wwPDB X-ray Structure Validation Summary Report (i)

May 27, 2024 – 10:12 am BST

PDB ID	:	8PWW
Title	:	PfRH5 bound to monoclonal antibody MAD8-151
Authors	:	Farrell, B.; Higgins, M.K.
Deposited on	:	2023-07-21
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 (i) 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 (i)) were used in the production of this report:

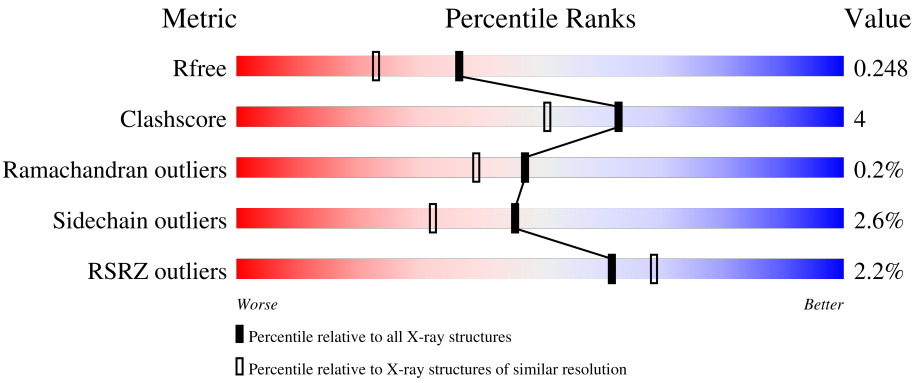
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.2

1 Overall quality at a glance (i)

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	4310 (1.96-1.92)
Clashscore	141614	1023 (1.94-1.94)
Ramachandran outliers	138981	1007 (1.94-1.94)
Sidechain outliers	138945	1007 (1.94-1.94)
RSRZ outliers	127900	4250 (1.96-1.92)

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 <=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	338	78% 11% 12%
1	C	338	4% 72% 11% 16%
2	B	250	2% 80% 10% 9%
2	D	250	78% 12% 10%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8983 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 Reticulocyte-binding protein homolog 5.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	299	Total	C	N	O	S	0	0	0
			2535	1632	429	459	15			
1	C	283	Total	C	N	O	S	0	1	0
			2403	1549	405	436	13			

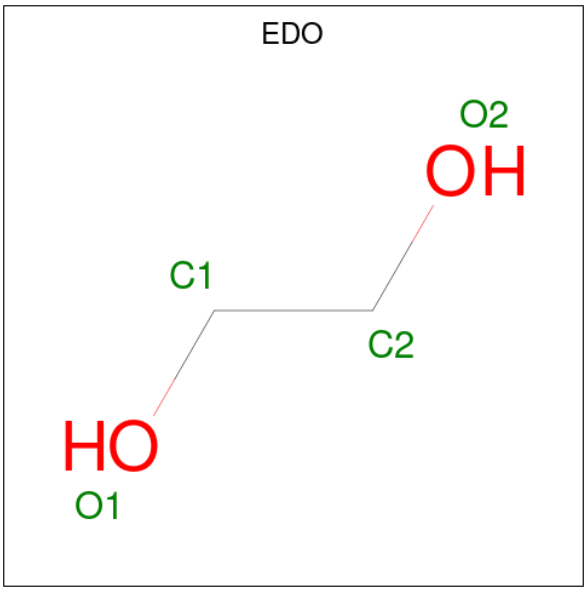
There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	203	TYR	CYS	conflict	UNP Q8IFM5
A	216	ALA	THR	conflict	UNP Q8IFM5
A	250	ALA	THR	conflict	UNP Q8IFM5
C	203	TYR	CYS	conflict	UNP Q8IFM5
C	216	ALA	THR	conflict	UNP Q8IFM5
C	250	ALA	THR	conflict	UNP Q8IFM5

- Molecule 2 is a protein called scFv fragment for antibody MAD8-151.

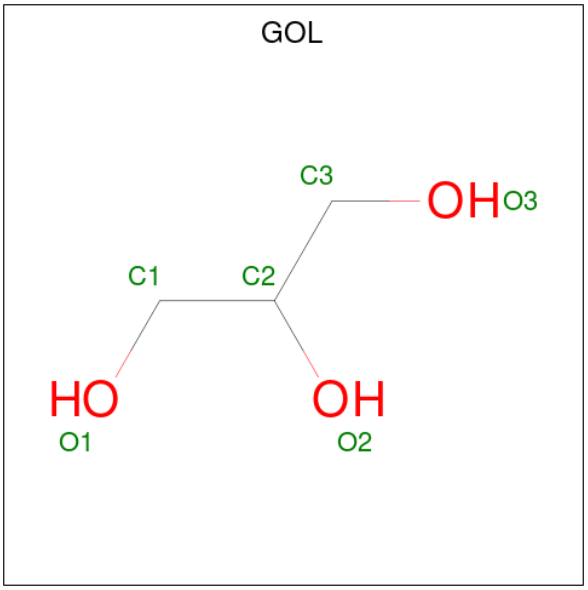
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	227	Total	C	N	O	S	0	1	0
			1738	1097	286	348	7			
2	D	225	Total	C	N	O	S	0	1	0
			1734	1095	285	347	7			

