



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 6, 2025 – 04:23 PM JST

PDB ID : 8YDV
Title : Crystal structure of the receptor binding domain of SARS-CoV-2 Omicron BA.5 variant spike protein in complex with CeSPIACE
Authors : Nakamura, S.; Numoto, N.; Fujiyoshi, Y.
Deposited on : 2024-02-21
Resolution : 2.20 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) : 1.21
EDS : 3.0
Percentile statistics : 20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4 : 9.0.004 (Gargrove)
Density-Fitness : 1.0.11
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.40

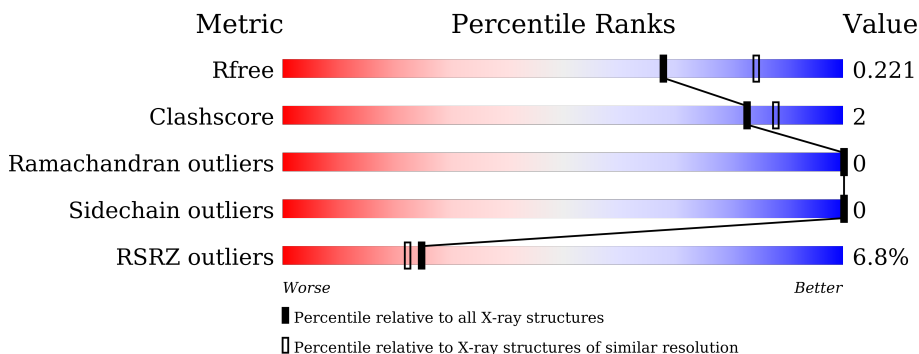
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.20 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}	164625	5791 (2.20-2.20)
Clashscore	180529	6634 (2.20-2.20)
Ramachandran outliers	177936	6560 (2.20-2.20)
Sidechain outliers	177891	6561 (2.20-2.20)
RSRZ outliers	164620	5791 (2.20-2.20)

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	39	 8% 92% 8%
1	C	39	 90% 10%
2	B	230	 10% 77% 7% 16%
2	D	230	 3% 82% 14%

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 4088 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 SARS-CoV-2 inhibiting peptide CeSPIACE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	39	330	212	52	63	3	0	0	0
1	C	39	330	212	52	63	3	0	0	0

- Molecule 2 is a protein called Spike protein S1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	194	1506	958	260	280	8	0	0	0
2	D	198	1571	1005	267	291	8	0	0	0

