



Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2020 – 07:25 am BST

PDB ID : 3FDE
Title : Mouse UHRF1 SRA domain bound with hemi-methylated CpG DNA, crystal structure in space group C222(1) at 1.4 Å resolution
Authors : Hashimoto, H.; Horton, J.R.; Zhang, X.; Cheng, X.
Deposited on : 2008-11-25
Resolution : 1.41 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

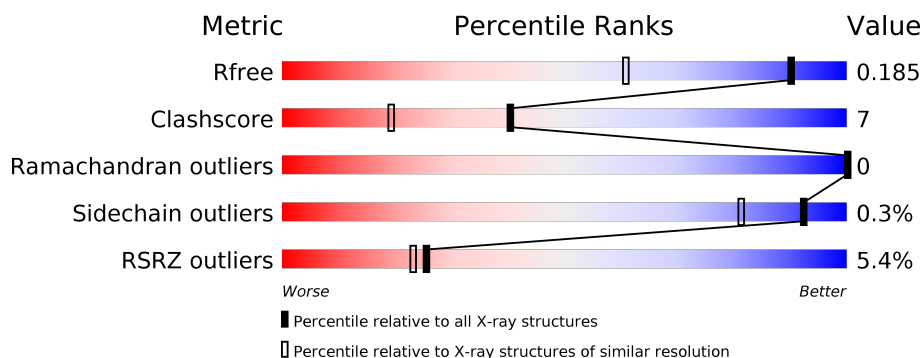
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.41 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	2579 (1.44-1.40)
Clashscore	141614	2696 (1.44-1.40)
Ramachandran outliers	138981	2632 (1.44-1.40)
Sidechain outliers	138945	2631 (1.44-1.40)
RSRZ outliers	127900	2528 (1.44-1.40)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	212	<div> <div>7%</div> <div> <div></div> <div>88%</div> <div>11%</div> </div> </div>
1	B	212	<div> <div>4%</div> <div> <div></div> <div>89%</div> <div>8%</div> </div> </div>
2	C	12	<div> <div>8%</div> <div> <div>25%</div> <div>67%</div> <div>8%</div> </div> </div>
2	D	12	<div> <div>8%</div> <div> <div></div> <div>92%</div> </div> </div>
3	E	12	<div> <div></div> <div> <div>83%</div> <div>17%</div> </div> </div>
3	F	12	<div> <div>8%</div> <div> <div>25%</div> <div>42%</div> <div>33%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	A	636	-	X	X	-
4	EDO	B	633	-	-	X	-
6	UNL	A	638	-	-	X	-
6	UNL	A	654	-	-	X	-
6	UNL	B	636	-	-	X	-
6	UNL	C	456	-	-	X	-
6	UNL	C	637	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 5159 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 E3 ubiquitin-protein ligase UHRF1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	210	Total	C	N	O	S	0	7	1
			1698	1051	326	315	6			
1	B	206	Total	C	N	O	S	0	4	0
			1651	1024	318	304	5			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	HIS	-	EXPRESSION TAG	UNP Q8VDF2
A	418	MET	-	EXPRESSION TAG	UNP Q8VDF2
B	417	HIS	-	EXPRESSION TAG	UNP Q8VDF2
B	418	MET	-	EXPRESSION TAG	UNP Q8VDF2

- Molecule 2 is a DNA chain called 5'-D(*CP*CP*AP*TP*GP*(5CM)P*GP*CP*TP*GP*A P*C)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	12	Total	C	N	O	P	0	0	0
			241	116	44	70	11			
2	C	12	Total	C	N	O	P	0	2	0
			283	136	52	82	13			

- Molecule 3 is a DNA chain called 5'-D(*GP*TP*CP*AP*GP*CP*GP*CP*AP*TP*GP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	12	Total	C	N	O	P	0	0	0
			246	117	48	70	11			
3	F	12	Total	C	N	O	P	0	0	0
			246	117	48	70	11			

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).



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

Continued on next page...

Continued from previous page...

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

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	B	1	Total Na 1 1	0	0
5	A	1	Total Na 1 1	0	0
5	F	1	Total Na 1 1	0	0
5	E	1	Total Na 1 1	0	0

- Molecule 6 is UNKNOWN LIGAND (three-letter code: UNL) (formula:).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	4	Total C O 12 2 10	0	0
6	A	2	Total O 4 4	0	0
6	C	3	Total O 6 6	0	0

