



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 13, 2024 – 05:05 pm GMT

PDB ID : 6QZD
Title : HLA-DR1 with SGP Influenza Matrix Peptide
Authors : Greenshields-Watson, A.; Rizkallah, P.J.
Deposited on : 2019-03-11
Resolution : 2.66 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 : 1.8.4, 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

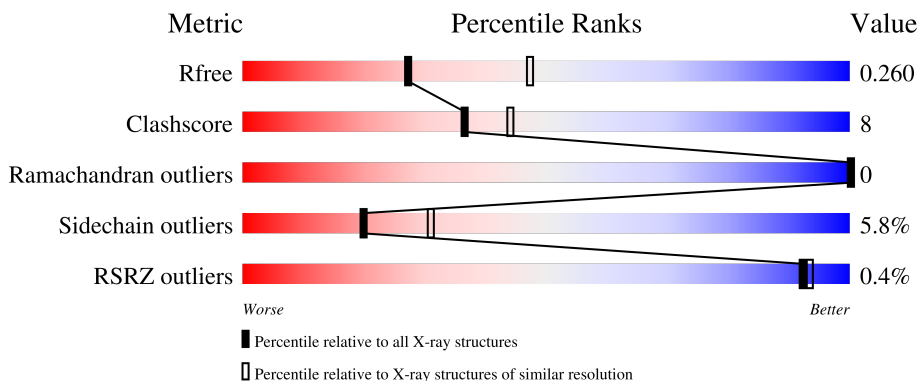
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


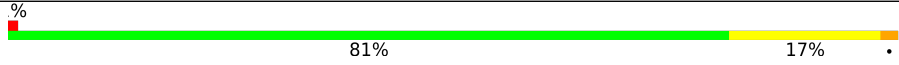
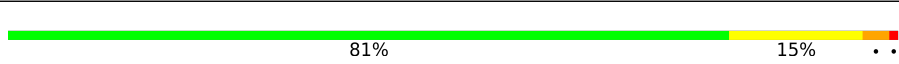
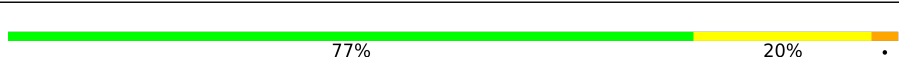
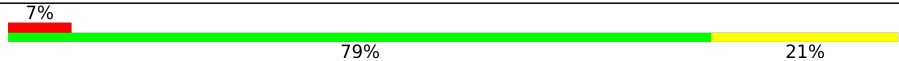
The reported resolution of this entry is 2.66 Å.

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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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	AAA	180	
1	DDD	180	
2	BBB	191	
2	EEE	191	
3	CCC	14	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	FFF	14	 64% 36%

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
6	GOL	AAA	204	-	-	-	X

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 6366 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
			Total	C	N	O	S			
1	AAA	180	1465	950	237	273	5	0	0	0
1	DDD	180	1469	952	237	275	5	0	0	0

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	191	1555	979	277	292	7	0	0	0
2	EEE	191	1559	981	277	294	7	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 3 is a protein called Matrix protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	CCC	14	107	65	19	23	0	0	0
3	FFF	14	107	65	19	23	0	0	0

- 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		
4	DDD	1	Total	C	O	0	0
			7	4	3		

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



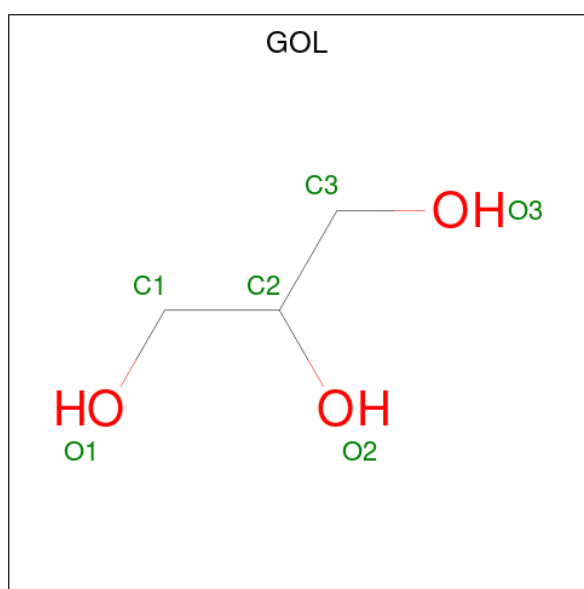
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	AAA	1	Total	C	O	0	0
			4	2	2		
5	AAA	1	Total	C	O	0	0
			4	2	2		

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		
5	BBB	1	Total	C	O	0	0
			4	2	2		
5	DDD	1	Total	C	O	0	0
			4	2	2		
5	EEE	1	Total	C	O	0	0
			4	2	2		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



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: O₄S).



