



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 5, 2023 – 04:04 PM EST

PDB ID : 7LFV  
Title : Crystal structure of the SARS CoV-1 Papain-like protease in complex with peptide inhibitor VIR251  
Authors : Olsen, S.K.; Lv, Z.  
Deposited on : 2021-01-18  
Resolution : 2.23 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](mailto: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) : 1.13  
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

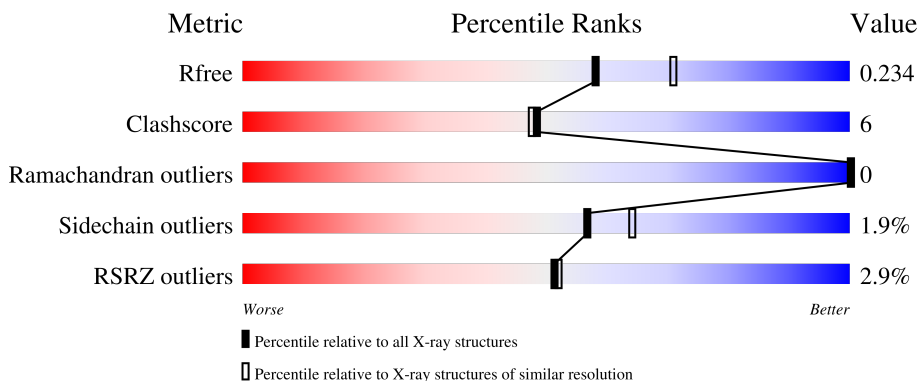
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.23 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	2391 (2.26-2.22)
Clashscore	141614	2539 (2.26-2.22)
Ramachandran outliers	138981	2489 (2.26-2.22)
Sidechain outliers	138945	2490 (2.26-2.22)
RSRZ outliers	127900	2353 (2.26-2.22)

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  $\geq 3$ , 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	325	 5% 86% 11%
1	B	325	 5% 80% 15%
2	D	5	 80% 20%
2	F	5	 80% 20%

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	EDO	A	2005	-	-	X	-

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 5221 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 papain-like protease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	B	312	2456	1558	409	472	17	0	0	0
1	A	314	2473	1570	412	474	17	0	0	0

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1	MET	-	initiating methionine	UNP P0C6U8
B	318	LEU	-	expression tag	UNP P0C6U8
B	319	GLU	-	expression tag	UNP P0C6U8
B	320	HIS	-	expression tag	UNP P0C6U8
B	321	HIS	-	expression tag	UNP P0C6U8
B	322	HIS	-	expression tag	UNP P0C6U8
B	323	HIS	-	expression tag	UNP P0C6U8
B	324	HIS	-	expression tag	UNP P0C6U8
B	325	HIS	-	expression tag	UNP P0C6U8
A	1	MET	-	initiating methionine	UNP P0C6U8
A	318	LEU	-	expression tag	UNP P0C6U8
A	319	GLU	-	expression tag	UNP P0C6U8
A	320	HIS	-	expression tag	UNP P0C6U8
A	321	HIS	-	expression tag	UNP P0C6U8
A	322	HIS	-	expression tag	UNP P0C6U8
A	323	HIS	-	expression tag	UNP P0C6U8
A	324	HIS	-	expression tag	UNP P0C6U8
A	325	HIS	-	expression tag	UNP P0C6U8

- Molecule 2 is a protein called Papain-like protease peptide inhibitor VIR251.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	D	5	34	22	5	7	0	0	0

*Continued on next page...*

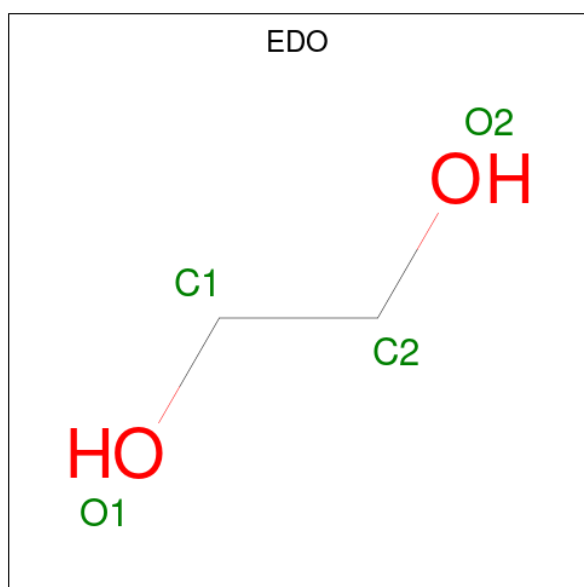
Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	F	5	34	22	5	7	0	0	0

