



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 6, 2020 – 10:42 AM BST

PDB ID : 6QZC
Title : HLA-DR1 with the QAR Peptide
Authors : Greenshields-Watson, A.; Rizkallah, P.J.
Deposited on : 2019-03-11
Resolution : 1.64 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

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.13.1
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) : 2.13.1

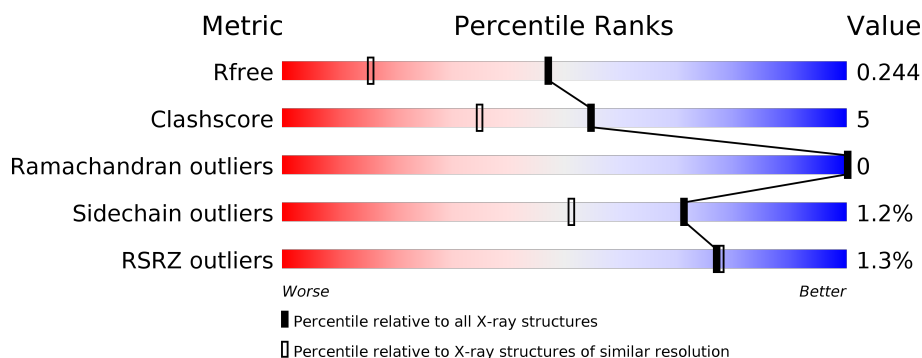
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.64 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	3122 (1.66-1.62)
Clashscore	141614	3268 (1.66-1.62)
Ramachandran outliers	138981	3215 (1.66-1.62)
Sidechain outliers	138945	3215 (1.66-1.62)
RSRZ outliers	127900	3079 (1.66-1.62)

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 ≥ 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	AAA	180	<div> <div>2%</div> <div> <div></div> <div>92%</div> <div>8%</div> </div> </div>
2	BBB	191	<div> <div></div> <div> <div>87%</div> <div>12%</div> </div> </div>
3	CCC	16	<div> <div>13%</div> <div> <div></div> <div>88%</div> <div>13%</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
7	SO4	BBB	203	-	-	-	X

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 3395 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	180	Total	C	N	O	S	0	1	0
			1475	955	238	276	6			

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	191	Total	C	N	O	S	0	0	0
			1540	968	272	293	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	0	MET	-	initiating methionine	UNP P04229

- Molecule 3 is a protein called M1-208-222-QAR-Peptide.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	CCC	16	Total	C	N	O	S	0	0	0
			126	75	27	22	2			

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C₄H₁₀O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	AAA	1	Total	C	O	0	0
			7	4	3		

- Molecule 5 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	BBB	1	Total	C	O	0	0
			4	2	2		

Continued on next page...

Continued from previous page...

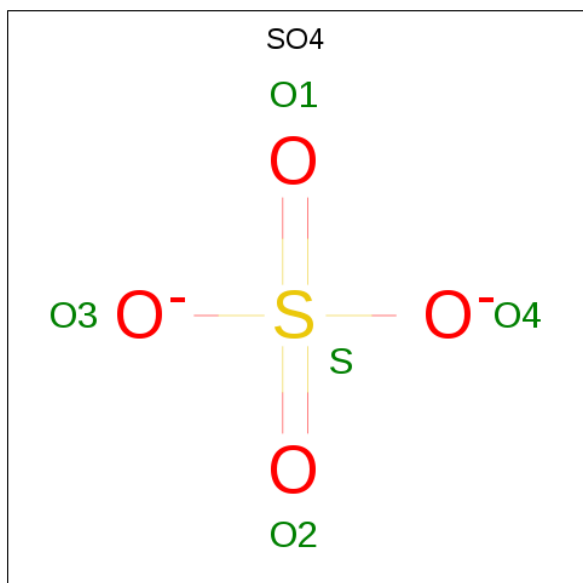
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	BBB	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 7 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	BBB	1	Total	O	S	0	0
			5	4	1		

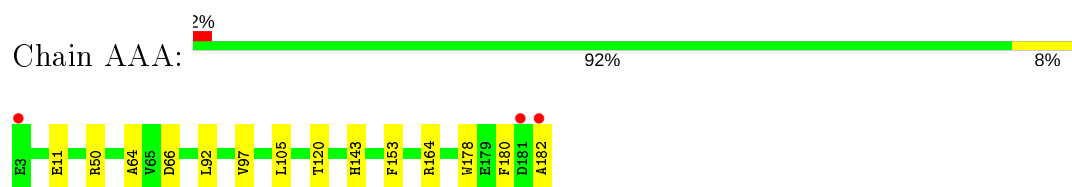
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	AAA	120	Total	O	0	0
			120	120		
8	BBB	89	Total	O	0	0
			89	89		
8	CCC	11	Total	O	0	0
			11	11		

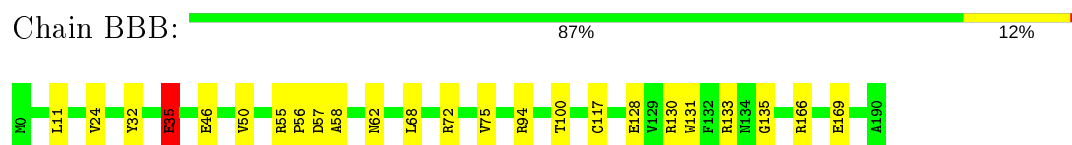
3 Residue-property plots [i](#)

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.

