Full wwPDB X-ray Structure Validation Report (i)

Aug 26, 2024 – 12:56 PM EDT

PDB ID	:	9C6W
Title	:	Crystal Structure of a single chain trimer composed of HLA-B*39:06 Y84C variant, beta-2microglobulin, and NRVMLPKAA peptide from NLRP2 (2 molecules/asymmetric unit)
Authors	:	Sharma, R.; Amdare, N.P.; Celikgil, A.; Garforth, S.J.; DiLorenzo, T.P.; Almo, S.C.; Ghosh, A.
Deposited on	:	2024-06-09
Resolution	:	1.70 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 (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 (1)) were used in the production of this report:

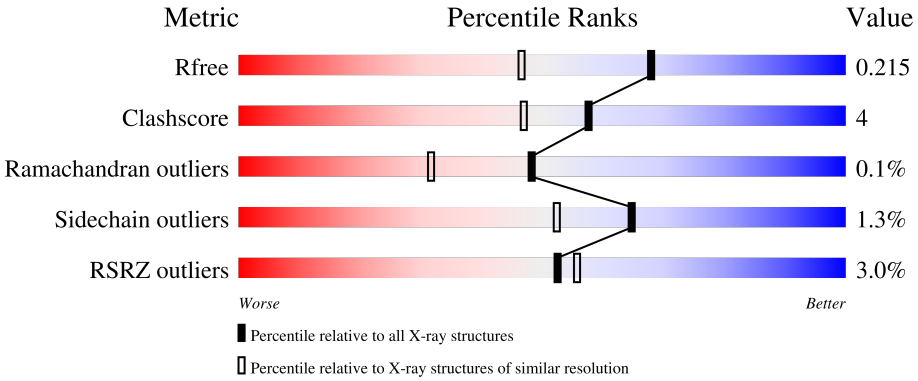
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.002 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.38.3

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 1.70 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	5161 (1.70-1.70)
Clashscore	180529	5671 (1.70-1.70)
Ramachandran outliers	177936	5594 (1.70-1.70)
Sidechain outliers	177891	5594 (1.70-1.70)
RSRZ outliers	164620	5159 (1.70-1.70)

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	419	3% 81% 11% • 8%
1	B	419	2% 85% 7% 8%

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
4	PEG	A	2316	-	-	X	-

2 Entry composition (i)

There are 8 unique types of molecules in this entry. The entry contains 7014 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 NACHT, LRR and PYD domains-containing protein 2,Beta-2 -microglobulin,MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	387	Total	C	N	O	S	0	12	0
			3214	2017	575	608	14			
1	B	386	Total	C	N	O	S	0	8	0
			3191	1997	569	612	13			

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	332	GLY	-	linker	UNP Q9NX02
A	333	CYS	-	linker	UNP Q9NX02
A	988	GLY	-	linker	UNP Q9NX02
A	989	ALA	-	linker	UNP Q9NX02
A	990	SER	-	linker	UNP Q9NX02
A	991	GLY	-	linker	UNP Q9NX02
A	992	GLY	-	linker	UNP Q9NX02
A	993	GLY	-	linker	UNP Q9NX02
A	994	GLY	-	linker	UNP Q9NX02
A	995	SER	-	linker	UNP Q9NX02
A	996	GLY	-	linker	UNP Q9NX02
A	997	GLY	-	linker	UNP Q9NX02
A	998	GLY	-	linker	UNP Q9NX02
A	999	GLY	-	linker	UNP Q9NX02
A	1000	SER	-	linker	UNP Q9NX02
A	1001	MET	-	linker	UNP Q9NX02
A	1981	GLY	-	linker	UNP P61769
A	1982	GLY	-	linker	UNP P61769
A	1983	GLY	-	linker	UNP P61769
A	1984	GLY	-	linker	UNP P61769
A	1985	SER	-	linker	UNP P61769
A	1986	GLY	-	linker	UNP P61769
A	1987	GLY	-	linker	UNP P61769
A	1988	GLY	-	linker	UNP P61769

Continued on next page...

Continued from previous page...