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	C	1	Total C O 4 2 2	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



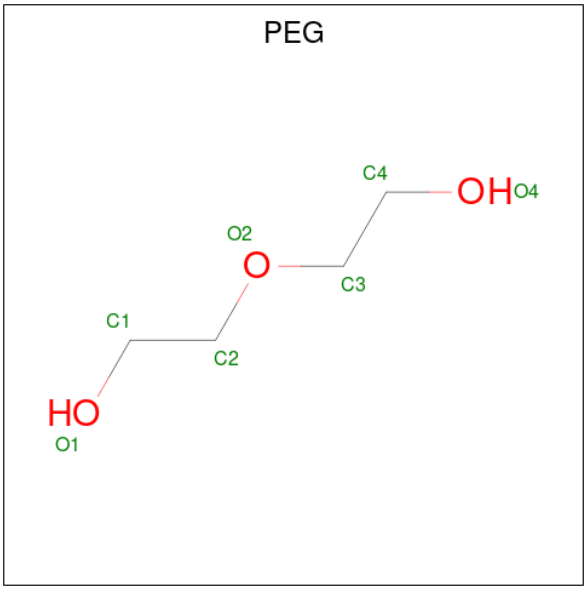
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 6 3 3	0	0
4	B	1	Total C O 6 3 3	0	0

Continued on next page...

Continued from previous page...

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

- Molecule 5 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).



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

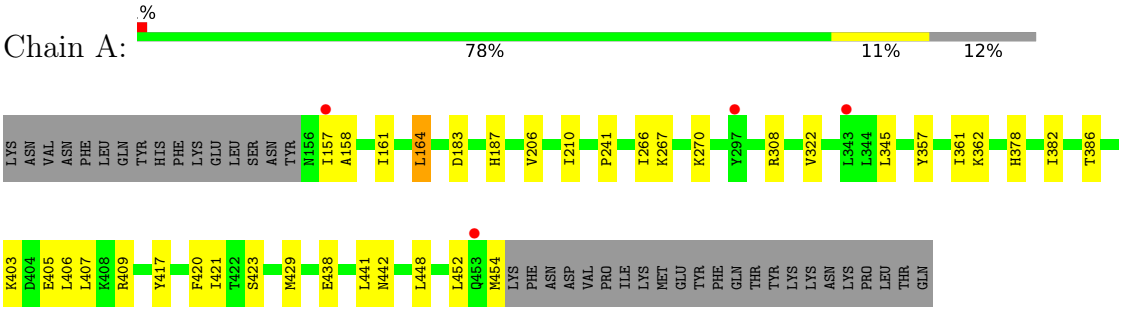
- Molecule 6 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	145	Total O 145 145	0	0
6	B	162	Total O 162 162	0	0
6	C	123	Total O 123 123	0	0
6	D	103	Total O 103 103	0	0

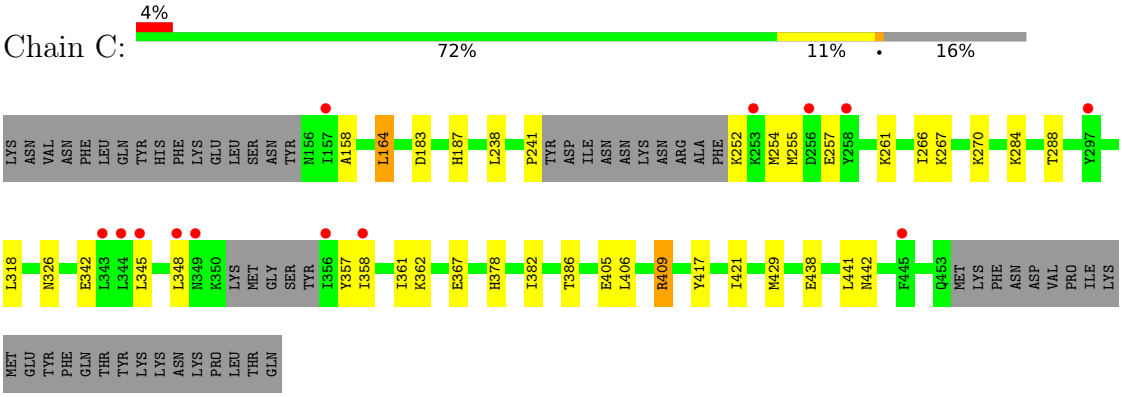
3 Residue-property plots (i)

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.