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Zn	0	0
			1	1		
3	A	1	Total	Zn	0	0
			1	1		

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



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

Continued on next page...

*Continued from previous page...*

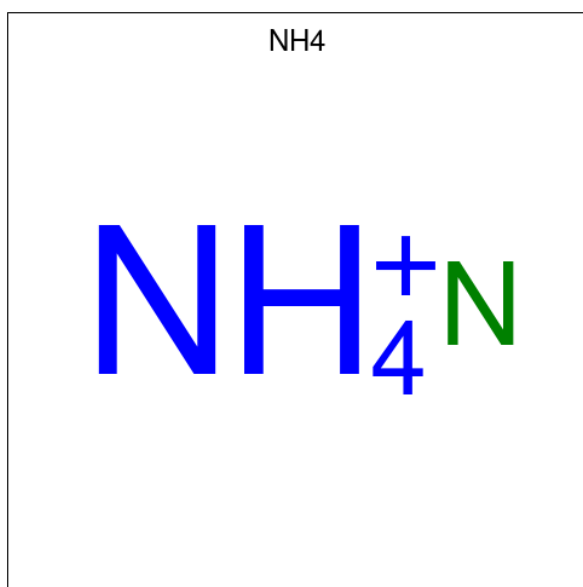
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	B	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0	0
4	A	1	Total 4	C 2	O 2	0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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

- Molecule 6 is AMMONIUM ION (three-letter code: NH4) (formula: H<sub>4</sub>N).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total N 1 1	0	0

- Molecule 7 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total Cl 1 1	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	74	Total O 74 74	0	0
8	A	46	Total O 46 46	0	0
8	F	1	Total O 1 1	0	0





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	103.61Å 103.61Å 265.35Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	53.34 – 2.23 89.73 – 2.23	Depositor EDS
% Data completeness (in resolution range)	99.9 (53.34-2.23) 94.9 (89.73-2.23)	Depositor EDS
$R_{merge}$	0.10	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 2.22Å)	Xtrriage
Refinement program	PHENIX 1.17.1_3660	Depositor
R, $R_{free}$	0.197 , 0.235 0.197 , 0.234	Depositor DCC
$R_{free}$ test set	2000 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.2	Xtrriage
Anisotropy	0.134	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 46.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	5221	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	65.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.57% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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, ZN, DPP, ACY, GVE, CL, NH4, 73O, 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	0.42	0/2529	0.53	0/3432
1	B	0.51	0/2511	0.70	3/3408 (0.1%)
2	D	1.82	0/3	0.37	0/2
2	F	1.73	0/3	1.40	0/2
All	All	0.47	0/5046	0.62	3/6844 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	76	LEU	CA-CB-CG	16.23	152.62	115.30
1	B	76	LEU	CB-CG-CD1	9.40	126.97	111.00
1	B	76	LEU	CB-CG-CD2	-8.05	97.31	111.00

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	2473	0	2417	23	0
1	B	2456	0	2397	35	1
2	D	34	0	12	0	0
2	F	34	0	11	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	1	0	0	0	0
3	B	1	0	0	0	0
4	A	36	0	54	14	0
4	B	48	0	72	6	0
5	A	5	0	0	0	0
5	B	10	0	0	0	0
6	A	1	0	0	0	0
7	A	1	0	0	0	0
8	A	46	0	0	2	0
8	B	74	0	0	2	0
8	F	1	0	0	0	0
All	All	5221	0	4963	60	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 6.