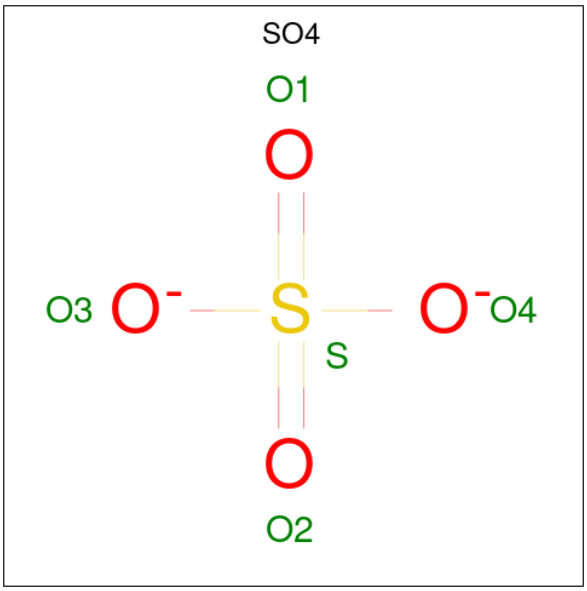
Chain	Residue	Modelled	Actual	Comment	Reference
A	1989	GLY	-	linker	UNP P61769
A	1990	SER	-	linker	UNP P61769
A	1991	GLY	-	linker	UNP P61769
A	1992	GLY	-	linker	UNP P61769
A	1993	GLY	-	linker	UNP P61769
A	1994	GLY	-	linker	UNP P61769
A	1995	SER	-	linker	UNP P61769
A	1996	GLY	-	linker	UNP P61769
A	1997	GLY	-	linker	UNP P61769
A	1998	GLY	-	linker	UNP P61769
A	1999	GLY	-	linker	UNP P61769
A	2000	SER	-	linker	UNP P61769
A	2001	GLY	-	linker	UNP P61769
A	2084	CYS	TYR	variant	UNP I3ZN83
B	332	GLY	-	linker	UNP Q9NX02
B	333	CYS	-	linker	UNP Q9NX02
B	988	GLY	-	linker	UNP Q9NX02
B	989	ALA	-	linker	UNP Q9NX02
B	990	SER	-	linker	UNP Q9NX02
B	991	GLY	-	linker	UNP Q9NX02
B	992	GLY	-	linker	UNP Q9NX02
B	993	GLY	-	linker	UNP Q9NX02
B	994	GLY	-	linker	UNP Q9NX02
B	995	SER	-	linker	UNP Q9NX02
B	996	GLY	-	linker	UNP Q9NX02
B	997	GLY	-	linker	UNP Q9NX02
B	998	GLY	-	linker	UNP Q9NX02
B	999	GLY	-	linker	UNP Q9NX02
B	1000	SER	-	linker	UNP Q9NX02
B	1001	MET	-	linker	UNP Q9NX02
B	1981	GLY	-	linker	UNP P61769
B	1982	GLY	-	linker	UNP P61769
B	1983	GLY	-	linker	UNP P61769
B	1984	GLY	-	linker	UNP P61769
B	1985	SER	-	linker	UNP P61769
B	1986	GLY	-	linker	UNP P61769
B	1987	GLY	-	linker	UNP P61769
B	1988	GLY	-	linker	UNP P61769
B	1989	GLY	-	linker	UNP P61769
B	1990	SER	-	linker	UNP P61769
B	1991	GLY	-	linker	UNP P61769
B	1992	GLY	-	linker	UNP P61769

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	1993	GLY	-	linker	UNP P61769
B	1994	GLY	-	linker	UNP P61769
B	1995	SER	-	linker	UNP P61769
B	1996	GLY	-	linker	UNP P61769
B	1997	GLY	-	linker	UNP P61769
B	1998	GLY	-	linker	UNP P61769
B	1999	GLY	-	linker	UNP P61769
B	2000	SER	-	linker	UNP P61769
B	2001	GLY	-	linker	UNP P61769
B	2084	CYS	TYR	variant	UNP I3ZN83

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



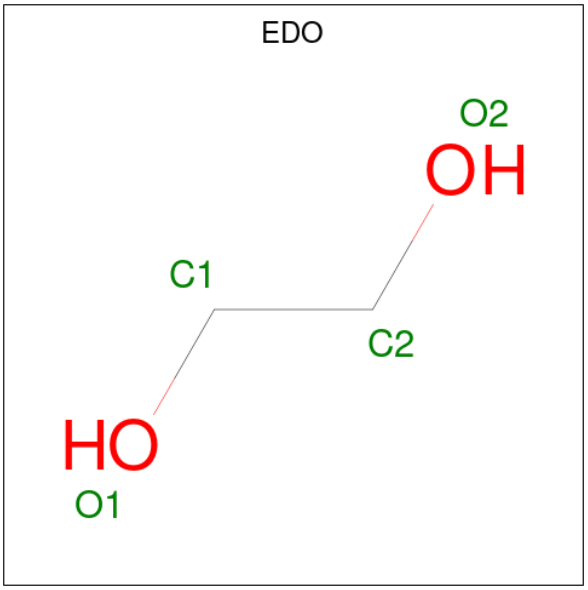
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0
2	A	1	Total O S 5 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0
2	B	1	Total O S 5 4 1	0	0