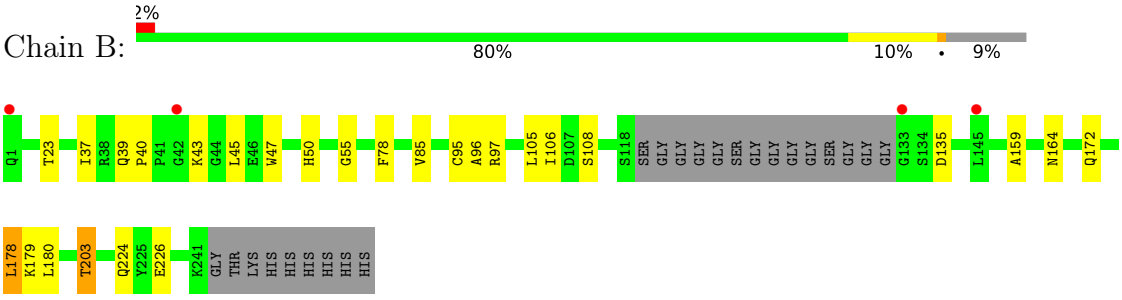
• Molecule 1: Reticulocyte-binding protein homolog 5



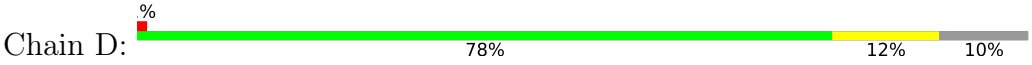
• Molecule 1: Reticulocyte-binding protein homolog 5

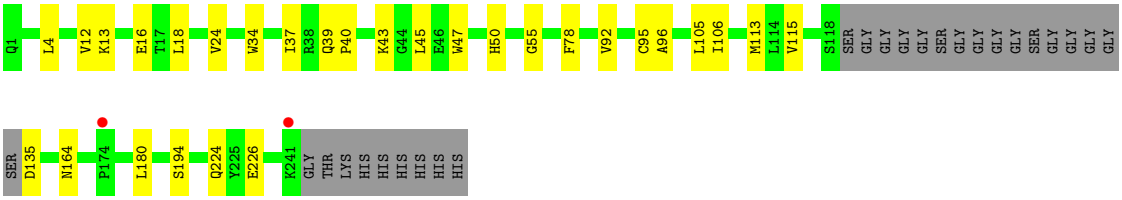


• Molecule 2: scFv fragment for antibody MAD8-151



• Molecule 2: scFv fragment for antibody MAD8-151




Q1
L4
V12
K13
E16
T17
L18
V24
W34
I37
R38
Q39
P40
K43
G44
L45
E46
W47
H50
G55
F78
V92
C95
A96
L105
I106
M113
L114
V115
S118
SER
GLY
GLY
GLY
GLY
SER
GLY
GLY
GLY
GLY
SER
GLY
GLY
GLY
GLY

SER
D135
N164
P174
L180
S194
Q224
Y225
E226
R241
GLY
THR
LYS
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	78.30Å 121.38Å 79.00Å 90.00° 90.18° 90.00°	Depositor
Resolution (Å)	79.00 – 1.95 79.00 – 1.94	Depositor EDS
% Data completeness (in resolution range)	97.5 (79.00-1.95) 96.6 (79.00-1.94)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.94Å)	Xtriage
Refinement program	BUSTER 2.10.4 (20-OCT-2021)	Depositor
R, R_{free}	0.230 , 0.252 0.223 , 0.248	Depositor DCC
R_{free} test set	5117 reflections (4.85%)	wwPDB-VP
Wilson B-factor (Å^2)	27.9	Xtriage
Anisotropy	0.456	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 30.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.014 for l,k,-h 0.257 for h,-k,-l 0.025 for l,-k,h	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8983	wwPDB-VP
Average B, all atoms (Å^2)	32.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.23% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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: EDO, GOL, PEG

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	0.40	0/2588	0.51	0/3473
1	C	0.37	0/2454	0.51	0/3292
2	B	0.43	0/1779	0.64	0/2420
2	D	0.44	0/1772	0.65	0/2411
All	All	0.41	0/8593	0.57	0/11596

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	2535	0	2557	21	0
1	C	2403	0	2430	22	0
2	B	1738	0	1695	12	0
2	D	1734	0	1687	17	0
3	A	4	0	6	0	0
3	C	4	0	6	0	0
4	A	6	0	8	0	0
4	B	6	0	8	0	0
4	D	6	0	8	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	A	7	0	10	0	0
5	C	7	0	10	0	0
6	A	145	0	0	0	0
6	B	162	0	0	1	0
6	C	123	0	0	1	0
6	D	103	0	0	0	0
All	All	8983	0	8425	72	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.

