



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 19, 2023 – 06:53 PM JST

PDB ID : 6M0J  
Title : Crystal structure of SARS-CoV-2 spike receptor-binding domain bound with ACE2  
Authors : Wang, X.; Lan, J.; Ge, J.; Yu, J.; Shan, S.  
Deposited on : 2020-02-21  
Resolution : 2.45 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

<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

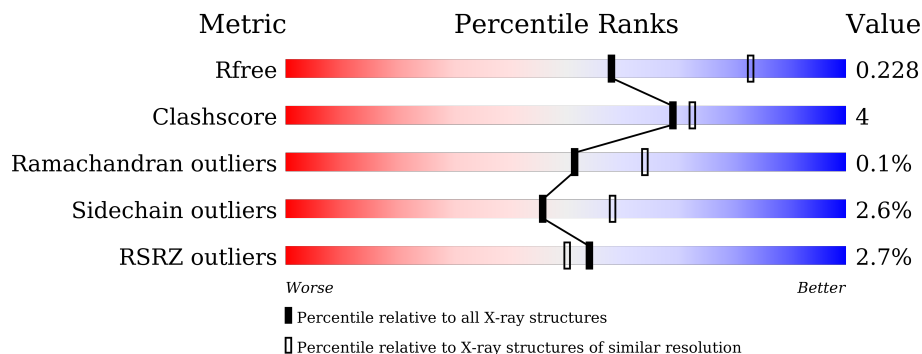
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.45 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	1544 (2.48-2.44)
Clashscore	141614	1613 (2.48-2.44)
Ramachandran outliers	138981	1598 (2.48-2.44)
Sidechain outliers	138945	1598 (2.48-2.44)
RSRZ outliers	127900	1523 (2.48-2.44)

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	A	603	 88% 10% ..
2	E	229	 72% 12% 15%

## 2 Entry composition i

There are 6 unique types of molecules in this entry. The entry contains 6571 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 Angiotensin-converting enzyme 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	597	4877	3120	808	920	29	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	616	HIS	-	expression tag	UNP Q9BYF1
A	617	HIS	-	expression tag	UNP Q9BYF1
A	618	HIS	-	expression tag	UNP Q9BYF1
A	619	HIS	-	expression tag	UNP Q9BYF1
A	620	HIS	-	expression tag	UNP Q9BYF1
A	621	HIS	-	expression tag	UNP Q9BYF1

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	E	194	1542	988	257	289	8	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	542	HIS	-	expression tag	UNP P0DTC2
E	543	HIS	-	expression tag	UNP P0DTC2
E	544	HIS	-	expression tag	UNP P0DTC2
E	545	HIS	-	expression tag	UNP P0DTC2
E	546	HIS	-	expression tag	UNP P0DTC2
E	547	HIS	-	expression tag	UNP P0DTC2

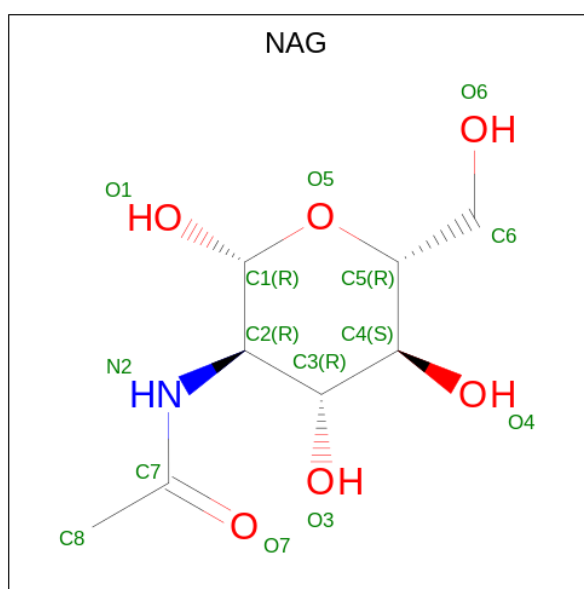
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Cl	0	0
			1	1		