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



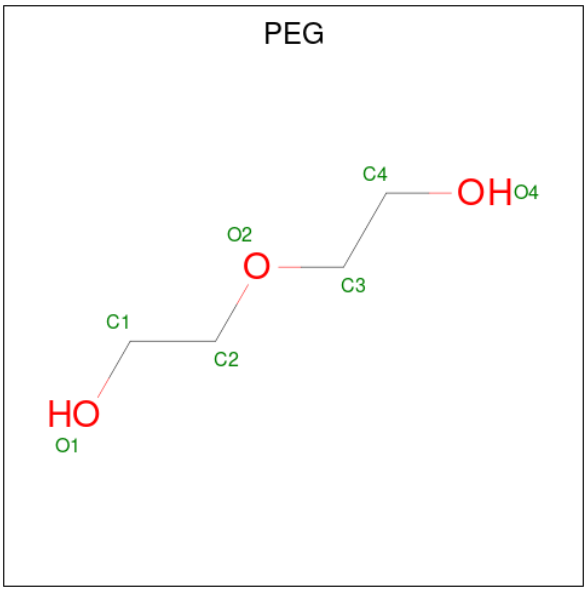
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 4 2 2	0	0
3	A	1	Total C O 8 4 4	0	1

Continued on next page...

Continued from previous page...

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

- Molecule 4 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

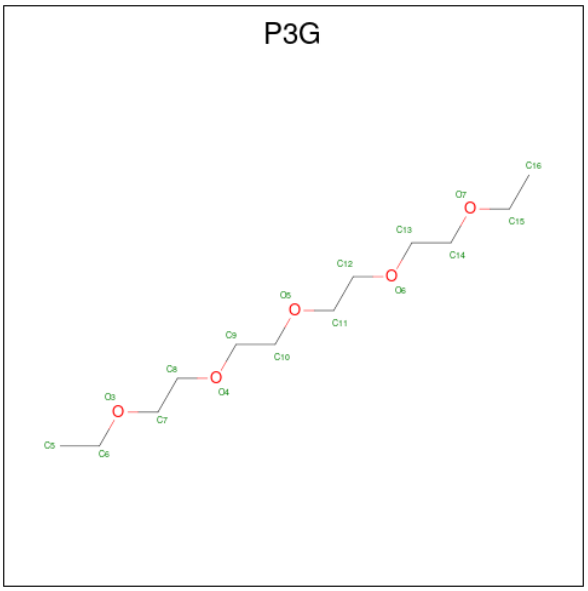


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 7 4 3	0	0
4	A	1	Total C O 7 4 3	0	0

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

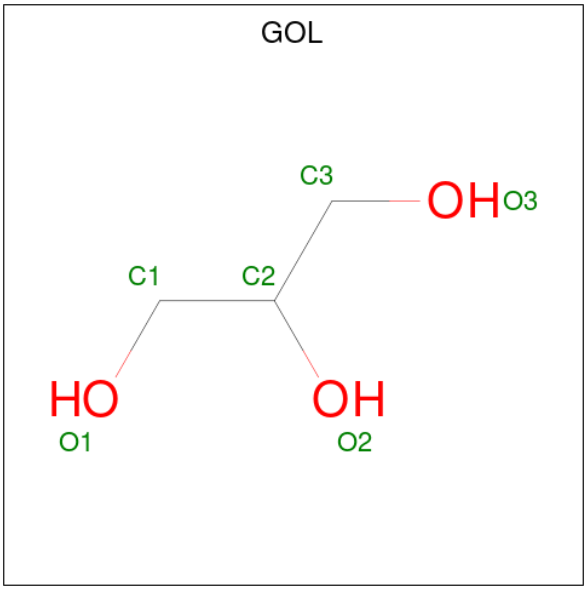
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	2	Total Cl 2 2	0	0

- Molecule 6 is 3,6,9,12,15-PENTAOXAHEPTADECANE (three-letter code: P3G) (formula: $C_{12}H_{26}O_5$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C O 17 12 5	0	0

- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 8 is water.