The worst 5 of 72 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:C:238:LEU:HD21	1:C:254:MET:HB3	1.63	0.80
1:A:382:ILE:O	1:A:386:THR:HG23	1.88	0.73
1:C:257:GLU:O	1:C:261:LYS:HG2	1.89	0.73
2:B:23:THR:HG23	6:B:412:HOH:O	1.95	0.66
1:C:348:LEU:HD22	1:C:358:ILE:HD13	1.78	0.65

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	297/338 (88%)	292 (98%)	5 (2%)	0	100	100
1	C	278/338 (82%)	276 (99%)	2 (1%)	0	100	100
2	B	224/250 (90%)	216 (96%)	7 (3%)	1 (0%)	34	24
2	D	222/250 (89%)	214 (96%)	7 (3%)	1 (0%)	29	17
All	All	1021/1176 (87%)	998 (98%)	21 (2%)	2 (0%)	47	39

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	55	GLY
2	D	55	GLY

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	288/327 (88%)	282 (98%)	6 (2%)	53	41
1	C	275/327 (84%)	268 (98%)	7 (2%)	47	35
2	B	198/208 (95%)	191 (96%)	7 (4%)	36	21
2	D	197/208 (95%)	192 (98%)	5 (2%)	47	35
All	All	958/1070 (90%)	933 (97%)	25 (3%)	46	32

5 of 25 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	326	ASN
1	C	409	ARG
2	D	224	GLN
1	C	406	LEU
1	C	438	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	77	GLN
1	C	326	ASN

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)

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	$\#\|Z\| > 2$
3	EDO	C	502	-	3,3,3	0.61	0	2,2,2	0.16	0
4	GOL	D	301	-	5,5,5	0.10	0	5,5,5	0.63	0
5	PEG	C	501	-	6,6,6	0.18	0	5,5,5	0.14	0
4	GOL	A	502	-	5,5,5	0.06	0	5,5,5	0.33	0
5	PEG	A	503	-	6,6,6	0.19	0	5,5,5	0.09	0
3	EDO	A	501	-	3,3,3	0.55	0	2,2,2	0.26	0
4	GOL	B	301	-	5,5,5	0.14	0	5,5,5	0.38	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	EDO	C	502	-	-	0/1/1/1	-
4	GOL	D	301	-	-	3/4/4/4	-
5	PEG	C	501	-	-	1/4/4/4	-
4	GOL	A	502	-	-	0/4/4/4	-
5	PEG	A	503	-	-	3/4/4/4	-
3	EDO	A	501	-	-	0/1/1/1	-
4	GOL	B	301	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 9 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	301	GOL	C1-C2-C3-O3
4	B	301	GOL	O2-C2-C3-O3
4	D	301	GOL	O1-C1-C2-O2
5	A	503	PEG	C1-C2-O2-C3
5	A	503	PEG	C4-C3-O2-C2

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^{th} 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	299/338 (88%)	0.24	4 (1%) 77 81	20, 33, 50, 70	0
1	C	283/338 (83%)	0.34	13 (4%) 32 39	21, 33, 56, 77	0
2	B	227/250 (90%)	0.23	4 (1%) 68 74	19, 28, 42, 57	0
2	D	225/250 (90%)	0.15	2 (0%) 84 87	20, 27, 40, 61	0
All	All	1034/1176 (87%)	0.25	23 (2%) 62 69	19, 30, 50, 77	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	42	GLY	4.9
2	B	133	GLY	4.7
1	C	297	TYR	4.6
2	B	1	GLN	3.8
1	A	343	LEU	3.8

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no monosaccharides in this entry.

6.4 Ligands

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^{th} 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
3	EDO	C	502	4/4	0.69	0.29	54,54,54,54	0
5	PEG	C	501	7/7	0.69	0.16	51,51,52,52	0
5	PEG	A	503	7/7	0.73	0.15	54,55,55,55	0
4	GOL	A	502	6/6	0.79	0.16	56,56,56,56	0
4	GOL	B	301	6/6	0.87	0.14	47,47,48,48	0
3	EDO	A	501	4/4	0.90	0.16	47,47,47,48	0
4	GOL	D	301	6/6	0.90	0.23	42,43,43,44	0

6.5 Other polymers (i)

There are no such residues in this entry.