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

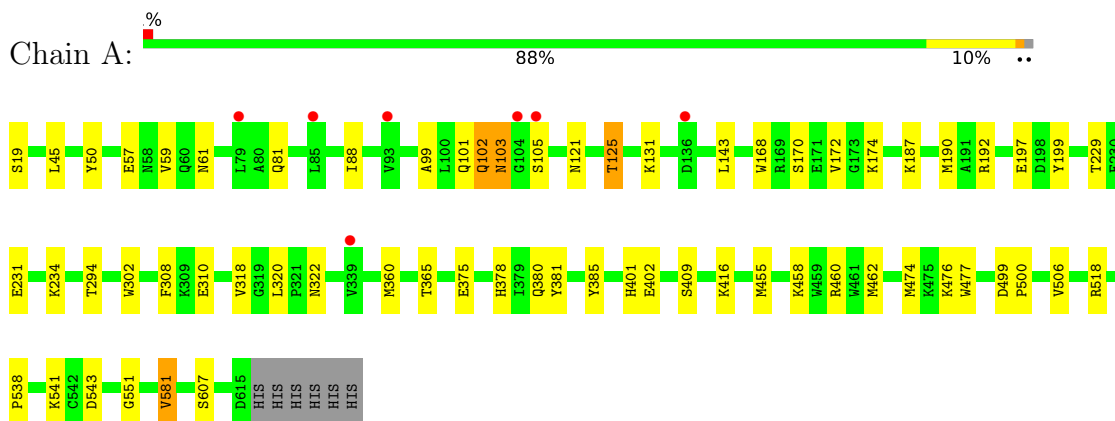
- Molecule 6 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	71	Total 71	O 71	0	0
6	E	9	Total 9	O 9	0	0

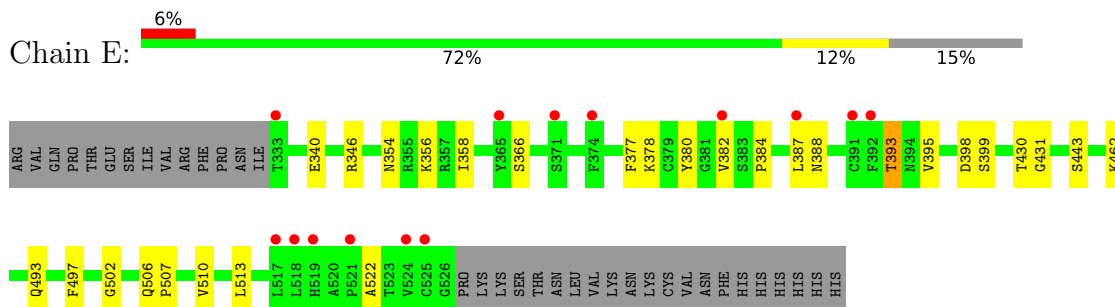
### 3 Residue-property plots [i](#)

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: Angiotensin-converting enzyme 2



- Molecule 2: Spike protein S1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	104.67Å 104.67Å 228.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.10 – 2.45 95.18 – 2.45	Depositor EDS
% Data completeness (in resolution range)	99.9 (53.10-2.45) 100.0 (95.18-2.45)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.45Å)	Xtrriage
Refinement program	PHENIX 1.15.2_3472	Depositor
R, $R_{free}$	0.192 , 0.227 0.192 , 0.228	Depositor DCC
$R_{free}$ test set	2426 reflections (5.10%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	60.7	Xtrriage
Anisotropy	0.239	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 47.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	6571	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, ZN, CL

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	0.44	0/5018	0.55	0/6818
2	E	0.40	0/1588	0.58	0/2161
All	All	0.43	0/6606	0.56	0/8979

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	4877	0	4646	32	0
2	E	1542	0	1460	15	0
3	A	1	0	0	0	0
4	A	1	0	0	0	0
5	A	56	0	52	0	0
5	E	14	0	13	0	0
6	A	71	0	0	1	0
6	E	9	0	0	0	0
All	All	6571	0	6171	47	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 4.