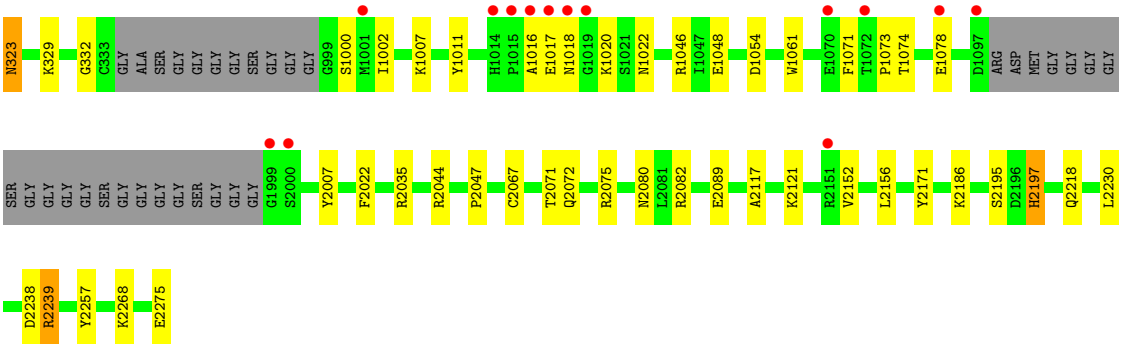
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	205	Total O 206 206	0	1
8	B	214	Total O 217 217	0	3

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.

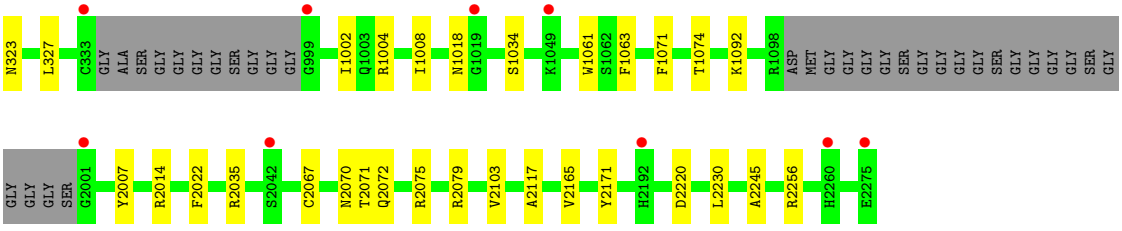
• Molecule 1: NACHT, LRR and PYD domains-containing protein 2,Beta-2-microglobulin,MHC class I antigen

Chain A:



• Molecule 1: NACHT, LRR and PYD domains-containing protein 2,Beta-2-microglobulin,MHC class I antigen

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.24Å 84.24Å 149.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.73 – 1.70 19.73 – 1.70	Depositor EDS
% Data completeness (in resolution range)	99.8 (19.73-1.70) 99.7 (19.73-1.70)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.62 (at 1.67Å)	Xtriage
Refinement program	PHENIX (1.21rc1_5156)	Depositor
R, R_{free}	0.187 , 0.215 0.187 , 0.215	Depositor DCC
R_{free} test set	5360 reflections (5.11%)	wwPDB-VP
Wilson B-factor (Å^2)	27.2	Xtriage
Anisotropy	0.205	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 54.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	7014	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 65.09 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 7.4870e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: P3G, GOL, CL, PEG, SO4, 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	A	0.37	0/3336	0.62	0/4530
1	B	0.36	0/3297	0.60	0/4477
All	All	0.36	0/6633	0.61	0/9007