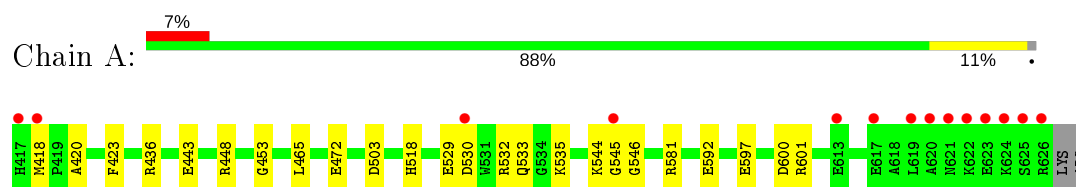
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	257	Total 257	O 257	0	0
7	B	269	Total 269	O 269	0	0
7	D	40	Total 40	O 40	0	0
7	E	37	Total 37	O 37	0	0
7	C	43	Total 43	O 43	0	0
7	F	34	Total 34	O 34	0	0

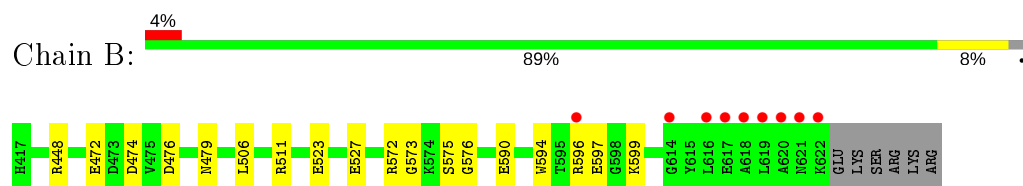
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: E3 ubiquitin-protein ligase UHRF1



- Molecule 1: E3 ubiquitin-protein ligase UHRF1



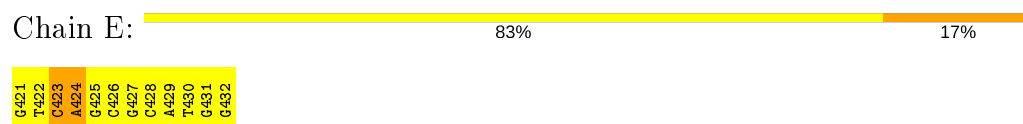
- Molecule 2: 5'-D(*CP*CP*AP*TP*GP*(5CM)P*GP*CP*TP*GP*AP*C)-3'



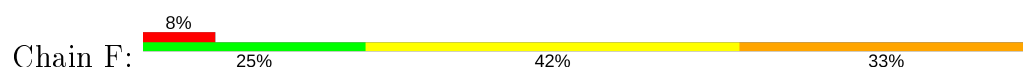
- Molecule 2: 5'-D(*CP*CP*AP*TP*GP*(5CM)P*GP*CP*TP*GP*AP*C)-3'



- Molecule 3: 5'-D(*GP*TP*CP*AP*GP*CP*GP*CP*AP*TP*GP*G)-3'



- Molecule 3: 5'-D(*GP*TP*CP*AP*GP*CP*GP*CP*AP*TP*GP*G)-3'





4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	84.68Å 103.69Å 149.84Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.79 – 1.41 32.79 – 1.41	Depositor EDS
% Data completeness (in resolution range)	96.1 (32.79-1.41) 96.0 (32.79-1.41)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.09 (at 1.41Å)	Xtriage
Refinement program	REFMAC 5.5.0044	Depositor
R, R_{free}	0.149 , 0.186 0.148 , 0.185	Depositor DCC
R_{free} test set	6067 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å ²)	14.3	Xtriage
Anisotropy	0.101	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 51.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.98	EDS
Total number of atoms	5159	wwPDB-VP
Average B, all atoms (Å ²)	19.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.99% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NA, 5CM, UNL, 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	1.09	3/1740 (0.2%)	1.00	4/2350 (0.2%)
1	B	1.13	4/1694 (0.2%)	0.99	5/2291 (0.2%)
2	C	1.68	1/271 (0.4%)	2.30	17/414 (4.1%)
2	D	2.10	6/246 (2.4%)	2.63	26/375 (6.9%)
3	E	1.86	4/276 (1.4%)	3.24	53/425 (12.5%)
3	F	1.76	1/276 (0.4%)	2.42	26/425 (6.1%)
All	All	1.32	19/4503 (0.4%)	1.61	131/6280 (2.1%)

All (19) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	405	DT	C5-C7	8.24	1.54	1.50
3	F	430	DT	C5-C7	7.75	1.54	1.50
1	A	592	GLU	CG-CD	7.53	1.63	1.51
1	A	529	GLU	CD-OE2	7.08	1.33	1.25
1	B	590	GLU	CG-CD	6.79	1.62	1.51
1	B	575	SER	CB-OG	6.60	1.50	1.42
3	E	429	DA	N7-C5	6.35	1.43	1.39
2	D	409	DC	N1-C6	6.25	1.40	1.37
2	D	408	DG	C8-N7	6.20	1.34	1.30
2	C	406	DG	C8-N7	5.79	1.34	1.30
2	D	408	DG	N7-C5	5.78	1.42	1.39
2	D	406	DG	N1-C2	5.73	1.42	1.37
2	D	409	DC	O4'-C1'	5.56	1.49	1.42
1	B	597	GLU	CG-CD	5.28	1.59	1.51
3	E	421	DG	C2-N2	-5.25	1.29	1.34
3	E	424	DA	C8-N7	5.22	1.35	1.31
3	E	431	DG	C4'-O4'	5.20	1.50	1.45
1	B	448	ARG	CG-CD	-5.04	1.39	1.51
1	A	529	GLU	CD-OE1	5.00	1.31	1.25

