2-C2
3	D	1474	NAG	C8-C7-N2-C2
3	D	1474	NAG	O7-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1480	NAG	1	0
3	D	1474	NAG	1	0

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
2	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	1332:THR	C	1333:9AT	N	1.84

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	388/432 (89%)	-0.11	11 (2%) 53 62	8, 17, 39, 51	5 (1%)
1	D	395/432 (91%)	0.13	17 (4%) 35 45	9, 22, 44, 65	5 (1%)
2	B	36/37 (97%)	-0.26	0 100 100	10, 20, 33, 37	0
2	E	36/37 (97%)	0.58	2 (5%) 24 33	14, 30, 45, 46	0
All	All	855/938 (91%)	0.02	30 (3%) 44 53	8, 20, 43, 65	10 (1%)

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	68	SER	6.5
1	D	450	GLN	3.7
2	E	1331	ALA	3.6
1	D	76	VAL	3.4
1	D	473	ALA	3.3

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, 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(Å ²)	Q<0.9
2	9AT	B	1333	8/8	0.84	0.18	42,44,45,45	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, 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
4	BO3	D	1476	4/4	0.82	0.12	44,45,45,48	0
4	BO3	D	1475	4/4	0.87	0.12	41,42,42,43	0
3	NAG	D	1474	14/15	0.93	0.09	22,29,33,36	0
3	NAG	A	1480	14/15	0.94	0.13	17,21,29,36	0

6.5 Other polymers [i](#)

There are no such residues in this entry.