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	3214	0	3075	35	0
1	B	3191	0	3022	21	0
2	A	35	0	0	0	0
2	B	40	0	0	1	0
3	A	36	0	54	7	0
3	B	36	0	54	4	0
4	A	14	0	20	4	0
5	B	2	0	0	0	0
6	B	17	0	26	0	0
7	B	6	0	8	0	0
8	A	206	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
8	B	217	0	0	3	0
All	All	7014	0	6259	57	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 (57) 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:B:1004:ARG:HH22	3:B:2314:EDO:H11	1.41	0.85
1:A:1017:GLU:HB2	1:A:1020:LYS:HB2	1.59	0.83
1:B:1018:ASN:HB3	3:B:2313:EDO:H22	1.74	0.69
1:A:2238:ASP:O	1:A:2239:ARG:HB3	1.93	0.68
1:B:2070[A]:ASN:ND2	8:B:2401:HOH:O	2.29	0.65
1:A:2067:CYS:O	1:A:2071[B]:THR:HG23	1.97	0.64
1:A:1000:SER:HB2	1:B:1002:ILE:HD11	1.79	0.64
1:B:1074:THR:HG22	3:B:2313:EDO:H12	1.80	0.61
1:B:2067:CYS:O	1:B:2071[B]:THR:HG23	1.99	0.61
1:A:1074:THR:HB	3:A:2309[A]:EDO:H12	1.82	0.61
1:A:2082[A]:ARG:HD3	8:A:2402:HOH:O	2.01	0.60
1:A:1018:ASN:O	3:A:2309[B]:EDO:O2	2.21	0.58
1:A:2072:GLN:HE22	1:A:2075[B]:ARG:CZ	2.19	0.55
1:B:323:ASN:N	1:B:2007:TYR:HH	2.05	0.54
1:A:323:ASN:N	1:A:2007:TYR:HH	2.05	0.54
1:A:323:ASN:N	1:A:2171:TYR:HH	2.06	0.53
1:A:1048:GLU:H	1:A:1048:GLU:CD	2.12	0.53
1:B:323:ASN:N	1:B:2171:TYR:HH	2.08	0.52
1:A:2047:PRO:HD2	3:A:2312:EDO:H11	1.90	0.52
1:A:2044:ARG:HH21	3:A:2312:EDO:H12	1.75	0.52
1:B:2075:ARG:HE	1:B:2079:ARG:HH22	1.57	0.52
1:B:1061:TRP:CE2	1:B:2117:ALA:HB2	2.46	0.51
1:A:2022:PHE:CD1	1:A:2071[B]:THR:HG22	2.46	0.50
1:A:1061:TRP:CE2	1:A:2117:ALA:HB2	2.46	0.50
1:B:2072[B]:GLN:HG3	8:B:2485:HOH:O	2.12	0.49
1:A:1016:ALA:HB1	1:A:1073:PRO:HG3	1.95	0.49
1:B:2220:ASP:OD2	1:B:2256:ARG:NH2	2.39	0.49
1:A:332:GLY:H	4:A:2316:PEG:H41	1.78	0.48
1:A:2022:PHE:CG	1:A:2071[B]:THR:HG22	2.49	0.47
1:A:2072:GLN:HE22	1:A:2075[B]:ARG:NH1	2.12	0.47
1:A:1078:GLU:HA	1:A:1078:GLU:OE1	2.15	0.47
1:B:2075:ARG:NE	1:B:2079:ARG:HH22	2.13	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:2314:EDO:H21	8:A:2449:HOH:O	2.14	0.46
1:B:2022:PHE:CD1	1:B:2071[B]:THR:HG22	2.51	0.46
1:B:2103:VAL:HG11	1:B:2165[A]:VAL:HG13	1.98	0.46
1:A:1016:ALA:HA	1:A:1022:ASN:HD22	1.81	0.45
1:A:2186:LYS:HD2	1:A:2186:LYS:HA	1.67	0.45
1:A:1054:ASP:OD2	1:A:2035[A]:ARG:HD3	2.16	0.44
1:A:329:LYS:NZ	4:A:2316:PEG:H11	2.32	0.44
1:A:1046:ARG:HD2	1:A:1046:ARG:H	1.83	0.44
1:B:2022:PHE:CG	1:B:2071[B]:THR:HG22	2.51	0.44
1:A:2044:ARG:NH2	3:A:2312:EDO:H12	2.33	0.43
1:A:2152:VAL:HG12	1:A:2156[B]:LEU:HD22	2.00	0.43
1:B:327:LEU:HD12	1:B:2070[A]:ASN:ND2	2.34	0.43
3:B:2316:EDO:H21	8:B:2462:HOH:O	2.19	0.43
1:A:329:LYS:HZ2	4:A:2316:PEG:H11	1.84	0.43
1:A:1002:ILE:HG12	1:A:2121:LYS:CG	2.48	0.43
1:A:2082[A]:ARG:CZ	1:A:2089:GLU:HG2	2.49	0.42
1:A:1020:LYS:HA	1:A:1020:LYS:HD3	1.90	0.42
1:A:1011:TYR:HB2	3:A:2310:EDO:H21	2.02	0.42
1:A:2080:ASN:HD21	4:A:2316:PEG:H22	1.84	0.41
1:B:2230:LEU:HD23	1:B:2245:ALA:HB2	2.02	0.41
1:A:2218:GLN:O	1:A:2257:TYR:HA	2.21	0.41
1:A:2195:SER:HG	1:A:2197:HIS:CE1	2.39	0.41
1:B:1008:ILE:HD12	1:B:1092:LYS:HD3	2.03	0.41
1:B:1034:SER:HB3	1:B:1063:PHE:CE2	2.56	0.41
1:B:2014:ARG:NH1	2:B:2302:SO4:O1	2.51	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	393/419 (94%)	385 (98%)	7 (2%)	1 (0%)	37 23

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
1	B	388/419 (93%)	382 (98%)	6 (2%)	0	100 100
All	All	781/838 (93%)	767 (98%)	13 (2%)	1 (0%)	48 32

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	2239	ARG

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	350/346 (101%)	342 (98%)	8 (2%)	45 29
1	B	346/346 (100%)	344 (99%)	2 (1%)	84 78
All	All	696/692 (101%)	686 (99%)	10 (1%)	65 49

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	323	ASN
1	A	1007	LYS
1	A	1071	PHE
1	A	2197	HIS
1	A	2230[A]	LEU
1	A	2230[B]	LEU
1	A	2268	LYS
1	A	2275	GLU
1	B	1071	PHE
1	B	2035	ARG

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

Mol	Chain	Res	Type
1	A	2072	GLN
1	A	2192	HIS

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no oligosaccharides in this entry.