All (131) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	418	MET	CG-SD-CE	16.90	127.25	100.20
3	E	423	DC	O4'-C4'-C3'	-15.98	96.41	106.00
3	E	426	DC	C5-C6-N1	14.21	128.10	121.00
2	C	402	DC	O4'-C1'-N1	-11.35	100.06	108.00
2	D	406	DG	C5-C6-N1	11.16	117.08	111.50
3	E	426	DC	C6-N1-C2	-10.59	116.06	120.30
3	E	428	DC	C6-N1-C2	-10.50	116.10	120.30
3	E	430	DT	C6-C5-C7	-10.40	116.66	122.90
3	E	426	DC	C2-N3-C4	10.31	125.06	119.90
2	D	404	DA	O4'-C1'-N9	-9.77	101.16	108.00
3	E	421	DG	C6-N1-C2	-9.36	119.48	125.10
3	F	428	DC	O4'-C4'-C3'	-9.35	100.39	106.00
3	F	424	DA	O4'-C1'-N9	-9.19	101.57	108.00
2	D	411	DG	O5'-P-OP2	-9.17	97.45	105.70
3	E	422	DT	O4'-C1'-N1	-9.05	101.67	108.00
2	C	404	DA	O4'-C1'-N9	-8.92	101.76	108.00
3	E	421	DG	N9-C4-C5	8.90	108.96	105.40
3	E	423	DC	C1'-O4'-C4'	-8.84	101.26	110.10
2	D	411	DG	OP1-P-OP2	8.81	132.82	119.60
3	F	429	DA	O4'-C1'-N9	-8.59	101.98	108.00
3	E	421	DG	C4-C5-N7	-8.54	107.38	110.80
1	B	511[A]	ARG	NE-CZ-NH2	-8.38	116.11	120.30
1	B	511[B]	ARG	NE-CZ-NH2	-8.38	116.11	120.30
2	D	412	DA	N1-C2-N3	-8.35	125.12	129.30
3	E	421	DG	N1-C2-N3	8.29	128.87	123.90
3	F	427	DG	O4'-C1'-N9	-8.09	102.34	108.00
3	E	426	DC	O4'-C1'-N1	-7.86	102.50	108.00
3	E	421	DG	N3-C2-N2	-7.86	114.40	119.90
3	E	430	DT	C5-C6-N1	-7.77	119.04	123.70
3	E	430	DT	C6-N1-C2	7.63	125.12	121.30
3	F	430	DT	O4'-C1'-N1	-7.21	102.95	108.00
3	E	426	DC	C4-C5-C6	-7.19	113.81	117.40
1	B	511[A]	ARG	NE-CZ-NH1	7.18	123.89	120.30
1	B	511[B]	ARG	NE-CZ-NH1	7.18	123.89	120.30
3	E	432	DG	C5-C6-O6	-7.11	124.34	128.60
3	E	422	DT	N3-C4-O4	7.09	124.15	119.90
2	C	410	DT	C4-C5-C7	7.09	123.25	119.00
3	E	421	DG	N1-C6-O6	-6.98	115.71	119.90
3	F	426	DC	O4'-C1'-N1	-6.92	103.16	108.00
2	D	412	DA	O4'-C1'-N9	-6.91	103.16	108.00
3	E	424	DA	C4-C5-N7	6.87	114.13	110.70
3	E	421	DG	C8-N9-C4	-6.83	103.67	106.40