There are 104 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	339	ASP	GLY	variant	UNP P0DTC2
B	371	PHE	SER	variant	UNP P0DTC2
B	373	PRO	SER	variant	UNP P0DTC2
B	376	ALA	THR	variant	UNP P0DTC2
B	405	ASN	ASP	variant	UNP P0DTC2
B	408	SER	ARG	variant	UNP P0DTC2
B	417	ASN	LYS	variant	UNP P0DTC2
B	440	LYS	ASN	variant	UNP P0DTC2
B	452	ARG	LEU	variant	UNP P0DTC2
B	477	ASN	SER	variant	UNP P0DTC2
B	478	LYS	THR	variant	UNP P0DTC2
B	484	ALA	GLU	variant	UNP P0DTC2
B	486	VAL	PHE	variant	UNP P0DTC2
B	498	ARG	GLN	variant	UNP P0DTC2
B	501	TYR	ASN	variant	UNP P0DTC2
B	505	HIS	TYR	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	527	SER	-	expression tag	UNP P0DTC2
B	528	ASN	-	expression tag	UNP P0DTC2
B	529	SER	-	expression tag	UNP P0DTC2
B	530	GLU	-	expression tag	UNP P0DTC2
B	531	ASN	-	expression tag	UNP P0DTC2
B	532	LEU	-	expression tag	UNP P0DTC2
B	533	TYR	-	expression tag	UNP P0DTC2
B	534	PHE	-	expression tag	UNP P0DTC2
B	535	GLN	-	expression tag	UNP P0DTC2
B	536	GLY	-	expression tag	UNP P0DTC2
B	537	SER	-	expression tag	UNP P0DTC2
B	538	HIS	-	expression tag	UNP P0DTC2
B	539	HIS	-	expression tag	UNP P0DTC2
B	540	HIS	-	expression tag	UNP P0DTC2
B	541	HIS	-	expression tag	UNP P0DTC2
B	542	HIS	-	expression tag	UNP P0DTC2
B	543	HIS	-	expression tag	UNP P0DTC2
B	544	HIS	-	expression tag	UNP P0DTC2
B	545	HIS	-	expression tag	UNP P0DTC2
B	546	HIS	-	expression tag	UNP P0DTC2
B	547	HIS	-	expression tag	UNP P0DTC2
B	548	GLY	-	expression tag	UNP P0DTC2
B	549	LEU	-	expression tag	UNP P0DTC2
B	550	ASN	-	expression tag	UNP P0DTC2
B	551	ASP	-	expression tag	UNP P0DTC2
B	552	ILE	-	expression tag	UNP P0DTC2
B	553	PHE	-	expression tag	UNP P0DTC2
B	554	GLU	-	expression tag	UNP P0DTC2
B	555	ALA	-	expression tag	UNP P0DTC2
B	556	GLN	-	expression tag	UNP P0DTC2
B	557	LYS	-	expression tag	UNP P0DTC2
B	558	ILE	-	expression tag	UNP P0DTC2
B	559	GLU	-	expression tag	UNP P0DTC2
B	560	TRP	-	expression tag	UNP P0DTC2
B	561	HIS	-	expression tag	UNP P0DTC2
B	562	GLU	-	expression tag	UNP P0DTC2
D	339	ASP	GLY	variant	UNP P0DTC2
D	371	PHE	SER	variant	UNP P0DTC2
D	373	PRO	SER	variant	UNP P0DTC2
D	376	ALA	THR	variant	UNP P0DTC2
D	405	ASN	ASP	variant	UNP P0DTC2
D	408	SER	ARG	variant	UNP P0DTC2

Continued on next page...

Continued from previous page...