5.6 Ligand geometry

Of 39 ligands modelled in this entry, 2 are monoatomic - leaving 37 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	$\#\|Z\| > 2$
3	EDO	B	2311	-	3,3,3	0.24	0	2,2,2	0.42	0
3	EDO	B	2313	-	3,3,3	0.24	0	2,2,2	0.27	0
2	SO4	B	2308	-	4,4,4	0.68	0	6,6,6	0.12	0
2	SO4	B	2302	-	4,4,4	0.70	0	6,6,6	0.15	0
3	EDO	A	2312	-	3,3,3	0.22	0	2,2,2	0.32	0
3	EDO	B	2317	-	3,3,3	0.25	0	2,2,2	0.35	0
2	SO4	A	2306	-	4,4,4	0.65	0	6,6,6	0.15	0
2	SO4	A	2303	-	4,4,4	0.68	0	6,6,6	0.11	0
3	EDO	A	2308	-	3,3,3	0.16	0	2,2,2	0.50	0
3	EDO	B	2316	-	3,3,3	0.23	0	2,2,2	0.18	0
3	EDO	B	2314	-	3,3,3	0.26	0	2,2,2	0.42	0
7	GOL	B	2321	-	5,5,5	0.31	0	5,5,5	0.38	0
3	EDO	A	2313	-	3,3,3	0.26	0	2,2,2	0.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	SO4	A	2304	-	4,4,4	0.69	0	6,6,6	0.07	0
2	SO4	B	2301	-	4,4,4	0.69	0	6,6,6	0.06	0
2	SO4	A	2302	-	4,4,4	0.70	0	6,6,6	0.14	0
3	EDO	B	2315	-	3,3,3	0.25	0	2,2,2	0.31	0
3	EDO	A	2309[A]	-	3,3,3	0.25	0	2,2,2	0.40	0
2	SO4	B	2305	-	4,4,4	0.66	0	6,6,6	0.12	0
3	EDO	A	2314	-	3,3,3	0.23	0	2,2,2	0.08	0
3	EDO	A	2315	-	3,3,3	0.25	0	2,2,2	0.29	0
2	SO4	B	2303	-	4,4,4	0.70	0	6,6,6	0.05	0
2	SO4	B	2307	-	4,4,4	0.70	0	6,6,6	0.08	0
3	EDO	A	2311	-	3,3,3	0.24	0	2,2,2	0.30	0
3	EDO	B	2312	-	3,3,3	0.31	0	2,2,2	0.39	0
2	SO4	A	2301	-	4,4,4	0.68	0	6,6,6	0.18	0
4	PEG	A	2316	-	6,6,6	0.23	0	5,5,5	0.39	0
6	P3G	B	2320	-	16,16,16	0.36	0	15,15,15	0.23	0
3	EDO	B	2318	-	3,3,3	0.24	0	2,2,2	0.46	0
3	EDO	A	2310	-	3,3,3	0.23	0	2,2,2	0.27	0
2	SO4	B	2306	-	4,4,4	0.66	0	6,6,6	0.07	0
2	SO4	A	2307	-	4,4,4	0.70	0	6,6,6	0.09	0
3	EDO	B	2319	-	3,3,3	0.22	0	2,2,2	0.43	0
2	SO4	A	2305	-	4,4,4	0.66	0	6,6,6	0.07	0
3	EDO	A	2309[B]	-	3,3,3	0.25	0	2,2,2	0.39	0
4	PEG	A	2317	-	6,6,6	0.27	0	5,5,5	0.28	0
2	SO4	B	2304	-	4,4,4	0.69	0	6,6,6	0.09	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	B	2311	-	-	1/1/1/1	-
3	EDO	B	2313	-	-	0/1/1/1	-
3	EDO	A	2312	-	-	0/1/1/1	-
3	EDO	B	2317	-	-	0/1/1/1	-
3	EDO	A	2308	-	-	0/1/1/1	-
3	EDO	B	2316	-	-	1/1/1/1	-
3	EDO	B	2314	-	-	0/1/1/1	-
7	GOL	B	2321	-	-	2/4/4/4	-
3	EDO	A	2313	-	-	0/1/1/1	-
3	EDO	B	2315	-	-	0/1/1/1	-
3	EDO	A	2309[A]	-	-	0/1/1/1	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	A	2314	-	-	1/1/1/1	-
3	EDO	A	2315	-	-	1/1/1/1	-
3	EDO	A	2311	-	-	0/1/1/1	-
3	EDO	B	2312	-	-	1/1/1/1	-
4	PEG	A	2316	-	-	1/4/4/4	-
6	P3G	B	2320	-	-	7/14/14/14	-
3	EDO	B	2318	-	-	1/1/1/1	-
3	EDO	A	2310	-	-	1/1/1/1	-
3	EDO	B	2319	-	-	0/1/1/1	-
3	EDO	A	2309[B]	-	-	0/1/1/1	-
4	PEG	A	2317	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (19) torsion outliers are listed below:

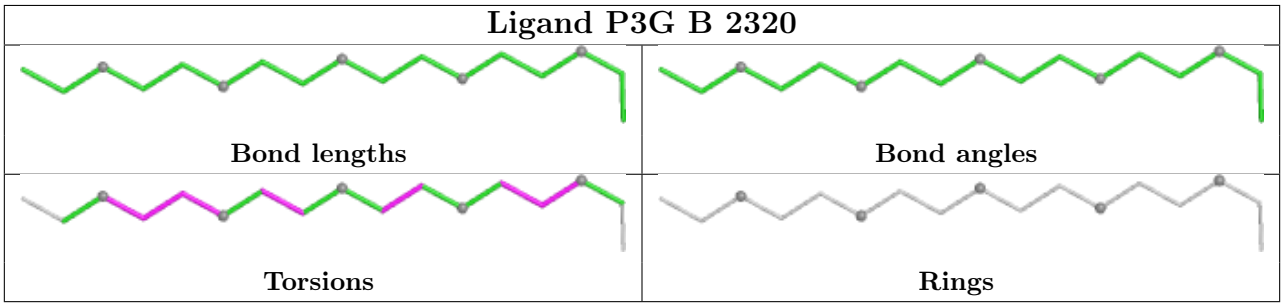
Mol	Chain	Res	Type	Atoms
7	B	2321	GOL	C1-C2-C3-O3
7	B	2321	GOL	O2-C2-C3-O3
4	A	2316	PEG	O1-C1-C2-O2
6	B	2320	P3G	O5-C10-C9-O4
3	B	2312	EDO	O1-C1-C2-O2
3	B	2316	EDO	O1-C1-C2-O2
6	B	2320	P3G	O3-C7-C8-O4
3	B	2311	EDO	O1-C1-C2-O2
3	B	2318	EDO	O1-C1-C2-O2
3	A	2315	EDO	O1-C1-C2-O2
4	A	2317	PEG	O2-C3-C4-O4
6	B	2320	P3G	C14-C13-O6-C12
4	A	2317	PEG	C1-C2-O2-C3
3	A	2310	EDO	O1-C1-C2-O2
6	B	2320	P3G	C13-C14-O7-C15
6	B	2320	P3G	C8-C7-O3-C6
6	B	2320	P3G	O5-C11-C12-O6
3	A	2314	EDO	O1-C1-C2-O2
6	B	2320	P3G	O6-C13-C14-O7

There are no ring outliers.

10 monomers are involved in 16 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	2313	EDO	2	0
2	B	2302	SO4	1	0
3	A	2312	EDO	3	0
3	B	2316	EDO	1	0
3	B	2314	EDO	1	0
3	A	2309[A]	EDO	1	0
3	A	2314	EDO	1	0
4	A	2316	PEG	4	0
3	A	2310	EDO	1	0
3	A	2309[B]	EDO	1	0

The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.