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	C	403	DC	C3'-C2'-C1'	-6.83	94.30	102.50
2	C	409	DC	C1'-O4'-C4'	-6.81	103.29	110.10
3	F	425	DG	N7-C8-N9	-6.79	109.70	113.10
2	C	403	DC	N3-C4-C5	-6.78	119.19	121.90
2	D	403	DC	N1-C2-O2	-6.71	114.87	118.90
3	E	422	DT	P-O3'-C3'	6.69	127.73	119.70
3	E	430	DT	C4-C5-C7	6.69	123.01	119.00
3	E	432	DG	N1-C6-O6	6.66	123.89	119.90
3	F	432	DG	N7-C8-N9	-6.66	109.77	113.10
3	E	426	DC	O4'-C4'-C3'	-6.63	101.85	104.50
2	C	403	DC	O4'-C4'-C3'	-6.57	101.87	104.50
3	E	422	DT	O4'-C4'-C3'	6.54	109.92	106.00
3	F	428	DC	O4'-C1'-C2'	6.53	111.13	105.90
2	C	406	DG	O4'-C4'-C3'	-6.53	101.89	104.50
3	E	431	DG	N1-C2-N3	-6.52	119.99	123.90
2	D	406	DG	C2-N3-C4	6.47	115.13	111.90
2	D	408	DG	C5-C6-N1	-6.43	108.28	111.50
3	E	423	DC	C2-N3-C4	6.43	123.11	119.90
3	E	424	DA	N1-C2-N3	-6.41	126.09	129.30
3	E	430	DT	C1'-O4'-C4'	-6.41	103.69	110.10
2	D	411	DG	O4'-C1'-C2'	6.38	111.00	105.90
3	E	431	DG	C8-N9-C4	6.33	108.93	106.40
3	F	430	DT	C1'-O4'-C4'	-6.33	103.77	110.10
3	E	426	DC	N3-C4-C5	-6.30	119.38	121.90
2	C	406	DG	C1'-O4'-C4'	-6.30	103.80	110.10
3	E	429	DA	O4'-C1'-N9	-6.25	103.62	108.00
3	F	427	DG	C6-N1-C2	-6.20	121.38	125.10
3	F	432	DG	N1-C6-O6	-6.16	116.21	119.90
3	E	423	DC	C4'-C3'-C2'	-6.14	97.58	103.10
2	D	409	DC	O4'-C4'-C3'	-6.13	102.05	104.50
3	F	430	DT	C6-C5-C7	-6.10	119.24	122.90
3	F	425	DG	C8-N9-C4	6.06	108.82	106.40
3	F	431	DG	C6-C5-N7	6.00	134.00	130.40
2	C	410	DT	P-O3'-C3'	5.98	126.88	119.70
3	E	424	DA	O5'-P-OP1	-5.96	100.34	105.70
2	D	405	DT	C4-C5-C7	5.93	122.56	119.00
2	C	406	DG	OP1-P-OP2	5.92	128.48	119.60
2	C	404	DA	OP1-P-OP2	5.89	128.44	119.60
3	E	431	DG	C2-N3-C4	5.84	114.82	111.90
2	D	410	DT	P-O3'-C3'	5.82	126.68	119.70
3	E	428	DC	N1-C2-O2	-5.79	115.43	118.90
2	D	406	DG	C5-C6-O6	-5.78	125.13	128.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	409	DC	C1'-O4'-C4'	-5.78	104.32	110.10
2	D	405	DT	N3-C4-O4	-5.71	116.47	119.90
3	F	427	DG	C5-C6-N1	5.69	114.35	111.50
2	C	404	DA	O5'-P-OP1	-5.68	100.58	105.70
3	E	428	DC	OP1-P-OP2	5.64	128.06	119.60
3	F	431	DG	C5-C6-O6	-5.63	125.22	128.60
3	E	430	DT	O4'-C1'-N1	-5.56	104.11	108.00
2	D	402	DC	C5-C6-N1	5.55	123.78	121.00
2	D	406	DG	C1'-O4'-C4'	-5.54	104.56	110.10
3	E	431	DG	C5-C6-N1	5.52	114.26	111.50
3	F	427	DG	C5-N7-C8	5.52	107.06	104.30
2	C	412	DA	C3'-C2'-C1'	-5.52	95.88	102.50
2	C	405	DT	O4'-C1'-N1	-5.50	104.15	108.00
1	A	503	ASP	CB-CG-OD1	5.47	123.23	118.30
1	A	600	ASP	CB-CG-OD1	-5.45	113.39	118.30
3	F	430	DT	C4'-C3'-C2'	5.44	108.00	103.10
2	C	403	DC	O3'-P-O5'	-5.44	93.67	104.00
3	E	425	DG	C3'-C2'-C1'	-5.42	96.00	102.50
3	E	425	DG	C5-C6-O6	-5.41	125.36	128.60
3	F	424	DA	N7-C8-N9	-5.39	111.10	113.80
3	F	432	DG	C5-N7-C8	5.39	107.00	104.30
2	D	406	DG	C4-C5-C6	-5.38	115.57	118.80
2	D	409	DC	O4'-C1'-N1	-5.37	104.24	108.00
3	E	428	DC	C5-C6-N1	5.36	123.68	121.00
3	E	429	DA	N1-C6-N6	-5.35	115.39	118.60
3	E	431	DG	N7-C8-N9	-5.35	110.42	113.10
3	E	428	DC	P-O3'-C3'	5.34	126.11	119.70
2	D	406	DG	OP1-P-OP2	5.33	127.59	119.60
3	E	426	DC	O4'-C1'-C2'	5.32	110.16	105.90
1	B	474	ASP	CB-CG-OD2	5.32	123.08	118.30
2	D	402	DC	O4'-C1'-C2'	5.28	110.13	105.90
3	F	426	DC	O4'-C4'-C3'	-5.26	102.39	104.50
3	F	432	DG	C6-N1-C2	-5.22	121.97	125.10
2	D	413	DC	N3-C4-N4	-5.20	114.36	118.00
3	F	430	DT	O4'-C4'-C3'	-5.19	102.43	104.50
2	C	403	DC	C4'-C3'-C2'	5.15	107.74	103.10
3	E	423	DC	P-O3'-C3'	5.14	125.87	119.70
3	F	427	DG	N7-C8-N9	-5.13	110.53	113.10
3	E	428	DC	O4'-C4'-C3'	-5.11	102.45	104.50
3	E	427	DG	O4'-C1'-N9	-5.11	104.43	108.00
3	E	425	DG	N3-C2-N2	-5.09	116.34	119.90
2	D	412	DA	O5'-P-OP2	5.09	116.81	110.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	421	DG	C5-N7-C8	5.09	106.84	104.30
2	D	404	DA	C8-N9-C4	-5.07	103.77	105.80
2	D	412	DA	C2-N3-C4	5.05	113.12	110.60
3	F	431	DG	C2-N3-C4	5.05	114.42	111.90
1	A	436	ARG	NE-CZ-NH1	5.00	122.80	120.30