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

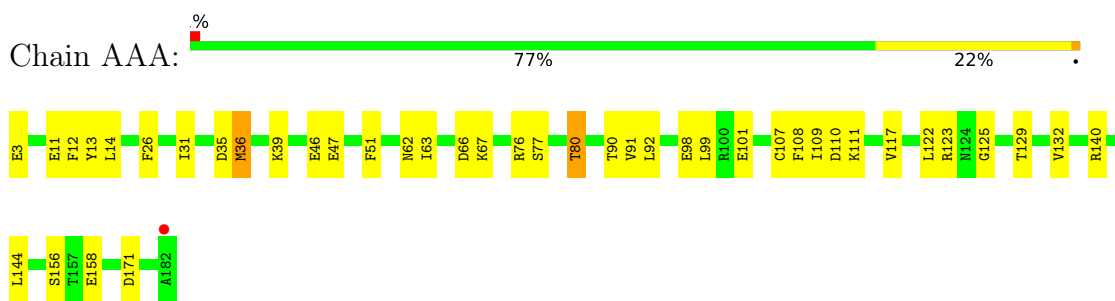
- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	AAA	12	Total	O	0	0
			12	12		
8	BBB	12	Total	O	0	0
			12	12		
8	DDD	14	Total	O	0	0
			14	14		
8	EEE	17	Total	O	0	0
			17	17		

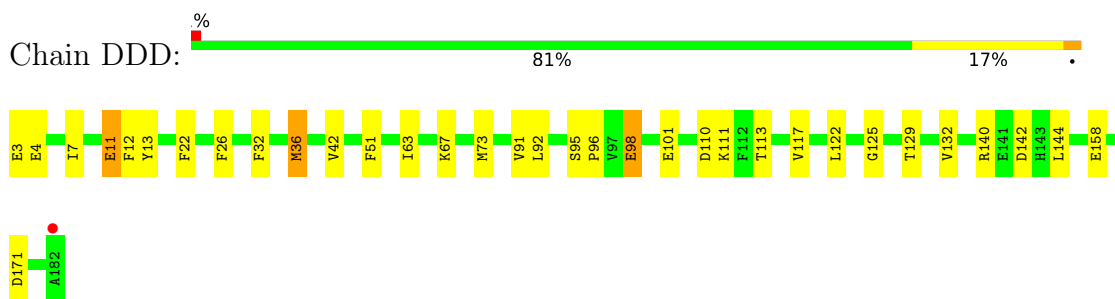
3 Residue-property plots [i](#)

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.

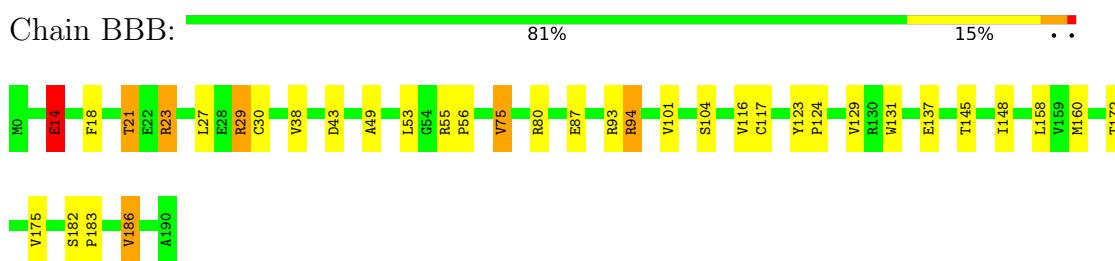
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



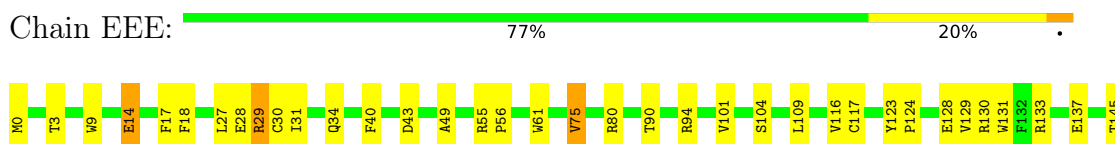
- Molecule 1: HLA class II histocompatibility antigen, DR alpha chain