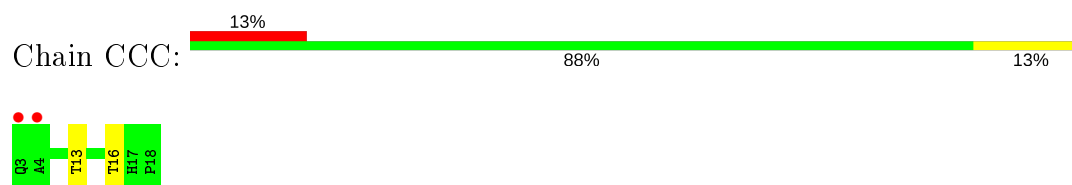
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain



- Molecule 3: M1-208-222-QAR-Peptide



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2	Depositor
Cell constants a, b, c, α , β , γ	90.06 Å 116.08 Å 43.68 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	71.16 – 1.64 71.16 – 1.64	Depositor EDS
% Data completeness (in resolution range)	99.8 (71.16-1.64) 99.8 (71.16-1.64)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.42 (at 1.64 Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.202 , 0.240 0.208 , 0.244	Depositor DCC
R_{free} test set	2819 reflections (4.95%)	wwPDB-VP
Wilson B-factor (Å ²)	25.9	Xtriage
Anisotropy	0.685	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 37.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	3395	wwPDB-VP
Average B, all atoms (Å ²)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.96% 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, CSO, PEG, EDO, SO4

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	AAA	0.77	0/1520	0.90	1/2073 (0.0%)
2	BBB	0.80	3/1571 (0.2%)	0.89	0/2134
3	CCC	0.75	0/127	0.88	0/168
All	All	0.79	3/3218 (0.1%)	0.89	1/4375 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BBB	46	GLU	CD-OE2	7.08	1.33	1.25
2	BBB	46	GLU	CG-CD	5.92	1.60	1.51
2	BBB	35	GLU	CD-OE1	5.66	1.31	1.25

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	AAA	50	ARG	NE-CZ-NH2	-5.30	117.65	120.30

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	AAA	1475	0	1400	10	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	BBB	1540	0	1448	21	1
3	CCC	126	0	128	6	0
4	AAA	7	0	10	2	0
5	AAA	8	0	12	2	0
5	BBB	8	0	12	3	0
6	AAA	6	0	8	0	0
7	BBB	5	0	0	0	0
8	AAA	120	0	0	0	0
8	BBB	89	0	0	4	0
8	CCC	11	0	0	1	0
All	All	3395	0	3018	32	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 5.

All (32) 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:BBB:166:ARG:NH2	2:BBB:169:GLU:OE2	2.09	0.84
2:BBB:94:ARG:HD2	5:BBB:201:EDO:H11	1.61	0.81
2:BBB:94:ARG:HD2	5:BBB:201:EDO:C1	2.16	0.75
2:BBB:11:LEU:HD11	8:BBB:340:HOH:O	1.88	0.74
2:BBB:128:GLU:OE1	2:BBB:130:ARG:NH2	2.33	0.60
1:AAA:182:ALA:HB2	8:BBB:387:HOH:O	2.01	0.58
1:AAA:120:THR:HG21	4:AAA:201:PEG:O1	2.05	0.57
2:BBB:11:LEU:HD12	3:CCC:13:THR:HG21	1.87	0.57
1:AAA:143:HIS:HE1	8:BBB:378:HOH:O	1.88	0.56
1:AAA:97:VAL:HG21	1:AAA:180:PHE:CD1	2.41	0.55
1:AAA:64:ALA:HB1	5:AAA:202:EDO:H21	1.89	0.54
2:BBB:11:LEU:HD12	3:CCC:13:THR:CG2	2.38	0.53
1:AAA:11:GLU:OE1	1:AAA:66:ASP:OD2	2.27	0.52
2:BBB:128:GLU:OE1	2:BBB:130:ARG:NH1	2.45	0.50
2:BBB:133:ARG:HD3	2:BBB:169:GLU:OE2	2.14	0.48
1:AAA:105:LEU:HG	1:AAA:153:PHE:CE1	2.51	0.46
2:BBB:117:CYS:HB2	2:BBB:131:TRP:CZ2	2.51	0.46
2:BBB:55:ARG:HB2	2:BBB:56:PRO:HD3	1.97	0.45
2:BBB:11:LEU:CD1	3:CCC:13:THR:HG22	2.47	0.45
2:BBB:94:ARG:HD2	5:BBB:201:EDO:H12	1.94	0.45
3:CCC:16:THR:O	8:CCC:101:HOH:O	2.21	0.44
2:BBB:24:VAL:HG12	2:BBB:75:VAL:HG13	2.00	0.44
2:BBB:24:VAL:CG1	2:BBB:75:VAL:HG13	2.49	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:BBB:11:LEU:CD1	3:CCC:13:THR:CG2	2.97	0.43
2:BBB:68:LEU:O	2:BBB:72:ARG:HD2	2.18	0.43
2:BBB:32:TYR:O	2:BBB:35:GLU:HG2	2.19	0.42
1:AAA:164:ARG:NE	4:AAA:201:PEG:H21	2.34	0.42
2:BBB:72:ARG:NH1	8:BBB:307:HOH:O	2.54	0.41
1:AAA:97:VAL:HG11	1:AAA:178:TRP:HZ2	1.86	0.41
2:BBB:58:ALA:O	2:BBB:62:ASN:HB2	2.21	0.41
2:BBB:57:ASP:OD1	3:CCC:16:THR:HB	2.20	0.41
1:AAA:64:ALA:CB	5:AAA:202:EDO:H21	2.51	0.40