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	1698	0	1618	22	0
1	B	1651	0	1566	17	0
2	C	283	0	162	1	0
2	D	241	0	138	0	0
3	E	246	0	134	2	0
3	F	246	0	136	3	0
4	A	32	0	48	7	0
4	B	20	0	30	8	0
4	C	4	0	4	0	0
4	D	4	0	4	0	0
4	E	8	0	11	0	0
4	F	20	0	29	3	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
5	E	1	0	0	0	0
5	F	1	0	0	0	0
6	A	4	0	0	4	0
6	B	12	0	0	5	0
6	C	6	0	0	5	0
7	A	257	0	0	11	0
7	B	269	0	0	8	0
7	C	43	0	0	1	0
7	D	40	0	0	0	0
7	E	37	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	F	34	0	0	1	0
All	All	5159	0	3880	60	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 7.

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)
6:B:215:UNL:O2	6:B:215:UNL:O1	1.71	1.08
6:C:456:UNL:O1	6:C:456:UNL:O2	1.71	1.07
1:A:546:GLY:HA2	7:B:853:HOH:O	1.52	1.06
1:B:479:ASN:ND2	6:B:655:UNL:O1	1.94	1.00
6:A:638:UNL:O1	6:A:638:UNL:O2	1.86	0.94
6:C:637:UNL:O1	6:C:637:UNL:O2	1.91	0.89
3:E:423:DC:H2''	3:E:424:DA:O5'	1.74	0.86
1:A:532[B]:ARG:HB2	4:A:636:EDO:H21	1.56	0.84
1:B:572[B]:ARG:HD3	1:B:576:GLY:O	1.76	0.84
6:B:322:UNL:O2	6:B:322:UNL:O1	1.99	0.81
3:E:423:DC:H2'	3:E:424:DA:C8	2.18	0.79
1:B:572[B]:ARG:HG2	4:B:633:EDO:H12	1.65	0.77
7:A:212:HOH:O	4:B:633:EDO:H21	1.86	0.75
4:F:513:EDO:H11	7:F:327:HOH:O	1.86	0.74
1:A:601:ARG:NH2	7:A:399:HOH:O	2.22	0.73
1:B:476:ASP:HB2	4:B:630:EDO:H22	1.76	0.67
6:C:359:UNL:O1	6:C:359:UNL:O2	2.16	0.64
1:A:472:GLU:O	4:B:633:EDO:H11	1.97	0.64
1:B:572[A]:ARG:CZ	7:B:253:HOH:O	2.46	0.64
1:B:472:GLU:CG	7:B:852:HOH:O	2.46	0.63
1:A:601:ARG:NH2	1:A:601:ARG:HG2	2.12	0.62
3:F:425:DG:O6	4:F:513:EDO:H12	1.98	0.62
1:B:572[A]:ARG:NH2	7:B:253:HOH:O	2.33	0.62
6:B:636:UNL:O1	6:B:636:UNL:O2	2.18	0.62
1:A:601:ARG:HG2	1:A:601:ARG:HH21	1.64	0.62
6:C:456:UNL:O1	6:C:637:UNL:O2	2.18	0.61
1:A:532[A]:ARG:HB2	4:A:636:EDO:H21	1.84	0.59
1:A:544:LYS:O	7:A:653:HOH:O	2.16	0.59
1:A:581:ARG:NH2	1:B:572[A]:ARG:HH12	2.02	0.58
1:A:545:GLY:HA2	7:A:700:HOH:O	2.04	0.57
4:A:630:EDO:C1	7:A:286:HOH:O	2.52	0.57
6:A:654:UNL:O1	1:B:572[A]:ARG:NH2	2.38	0.56