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

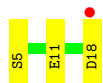
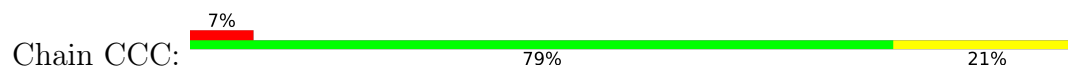


- Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

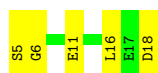




- Molecule 3: Matrix protein 1



- Molecule 3: Matrix protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	67.42Å 76.73Å 90.13Å 90.00° 93.50° 90.00°	Depositor
Resolution (Å)	55.54 – 2.66 55.54 – 2.66	Depositor EDS
% Data completeness (in resolution range)	99.0 (55.54-2.66) 99.0 (55.54-2.66)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.73 (at 2.65Å)	Xtrriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.200 , 0.254 0.202 , 0.260	Depositor DCC
R_{free} test set	1267 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å ²)	43.2	Xtrriage
Anisotropy	0.464	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 30.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	6366	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.59% 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: SO4, GOL, PEG, EDO

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	0.46	1/1510 (0.1%)	0.84	3/2060 (0.1%)
1	DDD	0.42	0/1514	0.79	0/2065
2	BBB	1.02	3/1595 (0.2%)	0.94	3/2166 (0.1%)
2	EEE	0.75	1/1599 (0.1%)	0.90	4/2171 (0.2%)
3	CCC	0.52	0/107	0.80	0/141
3	FFF	0.59	0/107	0.83	0/141
All	All	0.71	5/6432 (0.1%)	0.87	10/8744 (0.1%)

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	BBB	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	BBB	14	GLU	CD-OE2	28.74	1.57	1.25
2	EEE	14	GLU	CD-OE2	22.16	1.50	1.25
2	BBB	14	GLU	CD-OE1	20.09	1.47	1.25
2	BBB	29	ARG	CZ-NH2	8.67	1.44	1.33
1	AAA	123	ARG	CD-NE	5.45	1.55	1.46

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	BBB	14	GLU	OE1-CD-OE2	-13.80	106.73	123.30
2	EEE	29	ARG	NE-CZ-NH2	-8.54	116.03	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	EEE	94	ARG	CG-CD-NE	-6.58	97.99	111.80
2	EEE	3	THR	CA-CB-OG1	-6.24	95.90	109.00
1	AAA	123	ARG	CB-CG-CD	5.99	127.18	111.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	BBB	14	GLU	Sidechain

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	1465	0	1393	29	0
1	DDD	1469	0	1397	26	0
2	BBB	1555	0	1482	24	0
2	EEE	1559	0	1486	29	0
3	CCC	107	0	107	1	0
3	FFF	107	0	107	3	0
4	AAA	7	0	10	0	0
4	DDD	7	0	10	0	0
5	AAA	8	0	12	0	0
5	BBB	8	0	12	0	0
5	DDD	4	0	6	0	0
5	EEE	4	0	6	0	0
6	AAA	6	0	8	0	0
7	DDD	5	0	0	0	0
8	AAA	12	0	0	0	0
8	BBB	12	0	0	0	0
8	DDD	14	0	0	1	0
8	EEE	17	0	0	0	0
All	All	6366	0	6036	97	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 8.

The worst 5 of 97 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:AAA:3:GLU:HB3	2:BBB:18:PHE:HD1	1.07	1.14
1:AAA:3:GLU:HB3	2:BBB:18:PHE:CD1	1.98	0.98
1:AAA:98:GLU:HB2	1:AAA:101:GLU:HG3	1.48	0.95
2:EEE:14:GLU:OE2	2:EEE:29:ARG:NE	2.00	0.94
1:AAA:3:GLU:CB	2:BBB:18:PHE:HD1	1.84	0.91