All (60) 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:38:ASP:OD1	8:A:2101:HOH:O	1.89	0.89
1:A:292:THR:HG23	4:A:2005:EDO:H11	1.72	0.72
1:A:293:LYS:H	4:A:2005:EDO:H22	1.55	0.71
1:A:212:LEU:H	4:A:2009:EDO:H12	1.56	0.71
1:B:223:ILE:HD12	1:B:224:PRO:HD2	1.73	0.69
4:A:2006:EDO:O1	8:A:2103:HOH:O	2.14	0.63
1:B:178:ASN:OD1	1:B:180:GLU:HG2	2.00	0.62
1:A:241:SER:O	4:A:2010:EDO:H11	2.00	0.61
1:A:167:ARG:HG3	4:A:2008:EDO:H12	1.83	0.60
1:B:139:ARG:NH2	1:B:144:ASP:OD2	2.35	0.60
1:B:127:LYS:NZ	1:B:138:TYR:OH	2.35	0.59
1:A:195:GLN:HE22	1:A:317:PRO:HB3	1.68	0.59
1:B:67:SER:O	1:B:71:GLU:HG2	2.03	0.57
4:B:405:EDO:H21	1:A:97:PRO:HA	1.87	0.57
1:B:227:CYS:SG	1:B:228:GLY:N	2.79	0.56
1:B:190:CYS:SG	1:B:229:ARG:NH1	2.79	0.56
1:B:253:LYS:HE3	1:B:298:LYS:HE2	1.91	0.53
1:A:137:TYR:CE1	1:A:141:ARG:HD2	2.44	0.53
1:A:307:LYS:HG2	4:A:2010:EDO:H22	1.89	0.53
4:B:404:EDO:H22	4:B:405:EDO:O1	2.09	0.52
1:B:309:THR:HG23	1:B:310:SER:N	2.25	0.52
1:A:131:PRO:HG2	4:A:2006:EDO:H11	1.93	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:124:LEU:HA	4:B:412:EDO:H12	1.92	0.50
1:A:79:SER:OG	1:A:83:ARG:NH1	2.46	0.49
1:A:125:GLU:OE1	4:A:2010:EDO:H21	2.13	0.49
1:B:9:PHE:CZ	4:A:2005:EDO:H21	2.48	0.49
1:A:207:MET:HB3	4:A:2008:EDO:H11	1.94	0.49
1:B:186:LEU:HB2	1:B:198:THR:OG1	2.13	0.49
1:B:235:LEU:HD21	1:B:238:GLN:HB2	1.97	0.47
1:A:195:GLN:HE22	1:A:317:PRO:CB	2.27	0.47
1:A:302:THR:HG23	1:A:303:ASP:OD1	2.15	0.47
1:B:69:ALA:HB1	4:B:409:EDO:H11	1.97	0.47
1:B:278:THR:HG23	1:B:284:TYR:HB2	1.97	0.46
1:A:6:ILE:HG13	1:A:22:VAL:HG12	1.98	0.46
1:B:19:THR:H	4:B:403:EDO:H12	1.81	0.46
1:B:63:ASP:OD2	8:B:501:HOH:O	2.21	0.45
1:B:182:ALA:HA	1:B:239:GLU:O	2.16	0.45
1:B:191:LYS:HG2	1:B:229:ARG:CZ	2.47	0.45
1:B:114:LEU:HD11	1:B:153:LEU:HD21	1.98	0.45
1:B:187:ASN:O	1:B:233:GLN:HA	2.17	0.45
1:A:224:PRO:HA	1:A:230:ASP:OD2	2.17	0.44
1:B:191:LYS:H	1:B:229:ARG:NH1	2.15	0.44
1:B:220:GLY:HA2	1:B:234:TYR:HB3	1.99	0.44
1:A:137:TYR:O	1:A:141:ARG:HG3	2.18	0.44
1:A:185:VAL:HG12	1:A:236:VAL:HB	2.00	0.44
1:B:86:SER:OG	4:B:408:EDO:H12	2.18	0.43
1:A:71:GLU:HB3	4:A:2006:EDO:H22	2.00	0.43
1:A:227:CYS:SG	1:A:229:ARG:HB2	2.59	0.43
1:B:223:ILE:HG13	1:B:224:PRO:O	2.19	0.42
1:B:152:ILE:HG12	1:B:173:LEU:HD11	2.01	0.42
1:B:229:ARG:NH2	1:B:230:ASP:O	2.53	0.42
1:B:253:LYS:HG2	1:B:298:LYS:HG3	2.00	0.42
1:A:170:MET:HE3	1:A:174:LEU:HD11	2.00	0.42
1:B:207:MET:HE2	8:B:529:HOH:O	2.19	0.41
1:B:309:THR:HG23	1:B:310:SER:H	1.84	0.41
1:B:7:LYS:CD	4:A:2005:EDO:H12	2.50	0.41
1:B:23:ASP:OD1	1:B:25:SER:HB3	2.21	0.41
1:B:309:THR:CG2	1:B:310:SER:N	2.84	0.41
1:B:185:VAL:HB	1:B:237:GLN:HB3	2.03	0.40
1:B:7:LYS:HD2	4:A:2005:EDO:H12	2.02	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)
1:B:195:GLN:NE2	1:B:311:TYR:OH[9_555]	2.14	0.06