All (47) 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:458:LYS:HG2	1:A:462:MET:HE2	1.72	0.70
1:A:57:GLU:O	1:A:61:ASN:ND2	2.27	0.67
1:A:538:PRO:HD2	1:A:541:LYS:HD3	1.77	0.67
1:A:121:ASN:O	1:A:125:THR:HG23	1.95	0.67
2:E:384:PRO:HA	2:E:387:LEU:HG	1.78	0.64
1:A:192:ARG:NH2	1:A:197:GLU:O	2.31	0.63
1:A:229:THR:HB	1:A:581:VAL:HG13	1.80	0.62
1:A:320:LEU:HD13	1:A:380:GLN:HG2	1.80	0.62
1:A:131:LYS:HB3	1:A:143:LEU:HD23	1.82	0.62
1:A:103:ASN:OD1	1:A:103:ASN:N	2.37	0.57
2:E:443:SER:HB3	2:E:497:PHE:HB3	1.86	0.56
1:A:81:GLN:HE22	1:A:103:ASN:HB3	1.70	0.56
1:A:477:TRP:CE3	1:A:500:PRO:HG3	2.42	0.55
1:A:187:LYS:HD2	1:A:199:TYR:CZ	2.42	0.54
1:A:402:GLU:HB3	1:A:518:ARG:HD3	1.92	0.51
1:A:170:SER:O	1:A:174:LYS:HD2	2.10	0.51
1:A:474:MET:HE1	1:A:499:ASP:HB2	1.94	0.49
2:E:366:SER:OG	2:E:388:ASN:ND2	2.45	0.49
1:A:50:TYR:CE1	1:A:59:VAL:HG22	2.48	0.48
2:E:354:ASN:O	2:E:398:ASP:HA	2.13	0.47
1:A:231:GLU:OE2	1:A:234:LYS:NZ	2.30	0.47
1:A:460:ARG:HH21	1:A:506:VAL:HA	1.80	0.47
1:A:101:GLN:O	1:A:103:ASN:N	2.48	0.47
1:A:460:ARG:NH2	1:A:506:VAL:HA	2.29	0.47
2:E:431:GLY:HA3	2:E:513:LEU:O	2.15	0.46
2:E:462:LYS:HB3	2:E:462:LYS:HE2	1.87	0.45
1:A:302:TRP:CH2	1:A:310:GLU:HG3	2.52	0.44
2:E:358:ILE:HB	2:E:395:VAL:HB	2.01	0.43
1:A:168:TRP:CZ3	1:A:172:VAL:HG21	2.54	0.43
1:A:99:ALA:O	1:A:102:GLN:HG2	2.19	0.43
2:E:382:VAL:CG1	2:E:387:LEU:HB3	2.49	0.43
2:E:378:LYS:HA	2:E:378:LYS:HD3	1.93	0.42
1:A:45:LEU:HD12	1:A:45:LEU:HA	1.86	0.42
1:A:322:ASN:OD1	1:A:322:ASN:N	2.52	0.42
2:E:380:TYR:O	2:E:430:THR:HA	2.20	0.42
1:A:375:GLU:O	1:A:378:HIS:HB2	2.20	0.42
2:E:340:GLU:OE2	2:E:356:LYS:HE3	2.20	0.42
2:E:497:PHE:CE2	2:E:507:PRO:HB3	2.55	0.41
1:A:581:VAL:HG22	6:A:1058:HOH:O	2.19	0.41
1:A:402:GLU:HB3	1:A:518:ARG:CD	2.49	0.41

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:308:PHE:HZ	1:A:360:MET:HE3	1.86	0.41
1:A:318:VAL:O	1:A:551:GLY:HA3	2.20	0.41
2:E:399:SER:HA	2:E:510:VAL:O	2.21	0.40
2:E:393:THR:HA	2:E:522:ALA:HA	2.02	0.40
2:E:502:GLY:O	2:E:506:GLN:HG3	2.22	0.40
1:A:294:THR:HG23	1:A:365:THR:HA	2.03	0.40
1:A:416:LYS:HD2	1:A:543:ASP:HB3	2.04	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	596/603 (99%)	584 (98%)	11 (2%)	1 (0%)	47	57
2	E	193/229 (84%)	182 (94%)	11 (6%)	0	100	100
All	All	789/832 (95%)	766 (97%)	22 (3%)	1 (0%)	51	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	102	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	528/533 (99%)	514 (97%)	14 (3%)	44	57
2	E	168/202 (83%)	163 (97%)	5 (3%)	41	52
All	All	696/735 (95%)	677 (97%)	19 (3%)	46	57