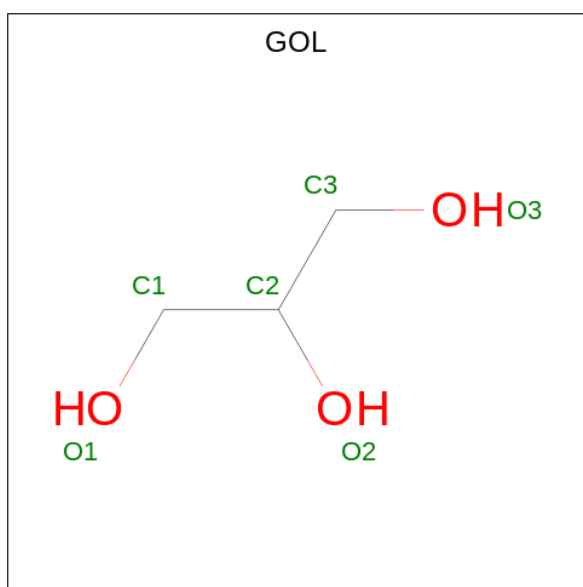
Chain	Residue	Modelled	Actual	Comment	Reference
D	417	ASN	LYS	variant	UNP P0DTC2
D	440	LYS	ASN	variant	UNP P0DTC2
D	452	ARG	LEU	variant	UNP P0DTC2
D	477	ASN	SER	variant	UNP P0DTC2
D	478	LYS	THR	variant	UNP P0DTC2
D	484	ALA	GLU	variant	UNP P0DTC2
D	486	VAL	PHE	variant	UNP P0DTC2
D	498	ARG	GLN	variant	UNP P0DTC2
D	501	TYR	ASN	variant	UNP P0DTC2
D	505	HIS	TYR	variant	UNP P0DTC2
D	527	SER	-	expression tag	UNP P0DTC2
D	528	ASN	-	expression tag	UNP P0DTC2
D	529	SER	-	expression tag	UNP P0DTC2
D	530	GLU	-	expression tag	UNP P0DTC2
D	531	ASN	-	expression tag	UNP P0DTC2
D	532	LEU	-	expression tag	UNP P0DTC2
D	533	TYR	-	expression tag	UNP P0DTC2
D	534	PHE	-	expression tag	UNP P0DTC2
D	535	GLN	-	expression tag	UNP P0DTC2
D	536	GLY	-	expression tag	UNP P0DTC2
D	537	SER	-	expression tag	UNP P0DTC2
D	538	HIS	-	expression tag	UNP P0DTC2
D	539	HIS	-	expression tag	UNP P0DTC2
D	540	HIS	-	expression tag	UNP P0DTC2
D	541	HIS	-	expression tag	UNP P0DTC2
D	542	HIS	-	expression tag	UNP P0DTC2
D	543	HIS	-	expression tag	UNP P0DTC2
D	544	HIS	-	expression tag	UNP P0DTC2
D	545	HIS	-	expression tag	UNP P0DTC2
D	546	HIS	-	expression tag	UNP P0DTC2
D	547	HIS	-	expression tag	UNP P0DTC2
D	548	GLY	-	expression tag	UNP P0DTC2
D	549	LEU	-	expression tag	UNP P0DTC2
D	550	ASN	-	expression tag	UNP P0DTC2
D	551	ASP	-	expression tag	UNP P0DTC2
D	552	ILE	-	expression tag	UNP P0DTC2
D	553	PHE	-	expression tag	UNP P0DTC2
D	554	GLU	-	expression tag	UNP P0DTC2
D	555	ALA	-	expression tag	UNP P0DTC2
D	556	GLN	-	expression tag	UNP P0DTC2
D	557	LYS	-	expression tag	UNP P0DTC2
D	558	ILE	-	expression tag	UNP P0DTC2

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	559	GLU	-	expression tag	UNP P0DTC2
D	560	TRP	-	expression tag	UNP P0DTC2
D	561	HIS	-	expression tag	UNP P0DTC2
D	562	GLU	-	expression tag	UNP P0DTC2

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	B	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total 31	O 31	0	0
4	B	121	Total 121	O 121	0	0
4	C	25	Total 25	O 25	0	0
4	D	120	Total 120	O 120	0	0

4 Data and refinement statistics i

Property	Value	Source
Space group	P 41	Depositor
Cell constants a, b, c, α , β , γ	87.48Å 87.48Å 118.23Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.98 – 2.20 48.98 – 2.20	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.98-2.20) 100.0 (48.98-2.20)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.24 (at 2.20Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.195 , 0.221 0.196 , 0.221	Depositor DCC
R_{free} test set	2255 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å ²)	43.6	Xtrriage
Anisotropy	0.313	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 39.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	0.041 for h,-k,-l	Xtrriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4088	wwPDB-VP
Average B, all atoms (Å ²)	51.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.74% 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: GOL

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	0.28	0/335	0.42	0/446
1	C	0.26	0/335	0.41	0/446
2	B	0.27	0/1545	0.50	0/2101
2	D	0.27	0/1616	0.48	0/2199
All	All	0.27	0/3831	0.48	0/5192

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 [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	330	0	330	3	0
1	C	330	0	330	4	0
2	B	1506	0	1401	8	0
2	D	1571	0	1488	5	0
3	B	12	0	16	0	0
3	C	12	0	16	0	0
3	D	30	0	40	0	0
4	A	31	0	0	0	0
4	B	121	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	C	25	0	0	1	0
4	D	120	0	0	0	0
All	All	4088	0	3621	17	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 2.