### 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	312/325 (96%)	301 (96%)	11 (4%)	0	100	100
1	B	310/325 (95%)	301 (97%)	9 (3%)	0	100	100
2	D	1/5 (20%)	1 (100%)	0	0	100	100
2	F	1/5 (20%)	1 (100%)	0	0	100	100
All	All	624/660 (94%)	604 (97%)	20 (3%)	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	271/282 (96%)	266 (98%)	5 (2%)	59	66
1	B	269/282 (95%)	264 (98%)	5 (2%)	57	64
All	All	540/564 (96%)	530 (98%)	10 (2%)	57	64

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

Mol	Chain	Res	Type
1	B	25	SER
1	B	67	SER
1	B	86	SER
1	B	181	SER
1	B	195	GLN
1	A	4	LYS
1	A	93	LYS
1	A	157	ASN
1	A	191	LYS
1	A	316	LYS

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	195	GLN
1	A	237	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

6 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
2	DPP	F	3	2	3,5,6	0.49	0	1,5,7	0.02	0
2	73O	D	2	2	12,13,14	0.55	0	11,16,18	0.74	0
2	GVE	F	5	1,2	7,7,7	1.02	0	7,7,7	1.73	3 (42%)
2	DPP	D	3	2	3,5,6	0.71	0	1,5,7	0.26	0
2	GVE	D	5	1,2	7,7,7	1.30	1 (14%)	7,7,7	1.08	0
2	73O	F	2	2	12,13,14	0.71	0	11,16,18	0.59	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
2	DPP	F	3	2	-	0/2/4/6	-
2	73O	D	2	2	-	3/6/7/9	0/1/1/1
2	GVE	F	5	1,2	-	4/6/6/6	-
2	DPP	D	3	2	-	0/2/4/6	-
2	GVE	D	5	1,2	-	6/6/6/6	-
2	73O	F	2	2	-	1/6/7/9	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	5	GVE	OXT-C	2.45	1.40	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	5	GVE	CB-CG-C	2.82	123.89	113.62
2	F	5	GVE	C1-CB-CG	2.65	120.71	112.84
2	F	5	GVE	CH3-OXT-C	-2.17	108.94	116.11

There are no chirality outliers.

All (14) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	2	73O	C54-C52-CA-N
2	D	2	73O	C54-C52-CA-C
2	D	5	GVE	C1-CB-CG-C
2	F	5	GVE	C1-CB-CG-C
2	F	5	GVE	CG-C-OXT-CH3
2	F	5	GVE	O-C-OXT-CH3
2	D	5	GVE	CG-C-OXT-CH3
2	D	5	GVE	O-C-OXT-CH3
2	F	2	73O	C54-C52-CA-C
2	D	5	GVE	N-C1-CB-CG
2	F	5	GVE	N-C1-CB-CG
2	D	2	73O	CA-C52-C54-C57
2	D	5	GVE	OXT-C-CG-CB
2	D	5	GVE	O-C-CG-CB



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](#)