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:533[A]:GLN:HG3	4:A:636:EDO:H12	1.89	0.55
4:B:632:EDO:H22	7:C:18:HOH:O	2.06	0.55
1:A:532[B]:ARG:NH1	7:A:662:HOH:O	2.36	0.54
1:B:572[A]:ARG:HG2	4:B:633:EDO:H12	1.90	0.54
1:B:572[B]:ARG:CG	4:B:633:EDO:H12	2.37	0.54
4:A:630:EDO:H11	7:A:286:HOH:O	2.07	0.54
1:B:472:GLU:CD	7:B:852:HOH:O	2.46	0.54
1:A:423:PHE:CZ	1:A:535:LYS:HD3	2.42	0.53
3:F:424:DA:N7	4:F:513:EDO:H22	2.23	0.52
2:C:410:DT:H2'	2:C:411:DG:C8	2.45	0.52
1:A:443:GLU:HG2	7:A:756:HOH:O	2.10	0.52
1:A:546:GLY:CA	7:B:853:HOH:O	2.32	0.51
1:A:530:ASP:HB3	4:A:636:EDO:H12	1.94	0.50
1:B:573:GLY:O	4:B:633:EDO:H22	2.12	0.50
4:A:630:EDO:H12	7:A:286:HOH:O	2.13	0.47
6:B:636:UNL:O1	6:C:637:UNL:O2	2.32	0.47
1:B:527:GLU:OE1	7:B:821:HOH:O	2.20	0.47
1:A:420:ALA:HA	1:A:448[B]:ARG:CZ	2.47	0.45
1:A:532[B]:ARG:NH2	7:A:662:HOH:O	2.48	0.45
1:A:597:GLU:HB2	7:A:803:HOH:O	2.20	0.41
1:A:453:GLY:O	1:A:465:LEU:HA	2.20	0.41
1:B:594:TRP:O	1:B:599:LYS:HE3	2.21	0.41
1:B:506:LEU:HD23	1:B:523[B]:GLU:HG3	2.02	0.41
6:A:654:UNL:O2	6:A:654:UNL:O1	2.39	0.40
1:B:472:GLU:HG2	7:B:852:HOH:O	2.15	0.40
1:A:518:HIS:HD2	6:A:638:UNL:O2	2.04	0.40
1:A:465:LEU:C	1:A:465:LEU:HD12	2.42	0.40
3:F:428:DC:H2'	3:F:429:DA:C8	2.56	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	215/212 (101%)	211 (98%)	4 (2%)	0	100	100
1	B	208/212 (98%)	205 (99%)	3 (1%)	0	100	100
All	All	423/424 (100%)	416 (98%)	7 (2%)	0	100	100

There are no Ramachandran outliers to report.

5.3.2 Protein sidechains ⓘ

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	172/173 (99%)	172 (100%)	0	100	100
1	B	166/173 (96%)	165 (99%)	1 (1%)	86	69
All	All	338/346 (98%)	337 (100%)	1 (0%)	92	81

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	596	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	518	HIS

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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	5CM	C	407[B]	2	15,21,22	0.80	0	19,30,33	1.34	2 (10%)
2	5CM	D	407	2	15,21,22	1.09	0	19,30,33	1.75	2 (10%)
2	5CM	C	407[A]	2	15,21,22	0.98	1 (6%)	19,30,33	1.38	3 (15%)

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	5CM	C	407[B]	2	-	1/4/21/22	0/2/2/2
2	5CM	D	407	2	-	1/4/21/22	0/2/2/2
2	5CM	C	407[A]	2	-	3/4/21/22	0/2/2/2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	407[A]	5CM	O4'-C1'	2.30	1.47	1.42

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	407	5CM	C2-N3-C4	5.37	122.50	116.02
2	C	407[B]	5CM	C2-N3-C4	4.00	120.84	116.02
2	C	407[A]	5CM	C2-N3-C4	3.91	120.74	116.02
2	C	407[A]	5CM	C2'-C1'-N1	2.72	120.55	114.27
2	C	407[B]	5CM	C5-C6-N1	-2.39	119.62	122.19
2	C	407[A]	5CM	C5-C6-N1	-2.26	119.76	122.19
2	D	407	5CM	C2'-C1'-N1	2.10	119.12	114.27

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	407	5CM	O4'-C1'-N1-C6
2	C	407[B]	5CM	O4'-C1'-N1-C6
2	C	407[A]	5CM	O4'-C1'-N1-C6
2	C	407[A]	5CM	O4'-C4'-C5'-O5'
2	C	407[A]	5CM	C3'-C4'-C5'-O5'