The worst 5 of 17 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:12:ILE:HG23	1:C:27:MET:HE2	1.72	0.70
2:B:388:ASN:HA	2:B:526:GLY:HA3	1.84	0.59
2:D:359:SER:HA	2:D:524:VAL:HG22	1.85	0.58
2:B:377:PHE:HB2	2:B:434:ILE:HG12	1.88	0.55
2:B:359:SER:HA	2:B:524:VAL:HG22	1.93	0.51

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	37/39 (95%)	37 (100%)	0	0	100	100
1	C	37/39 (95%)	36 (97%)	1 (3%)	0	100	100
2	B	190/230 (83%)	186 (98%)	4 (2%)	0	100	100
2	D	196/230 (85%)	193 (98%)	3 (2%)	0	100	100
All	All	460/538 (86%)	452 (98%)	8 (2%)	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	36/36 (100%)	36 (100%)	0	100	100
1	C	36/36 (100%)	36 (100%)	0	100	100
2	B	157/199 (79%)	157 (100%)	0	100	100
2	D	170/199 (85%)	170 (100%)	0	100	100
All	All	399/470 (85%)	399 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

9 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	GOL	B	601	-	5,5,5	1.02	0	5,5,5	0.94	0
3	GOL	C	101	-	5,5,5	0.93	0	5,5,5	0.95	0
3	GOL	B	602	-	5,5,5	0.85	0	5,5,5	1.06	0
3	GOL	D	602	-	5,5,5	0.89	0	5,5,5	1.00	0
3	GOL	D	605	-	5,5,5	0.92	0	5,5,5	0.95	0
3	GOL	D	601	-	5,5,5	0.76	0	5,5,5	1.02	0
3	GOL	C	102	-	5,5,5	0.80	0	5,5,5	1.07	0
3	GOL	D	603	-	5,5,5	0.91	0	5,5,5	0.97	0
3	GOL	D	604	-	5,5,5	0.86	0	5,5,5	1.02	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	GOL	B	601	-	-	2/4/4/4	-
3	GOL	C	101	-	-	1/4/4/4	-
3	GOL	B	602	-	-	0/4/4/4	-
3	GOL	D	602	-	-	0/4/4/4	-
3	GOL	D	605	-	-	0/4/4/4	-
3	GOL	D	601	-	-	0/4/4/4	-
3	GOL	C	102	-	-	2/4/4/4	-
3	GOL	D	603	-	-	2/4/4/4	-
3	GOL	D	604	-	-	4/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 11 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	C	102	GOL	C1-C2-C3-O3
3	D	603	GOL	O1-C1-C2-O2
3	D	603	GOL	O1-C1-C2-C3
3	B	601	GOL	C1-C2-C3-O3
3	D	604	GOL	O1-C1-C2-C3

There are no ring outliers.

No monomer is involved in short contacts.

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 [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	39/39 (100%)	0.07	3 (7%) 21 18	34, 42, 70, 98	0
1	C	39/39 (100%)	-0.15	0 100 100	33, 42, 65, 93	0
2	B	194/230 (84%)	0.55	22 (11%) 11 9	32, 47, 98, 109	0
2	D	198/230 (86%)	0.11	7 (3%) 47 44	32, 48, 79, 92	0
All	All	470/538 (87%)	0.27	32 (6%) 25 22	32, 46, 92, 109	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	371	PHE	6.6
2	B	368	LEU	5.3
2	B	369	TYR	5.0
2	D	372	ALA	4.8
2	D	371	PHE	4.0

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, 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	GOL	B	602	6/6	0.86	0.15	63,67,73,74	0
3	GOL	D	605	6/6	0.86	0.16	53,56,64,67	0
3	GOL	C	101	6/6	0.87	0.19	53,64,70,70	0
3	GOL	B	601	6/6	0.87	0.15	49,61,64,75	0
3	GOL	D	602	6/6	0.88	0.12	52,55,65,66	0
3	GOL	D	604	6/6	0.89	0.15	55,67,69,69	0
3	GOL	D	601	6/6	0.90	0.14	42,49,54,55	0
3	GOL	D	603	6/6	0.90	0.12	63,69,73,78	0
3	GOL	C	102	6/6	0.91	0.15	39,42,50,57	0

6.5 Other polymers [i](#)

There are no such residues in this entry.