Of 28 ligands modelled in this entry, 3 are monoatomic and 1 is modelled with single atom - leaving 24 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	$\# Z  > 2$
4	EDO	B	403	-	3,3,3	0.66	0	2,2,2	0.47	0
4	EDO	B	405	-	3,3,3	0.55	0	2,2,2	0.35	0
4	EDO	A	2010	-	3,3,3	0.43	0	2,2,2	0.26	0
5	SO4	A	2012	-	4,4,4	0.17	0	6,6,6	0.49	0
4	EDO	A	2007	-	3,3,3	0.41	0	2,2,2	0.46	0
4	EDO	B	406	-	3,3,3	0.44	0	2,2,2	0.57	0
4	EDO	B	408	-	3,3,3	0.54	0	2,2,2	0.20	0
4	EDO	B	409	-	3,3,3	0.50	0	2,2,2	0.37	0
4	EDO	A	2008	-	3,3,3	0.45	0	2,2,2	0.35	0
4	EDO	A	2009	-	3,3,3	0.52	0	2,2,2	0.20	0
4	EDO	B	411	-	3,3,3	0.63	0	2,2,2	0.36	0
4	EDO	B	413	-	3,3,3	0.54	0	2,2,2	0.17	0
4	EDO	A	2004	-	3,3,3	0.50	0	2,2,2	0.38	0
4	EDO	B	402	-	3,3,3	0.60	0	2,2,2	0.15	0
4	EDO	B	404	-	3,3,3	0.42	0	2,2,2	0.54	0
4	EDO	A	2006	-	3,3,3	0.43	0	2,2,2	0.31	0
5	SO4	B	415	-	4,4,4	0.29	0	6,6,6	0.60	0
4	EDO	B	410	-	3,3,3	0.49	0	2,2,2	0.34	0
4	EDO	B	412	-	3,3,3	0.44	0	2,2,2	0.33	0
4	EDO	A	2003	-	3,3,3	0.49	0	2,2,2	0.32	0
4	EDO	A	2001	-	3,3,3	0.48	0	2,2,2	0.42	0
4	EDO	A	2005	-	3,3,3	0.53	0	2,2,2	0.24	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	B	414	-	4,4,4	0.16	0	6,6,6	0.15	0
4	EDO	B	407	-	3,3,3	0.46	0	2,2,2	0.48	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	EDO	B	403	-	-	0/1/1/1	-
4	EDO	B	405	-	-	1/1/1/1	-
4	EDO	A	2010	-	-	1/1/1/1	-
4	EDO	A	2007	-	-	0/1/1/1	-
4	EDO	B	406	-	-	0/1/1/1	-
4	EDO	B	408	-	-	0/1/1/1	-
4	EDO	B	409	-	-	0/1/1/1	-
4	EDO	A	2008	-	-	0/1/1/1	-
4	EDO	A	2009	-	-	0/1/1/1	-
4	EDO	B	411	-	-	1/1/1/1	-
4	EDO	B	413	-	-	1/1/1/1	-
4	EDO	A	2004	-	-	0/1/1/1	-
4	EDO	B	402	-	-	0/1/1/1	-
4	EDO	B	404	-	-	0/1/1/1	-
4	EDO	A	2006	-	-	0/1/1/1	-
4	EDO	B	410	-	-	0/1/1/1	-
4	EDO	B	412	-	-	0/1/1/1	-
4	EDO	A	2003	-	-	0/1/1/1	-
4	EDO	A	2001	-	-	0/1/1/1	-
4	EDO	A	2005	-	-	1/1/1/1	-
4	EDO	B	407	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	405	EDO	O1-C1-C2-O2
4	B	411	EDO	O1-C1-C2-O2
4	A	2010	EDO	O1-C1-C2-O2
4	A	2005	EDO	O1-C1-C2-O2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
4	B	413	EDO	O1-C1-C2-O2

There are no ring outliers.