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	387/419 (92%)	0.20	14 (3%) 46 49	12, 33, 61, 80	12 (3%)
1	B	386/419 (92%)	0.06	9 (2%) 61 64	13, 32, 53, 82	8 (2%)
All	All	773/838 (92%)	0.13	23 (2%) 52 56	12, 33, 57, 82	20 (2%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1017	GLU	4.9
1	A	1016	ALA	3.9
1	B	999	GLY	3.7
1	B	2275	GLU	3.3
1	A	1019	GLY	3.3
1	A	2000	SER	3.1
1	A	1018	ASN	3.1
1	A	1014	HIS	3.0
1	A	1015	PRO	2.9
1	B	1049	LYS	2.7
1	A	1078	GLU	2.5
1	B	2042	SER	2.4
1	A	2151	ARG	2.4
1	A	1097	ASP	2.4
1	B	2260	HIS	2.3
1	B	333	CYS	2.2
1	B	2192	HIS	2.1
1	A	1072	THR	2.0
1	A	1001	MET	2.0
1	A	1999	GLY	2.0
1	B	1019	GLY	2.0
1	B	2001	GLY	2.0
1	A	1070	GLU	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^{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	B	2314	4/4	0.48	0.23	58,59,65,67	0
2	SO4	A	2303	5/5	0.49	0.15	94,96,108,115	0
3	EDO	B	2319	4/4	0.52	0.19	41,53,59,59	0
3	EDO	B	2317	4/4	0.61	0.17	52,54,55,57	0
6	P3G	B	2320	17/17	0.65	0.17	43,48,59,61	0
3	EDO	B	2318	4/4	0.67	0.20	36,43,43,50	0
3	EDO	B	2313	4/4	0.70	0.19	59,60,65,70	0
3	EDO	A	2315	4/4	0.70	0.18	50,50,54,55	0
2	SO4	B	2301	5/5	0.72	0.13	53,56,69,74	0
4	PEG	A	2316	7/7	0.74	0.18	46,54,58,61	0
3	EDO	B	2311	4/4	0.74	0.20	53,55,55,58	0
2	SO4	B	2307	5/5	0.75	0.09	64,65,77,77	5
2	SO4	B	2304	5/5	0.76	0.09	70,74,75,79	0
3	EDO	B	2315	4/4	0.77	0.15	34,55,57,62	0
3	EDO	A	2311	4/4	0.78	0.14	48,48,50,53	0
2	SO4	A	2304	5/5	0.79	0.12	45,56,59,61	5
3	EDO	A	2312	4/4	0.79	0.16	44,46,50,56	0
7	GOL	B	2321	6/6	0.81	0.12	57,62,65,65	0
2	SO4	B	2303	5/5	0.82	0.09	59,62,67,71	0
2	SO4	A	2302	5/5	0.84	0.09	55,60,68,68	0
2	SO4	A	2305	5/5	0.85	0.12	37,39,46,48	5
3	EDO	B	2316	4/4	0.85	0.13	41,43,47,51	0
3	EDO	A	2309[A]	4/4	0.85	0.28	79,80,81,82	4
3	EDO	A	2309[B]	4/4	0.85	0.28	76,80,80,81	4
2	SO4	B	2306	5/5	0.86	0.11	46,47,54,58	5
4	PEG	A	2317	7/7	0.86	0.12	51,54,62,71	0
2	SO4	B	2308	5/5	0.86	0.12	42,44,46,53	5

Continued on next page...

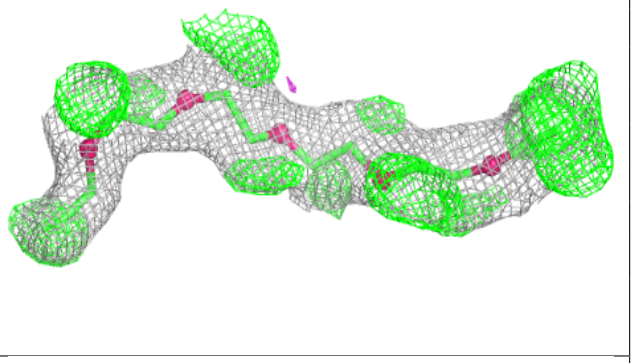
Continued from previous page...

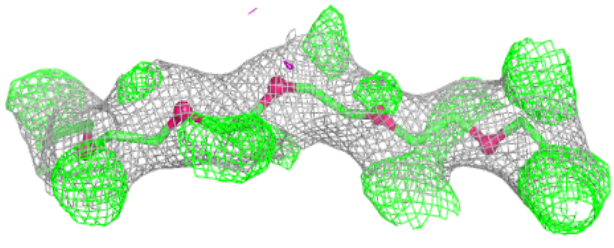
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	EDO	A	2314	4/4	0.86	0.13	42,42,44,52	0
2	SO4	A	2307	5/5	0.87	0.07	56,57,71,71	0
3	EDO	A	2313	4/4	0.87	0.10	44,49,53,54	0
3	EDO	A	2310	4/4	0.88	0.17	52,52,55,55	0
2	SO4	B	2305	5/5	0.88	0.10	44,51,57,58	5
3	EDO	A	2308	4/4	0.88	0.13	29,41,47,48	0
5	CL	B	2310	1/1	0.89	0.09	71,71,71,71	0
5	CL	B	2309	1/1	0.90	0.13	68,68,68,68	0
3	EDO	B	2312	4/4	0.90	0.10	29,39,41,54	0
2	SO4	A	2306	5/5	0.91	0.12	35,37,42,49	5
2	SO4	B	2302	5/5	0.93	0.08	42,51,55,63	0
2	SO4	A	2301	5/5	0.98	0.07	30,36,38,43	0

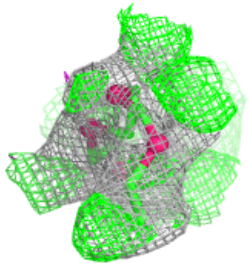
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

Electron density around P3G B 2320:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative) and green (positive)







6.5 Other polymers

There are no such residues in this entry.