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	19	SER
1	A	88	ILE
1	A	103	ASN
1	A	105	SER
1	A	125	THR
1	A	190	MET
1	A	381	TYR
1	A	385	TYR
1	A	401	HIS
1	A	409	SER
1	A	455	MET
1	A	476	LYS
1	A	581	VAL
1	A	607	SER
2	E	346	ARG
2	E	377	PHE
2	E	393	THR
2	E	493[A]	GLN
2	E	493[B]	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	81	GLN
1	A	102	GLN
1	A	154	ASN
1	A	194	ASN
2	E	388	ASN

### 5.3.3 RNA

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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
5	NAG	A	906	1	14,14,15	0.67	0	17,19,21	0.63	1 (5%)
5	NAG	A	905	1	14,14,15	0.51	0	17,19,21	0.55	0
5	NAG	E	601	2	14,14,15	0.48	0	17,19,21	0.62	1 (5%)
5	NAG	A	903	1	14,14,15	0.30	0	17,19,21	0.87	1 (5%)
5	NAG	A	904	1	14,14,15	0.45	0	17,19,21	1.11	1 (5%)

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	A	906	1	-	2/6/23/26	0/1/1/1
5	NAG	A	905	1	-	2/6/23/26	0/1/1/1
5	NAG	E	601	2	-	0/6/23/26	0/1/1/1
5	NAG	A	903	1	-	0/6/23/26	0/1/1/1
5	NAG	A	904	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	904	NAG	C1-O5-C5	4.09	117.74	112.19
5	A	903	NAG	C1-O5-C5	2.51	115.60	112.19
5	E	601	NAG	C1-O5-C5	2.22	115.20	112.19
5	A	906	NAG	C1-O5-C5	2.04	114.96	112.19

There are no chirality outliers.

All (6) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	906	NAG	O5-C5-C6-O6
5	A	904	NAG	O5-C5-C6-O6
5	A	906	NAG	C4-C5-C6-O6
5	A	904	NAG	C4-C5-C6-O6
5	A	905	NAG	C4-C5-C6-O6
5	A	905	NAG	O5-C5-C6-O6

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](#)

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	597/603 (99%)	0.27	7 (1%) 79 77	44, 59, 90, 127	0
2	E	194/229 (84%)	0.52	14 (7%) 15 11	54, 71, 114, 129	0
All	All	791/832 (95%)	0.33	21 (2%) 54 50	44, 61, 104, 129	0

All (21) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	104	GLY	5.7
2	E	517	LEU	5.5
1	A	105	SER	4.4
2	E	333	THR	4.4
2	E	521	PRO	4.1
2	E	387	LEU	4.0
2	E	524	VAL	3.7
2	E	519	HIS	3.4
1	A	136	ASP	3.4
2	E	371	SER	2.9
2	E	518	LEU	2.9
2	E	382	VAL	2.8
2	E	374	PHE	2.7
2	E	392	PHE	2.3
1	A	93	VAL	2.2
2	E	365	TYR	2.2
2	E	525	CYS	2.2
1	A	79	LEU	2.1
2	E	391	CYS	2.1
1	A	339	VAL	2.0
1	A	85	LEU	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](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	A	905	14/15	0.64	0.21	88,100,104,112	0
5	NAG	E	601	14/15	0.70	0.18	87,100,111,112	0
5	NAG	A	903	14/15	0.84	0.19	80,89,95,95	0
5	NAG	A	904	14/15	0.84	0.16	79,86,91,93	0
5	NAG	A	906	14/15	0.85	0.17	93,95,104,105	0
4	CL	A	902	1/1	0.97	0.10	52,52,52,52	0
3	ZN	A	901	1/1	0.98	0.07	95,95,95,95	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.