11 monomers are involved in 20 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	B	403	EDO	1	0
4	B	405	EDO	2	0
4	A	2010	EDO	3	0
4	B	408	EDO	1	0
4	B	409	EDO	1	0
4	A	2008	EDO	2	0
4	A	2009	EDO	1	0
4	B	404	EDO	1	0
4	A	2006	EDO	3	0
4	B	412	EDO	1	0
4	A	2005	EDO	5	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	314/325 (96%)	-0.33	2 (0%) 89 89	43, 62, 95, 147	0
1	B	312/325 (96%)	0.02	16 (5%) 28 27	38, 56, 132, 177	0
2	D	1/5 (20%)	-0.18	0 100 100	63, 63, 63, 63	0
2	F	1/5 (20%)	-0.55	0 100 100	39, 39, 39, 39	0
All	All	628/660 (95%)	-0.15	18 (2%) 51 52	38, 60, 112, 177	0

All (18) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	234	TYR	7.8
1	B	193	CYS	5.5
1	B	231	ALA	5.4
1	B	226	VAL	5.2
1	B	195	GLN	4.9
1	B	227	CYS	4.3
1	B	225	CYS	3.8
1	B	196	LYS	3.3
1	A	81	LEU	3.1
1	B	223	ILE	3.1
1	B	194	GLY	3.0
1	B	197	THR	3.0
1	B	311	TYR	2.9
1	B	192	HIS	2.9
1	A	226	VAL	2.4
1	B	312	THR	2.4
1	B	224	PRO	2.3
1	B	221	VAL	2.1

## 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
2	GVE	F	5	8/8	0.90	0.16	43,47,60,76	0
2	GVE	D	5	8/8	0.92	0.14	50,65,73,78	0
2	73O	D	2	13/14	0.92	0.20	69,90,107,112	0
2	73O	F	2	13/14	0.94	0.13	47,56,71,76	0
2	DPP	D	3	6/7	0.95	0.09	66,70,72,74	0
2	DPP	F	3	6/7	0.96	0.11	46,48,49,49	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
4	EDO	B	408	4/4	0.67	0.22	83,87,87,88	0
3	ZN	B	401	1/1	0.68	0.05	122,122,122,122	0
4	EDO	A	2005	4/4	0.68	0.35	78,83,83,83	0
4	EDO	B	403	4/4	0.69	0.23	54,55,58,64	0
4	EDO	B	411	4/4	0.71	0.26	71,75,75,75	0
4	EDO	B	407	4/4	0.74	0.23	83,83,85,90	0
4	EDO	B	413	4/4	0.75	0.28	79,82,86,90	0
4	EDO	A	2004	4/4	0.76	0.24	71,75,79,79	0
4	EDO	A	2007	4/4	0.76	0.18	87,87,91,92	0
4	EDO	B	409	4/4	0.81	0.19	63,67,70,70	0
4	EDO	A	2003	4/4	0.82	0.17	69,73,74,79	0
4	EDO	A	2010	4/4	0.82	0.32	86,87,88,88	0
4	EDO	A	2009	4/4	0.83	0.15	78,79,82,82	0
4	EDO	B	404	4/4	0.85	0.16	84,85,86,89	0
7	CL	A	2013	1/1	0.86	0.11	85,85,85,85	0
4	EDO	B	405	4/4	0.87	0.19	65,69,72,74	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
4	EDO	A	2001	4/4	0.88	0.10	76,78,78,78	0
4	EDO	B	402	4/4	0.88	0.17	75,75,76,76	0
4	EDO	A	2006	4/4	0.89	0.51	67,70,73,75	0
4	EDO	B	410	4/4	0.91	0.11	71,75,75,76	0
4	EDO	B	412	4/4	0.91	0.38	82,85,90,92	0
4	EDO	B	406	4/4	0.92	0.17	70,73,73,76	0
4	EDO	A	2008	4/4	0.93	0.29	71,72,73,73	0
3	ZN	A	2002	1/1	0.95	0.11	92,92,92,92	0
6	NH4	A	2011	1/1	0.97	0.24	77,77,77,77	0
5	SO4	A	2012	5/5	0.98	0.12	73,74,76,82	0
5	SO4	B	415	5/5	0.99	0.12	56,59,65,65	0
5	SO4	B	414	5/5	0.99	0.12	66,72,74,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.