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)
2:BBB:135:GLY:O	2:BBB:135:GLY:O[2_665]	2.00	0.20

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	AAA	179/180 (99%)	178 (99%)	1 (1%)	0	100	100
2	BBB	188/191 (98%)	183 (97%)	5 (3%)	0	100	100
3	CCC	14/16 (88%)	14 (100%)	0	0	100	100
All	All	381/387 (98%)	375 (98%)	6 (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	AAA	163/164 (99%)	162 (99%)	1 (1%)	86	75
2	BBB	165/171 (96%)	162 (98%)	3 (2%)	59	34
3	CCC	13/13 (100%)	13 (100%)	0	100	100
All	All	341/348 (98%)	337 (99%)	4 (1%)	71	51

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	AAA	92	LEU
2	BBB	35	GLU
2	BBB	50	VAL
2	BBB	100	THR

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

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	CSO	BBB	30	2	3,6,7	0.72	0	0,6,8	0.00	-

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	CSO	BBB	30	2	-	0/1/5/7	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

7 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	$\# Z > 2$
5	EDO	BBB	201	-	3,3,3	0.08	0	2,2,2	0.68	0
7	SO4	BBB	203	-	4,4,4	0.38	0	6,6,6	0.06	0
6	GOL	AAA	204	-	5,5,5	0.10	0	5,5,5	0.32	0
4	PEG	AAA	201	-	6,6,6	0.19	0	5,5,5	0.25	0
5	EDO	AAA	203	-	3,3,3	0.18	0	2,2,2	0.30	0
5	EDO	AAA	202	-	3,3,3	0.28	0	2,2,2	0.40	0
5	EDO	BBB	202	-	3,3,3	0.19	0	2,2,2	0.36	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
5	EDO	BBB	201	-	-	1/1/1/1	-
6	GOL	AAA	204	-	-	2/4/4/4	-
4	PEG	AAA	201	-	-	3/4/4/4	-
5	EDO	AAA	203	-	-	0/1/1/1	-
5	EDO	AAA	202	-	-	1/1/1/1	-
5	EDO	BBB	202	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (7) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
6	AAA	204	GOL	C1-C2-C3-O3
6	AAA	204	GOL	O2-C2-C3-O3
4	AAA	201	PEG	O1-C1-C2-O2
4	AAA	201	PEG	O2-C3-C4-O4
5	AAA	202	EDO	O1-C1-C2-O2
5	BBB	201	EDO	O1-C1-C2-O2
4	AAA	201	PEG	C4-C3-O2-C2

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	BBB	201	EDO	3	0
4	AAA	201	PEG	2	0
5	AAA	202	EDO	2	0

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 [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	AAA	180/180 (100%)	-0.28	3 (1%) 70 71	22, 30, 51, 75	0
2	BBB	190/191 (99%)	-0.08	0 100 100	22, 44, 66, 72	0
3	CCC	16/16 (100%)	0.55	2 (12%) 3 3	30, 39, 73, 79	0
All	All	386/387 (99%)	-0.15	5 (1%) 77 78	22, 37, 64, 79	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	182	ALA	6.0
3	CCC	4	ALA	4.5
3	CCC	3	GLN	4.0
1	AAA	3	GLU	2.4
1	AAA	181	ASP	2.3

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	CSO	BBB	30	7/8	0.96	0.06	31,33,44,51	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(Å ²)	Q<0.9
5	EDO	AAA	203	4/4	0.29	0.35	61,73,73,73	0
7	SO4	BBB	203	5/5	0.53	0.53	158,158,160,164	0
4	PEG	AAA	201	7/7	0.76	0.27	49,53,57,57	0
6	GOL	AAA	204	6/6	0.77	0.19	51,54,57,61	0
5	EDO	BBB	201	4/4	0.84	0.17	41,42,43,49	0
5	EDO	BBB	202	4/4	0.87	0.20	36,42,45,46	0
5	EDO	AAA	202	4/4	0.95	0.19	36,38,39,39	0

6.5 Other polymers [i](#)

There are no such residues in this entry.