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	AAA	178/180 (99%)	173 (97%)	5 (3%)	0	100	100
1	DDD	178/180 (99%)	172 (97%)	6 (3%)	0	100	100
2	BBB	189/191 (99%)	179 (95%)	10 (5%)	0	100	100
2	EEE	189/191 (99%)	182 (96%)	7 (4%)	0	100	100
3	CCC	12/14 (86%)	11 (92%)	1 (8%)	0	100	100
3	FFF	12/14 (86%)	11 (92%)	1 (8%)	0	100	100
All	All	758/770 (98%)	728 (96%)	30 (4%)	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	161/164 (98%)	153 (95%)	8 (5%)	24	38
1	DDD	162/164 (99%)	155 (96%)	7 (4%)	29	44
2	BBB	170/172 (99%)	159 (94%)	11 (6%)	17	26
2	EEE	171/172 (99%)	161 (94%)	10 (6%)	20	31
3	CCC	11/11 (100%)	9 (82%)	2 (18%)	1	1
3	FFF	11/11 (100%)	9 (82%)	2 (18%)	1	1
All	All	686/694 (99%)	646 (94%)	40 (6%)	20	31

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	EEE	0	MET
2	EEE	163	THR
2	EEE	27	LEU
2	EEE	75	VAL
2	EEE	186	VAL

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 [i](#)

10 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
4	PEG	DDD	202	-	6,6,6	0.23	0	5,5,5	0.24	0
5	EDO	EEE	201	-	3,3,3	0.08	0	2,2,2	0.19	0
6	GOL	AAA	204	-	5,5,5	0.10	0	5,5,5	0.40	0
7	SO4	DDD	201	-	4,4,4	0.37	0	6,6,6	0.06	0
5	EDO	BBB	202	-	3,3,3	0.09	0	2,2,2	0.11	0
5	EDO	AAA	203	-	3,3,3	0.06	0	2,2,2	0.12	0
4	PEG	AAA	201	-	6,6,6	0.22	0	5,5,5	0.18	0
5	EDO	AAA	202	-	3,3,3	0.11	0	2,2,2	0.13	0
5	EDO	BBB	201	-	3,3,3	0.12	0	2,2,2	0.34	0
5	EDO	DDD	203	-	3,3,3	0.13	0	2,2,2	0.26	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
4	PEG	DDD	202	-	-	2/4/4/4	-
5	EDO	EEE	201	-	-	1/1/1/1	-
6	GOL	AAA	204	-	-	0/4/4/4	-
5	EDO	BBB	202	-	-	0/1/1/1	-
5	EDO	AAA	203	-	-	0/1/1/1	-
4	PEG	AAA	201	-	-	2/4/4/4	-
5	EDO	AAA	202	-	-	1/1/1/1	-
5	EDO	BBB	201	-	-	1/1/1/1	-
5	EDO	DDD	203	-	-	1/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	DDD	202	PEG	O2-C3-C4-O4
4	DDD	202	PEG	O1-C1-C2-O2
5	AAA	202	EDO	O1-C1-C2-O2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
5	EEE	201	EDO	O1-C1-C2-O2
4	AAA	201	PEG	O2-C3-C4-O4

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 [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%)	-0.23	1 (0%) 89 89	27, 41, 59, 105	0
1	DDD	180/180 (100%)	-0.24	1 (0%) 89 89	26, 41, 67, 88	0
2	BBB	191/191 (100%)	-0.38	0 100 100	25, 40, 59, 74	0
2	EEE	191/191 (100%)	-0.36	0 100 100	23, 39, 58, 85	0
3	CCC	14/14 (100%)	0.13	1 (7%) 16 12	36, 48, 74, 86	0
3	FFF	14/14 (100%)	0.12	0 100 100	36, 46, 81, 88	0
All	All	770/770 (100%)	-0.29	3 (0%) 92 93	23, 40, 62, 105	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	AAA	182	ALA	7.3
1	DDD	182	ALA	4.4
3	CCC	18	ASP	2.8

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(Å ²)	Q<0.9
6	GOL	AAA	204	6/6	0.72	0.47	83,90,98,99	0
7	SO4	DDD	201	5/5	0.80	0.40	120,120,125,129	0
4	PEG	AAA	201	7/7	0.83	0.20	45,50,62,65	0
4	PEG	DDD	202	7/7	0.85	0.23	55,56,58,60	0
5	EDO	AAA	202	4/4	0.87	0.31	54,55,55,62	0
5	EDO	DDD	203	4/4	0.90	0.26	57,64,68,73	0
5	EDO	BBB	202	4/4	0.91	0.17	47,51,53,53	0
5	EDO	BBB	201	4/4	0.92	0.17	40,42,45,50	0
5	EDO	AAA	203	4/4	0.93	0.30	49,50,50,52	0
5	EDO	EEE	201	4/4	0.96	0.15	43,45,47,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.