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 35 ligands modelled in this entry, 9 are unknown and 4 are monoatomic - leaving 22 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	F	513	-	3,3,3	0.57	0	2,2,2	0.61	0
4	EDO	A	636	-	3,3,3	0.24	0	2,2,2	4.10	2 (100%)
4	EDO	B	630	-	3,3,3	0.93	0	2,2,2	0.70	0
4	EDO	A	634	-	3,3,3	0.69	0	2,2,2	0.33	0
4	EDO	D	502	5	3,3,3	1.06	0	2,2,2	0.09	0
4	EDO	E	501	5	3,3,3	0.50	0	2,2,2	0.43	0
4	EDO	B	632	-	3,3,3	0.56	0	2,2,2	0.90	0
4	EDO	E	503	-	3,3,3	0.55	0	2,2,2	0.73	0
4	EDO	F	506	5	3,3,3	0.50	0	2,2,2	0.89	0
4	EDO	F	509	-	3,3,3	0.53	0	2,2,2	0.99	0
4	EDO	A	635	-	3,3,3	0.49	0	2,2,2	0.31	0
4	EDO	A	631	-	3,3,3	0.85	0	2,2,2	0.98	0
4	EDO	A	629	-	3,3,3	0.21	0	2,2,2	1.96	1 (50%)
4	EDO	C	505	5	3,3,3	0.78	0	2,2,2	0.21	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	EDO	A	633	-	3,3,3	0.34	0	2,2,2	0.64	0
4	EDO	A	632	-	3,3,3	0.45	0	2,2,2	0.58	0
4	EDO	F	510	-	3,3,3	1.18	0	2,2,2	0.93	0
4	EDO	B	633	-	3,3,3	0.85	0	2,2,2	0.98	0
4	EDO	F	511	-	3,3,3	0.91	0	2,2,2	0.06	0
4	EDO	B	629	-	3,3,3	1.53	1 (33%)	2,2,2	0.57	0
4	EDO	B	631	-	3,3,3	0.39	0	2,2,2	0.96	0
4	EDO	A	630	-	3,3,3	0.57	0	2,2,2	0.43	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	F	513	-	-	0/1/1/1	-
4	EDO	A	636	-	-	1/1/1/1	-
4	EDO	B	630	-	-	1/1/1/1	-
4	EDO	A	634	-	-	1/1/1/1	-
4	EDO	D	502	5	-	0/1/1/1	-
4	EDO	E	501	5	-	0/1/1/1	-
4	EDO	B	632	-	-	1/1/1/1	-
4	EDO	E	503	-	-	0/1/1/1	-
4	EDO	F	506	5	-	0/1/1/1	-
4	EDO	F	509	-	-	1/1/1/1	-
4	EDO	A	635	-	-	0/1/1/1	-
4	EDO	A	631	-	-	1/1/1/1	-
4	EDO	A	629	-	-	0/1/1/1	-
4	EDO	C	505	5	-	0/1/1/1	-
4	EDO	A	633	-	-	0/1/1/1	-
4	EDO	A	632	-	-	0/1/1/1	-
4	EDO	F	510	-	-	1/1/1/1	-
4	EDO	B	633	-	-	0/1/1/1	-
4	EDO	F	511	-	-	1/1/1/1	-
4	EDO	B	629	-	-	0/1/1/1	-
4	EDO	B	631	-	-	0/1/1/1	-
4	EDO	A	630	-	-	1/1/1/1	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	629	EDO	O2-C2	2.26	1.53	1.42

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	636	EDO	O2-C2-C1	5.09	148.55	111.91
4	A	636	EDO	O1-C1-C2	-2.77	92.00	111.91
4	A	629	EDO	O1-C1-C2	-2.76	92.02	111.91

There are no chirality outliers.

All (9) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	B	632	EDO	O1-C1-C2-O2
4	F	511	EDO	O1-C1-C2-O2
4	A	631	EDO	O1-C1-C2-O2
4	B	630	EDO	O1-C1-C2-O2
4	A	636	EDO	O1-C1-C2-O2
4	F	509	EDO	O1-C1-C2-O2
4	A	630	EDO	O1-C1-C2-O2
4	A	634	EDO	O1-C1-C2-O2
4	F	510	EDO	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 18 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	F	513	EDO	3	0
4	A	636	EDO	4	0
4	B	630	EDO	1	0
4	B	632	EDO	1	0
4	B	633	EDO	6	0
4	A	630	EDO	3	0

5.7 Other polymers [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
2	C	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	C	406:DG	O3'	407[A]:5CM	P	1.76
1	C	406:DG	O3'	407[B]:5CM	P	1.75

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, 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	A	210/212 (99%)	0.17	14 (6%) 17 15	8, 15, 33, 64	0
1	B	206/212 (97%)	0.07	9 (4%) 34 34	7, 13, 30, 62	0
2	C	11/12 (91%)	0.20	1 (9%) 9 8	12, 17, 48, 56	0
2	D	11/12 (91%)	-0.24	0 100 100	12, 21, 29, 30	0
3	E	12/12 (100%)	0.00	0 100 100	15, 23, 33, 35	0
3	F	12/12 (100%)	0.10	1 (8%) 11 9	13, 21, 42, 42	0
All	All	462/472 (97%)	0.11	25 (5%) 25 23	7, 14, 35, 64	0

All (25) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	619	LEU	8.4
1	B	620	ALA	7.9
1	A	626	ARG	6.7
1	B	621	ASN	6.7
1	A	417	HIS	5.5
1	B	616	LEU	5.4
1	B	618	ALA	5.1
1	A	625	SER	4.8
1	A	624	LYS	3.7
1	A	418	MET	3.5
1	B	614	GLY	3.4
1	A	623	GLU	3.2
1	A	545	GLY	3.1
1	A	619	LEU	3.1
1	B	622	LYS	3.0
1	B	596	ARG	2.8
1	A	620	ALA	2.6
1	A	530	ASP	2.6
1	A	617	GLU	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
3	F	421	DG	2.4
1	A	621	ASN	2.2
2	C	413	DC	2.2
1	A	613	GLU	2.1
1	B	617	GLU	2.1
1	A	622	LYS	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, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å ²)	Q<0.9
2	5CM	C	407[B]	20/21	0.98	0.10	5,8,11,11	20
2	5CM	C	407[A]	20/21	0.98	0.10	4,7,9,9	20
2	5CM	D	407	20/21	0.99	0.09	7,9,11,12	0

6.3 Carbohydrates [i](#)

There are no carbohydrates 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
4	EDO	B	633	4/4	0.51	0.36	30,34,34,39	0
6	UNL	B	655	5/-	0.68	0.27	33,33,33,33	0
4	EDO	B	631	4/4	0.74	0.28	45,45,49,51	0
6	UNL	B	215	2/-	0.75	0.28	29,29,29,29	0
6	UNL	A	654	2/-	0.75	0.14	31,31,31,40	0
4	EDO	F	513	4/4	0.77	0.22	34,41,44,46	0
4	EDO	B	632	4/4	0.80	0.14	25,31,40,50	0
4	EDO	B	629	4/4	0.82	0.14	22,27,28,29	0
4	EDO	A	635	4/4	0.84	0.15	47,49,50,52	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	EDO	A	636	4/4	0.85	0.23	23,27,36,38	0
4	EDO	F	510	4/4	0.86	0.12	22,23,24,29	0
4	EDO	B	630	4/4	0.89	0.15	20,35,35,38	0
4	EDO	A	632	4/4	0.89	0.15	30,32,40,41	0
4	EDO	A	634	4/4	0.91	0.18	25,27,30,32	0
4	EDO	F	509	4/4	0.92	0.08	23,25,32,34	0
6	UNL	B	322	3/-	0.92	0.12	30,30,35,42	0
6	UNL	C	456	2/-	0.93	0.19	34,34,34,34	0
6	UNL	A	638	2/-	0.93	0.12	31,31,31,34	0
4	EDO	D	502	4/4	0.94	0.09	18,19,21,21	0
4	EDO	A	630	4/4	0.94	0.16	20,33,37,42	0
4	EDO	E	503	4/4	0.95	0.07	31,34,37,37	0
6	UNL	C	637	2/-	0.95	0.24	29,29,29,32	0
4	EDO	A	631	4/4	0.96	0.10	17,18,31,39	0
4	EDO	A	633	4/4	0.96	0.14	31,36,38,38	0
4	EDO	A	629	4/4	0.96	0.11	19,24,27,36	0
6	UNL	C	359	2/-	0.96	0.08	27,27,27,38	0
6	UNL	B	636	2/-	0.97	0.06	26,26,26,32	0
4	EDO	F	511	4/4	0.97	0.05	23,26,30,32	0
4	EDO	F	506	4/4	0.98	0.08	12,14,15,17	0
4	EDO	C	505	4/4	0.98	0.07	13,13,14,16	0
4	EDO	E	501	4/4	0.99	0.05	15,18,19,21	0
5	NA	A	637	1/1	1.00	0.19	11,11,11,11	0
5	NA	B	634	1/1	1.00	0.21	9,9,9,9	0
5	NA	F	503	1/1	1.00	0.06	12,12,12,12	0
5	NA	E	504	1/1	1.00	0.04	16,16,16,16	0

6.5 Other polymers ⓘ

There are no such residues in this entry.