Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 09:18 PM GMT

PDB ID	:	1OE5
Title	:	Xenopus SMUG1, an anti-mutator uracil-DNA Glycosylase
Authors	:	Wibley, J.E.A.; Pearl, L.H.
Deposited on	:	2003-03-19
Resolution	:	2.30 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

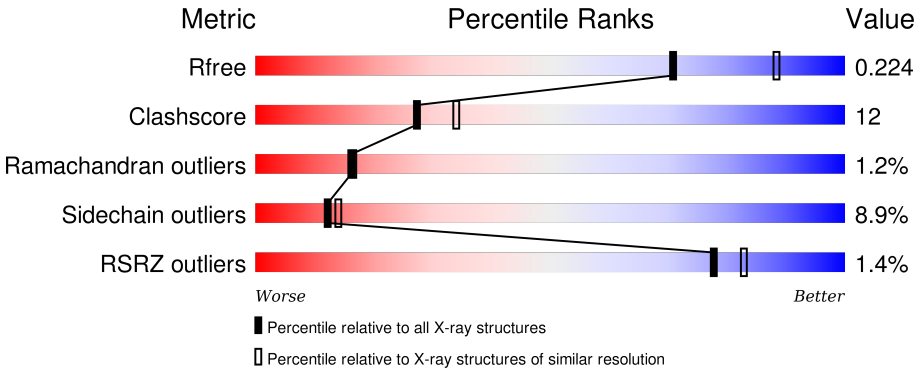
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 2.30 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}	91344	3852 (2.30-2.30)
Clashscore	102246	4452 (2.30-2.30)
Ramachandran outliers	100387	4410 (2.30-2.30)
Sidechain outliers	100360	4409 (2.30-2.30)
RSRZ outliers	91569	3857 (2.30-2.30)

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. 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	247	71% 22%
1	B	247	2% 70% 22% 5%
2	E	12	50% 50%
3	F	13	23% 8% 69% 23%

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	URA	A	1282	-	X	-	-
5	GOL	A	1283	-	-	-	X
6	IPA	A	1284	-	-	-	X
6	IPA	B	1281	-	-	-	X
7	EPE	A	1285	-	-	-	X
7	EPE	B	1282	-	-	-	X
8	DUR	F	1405	-	-	X	X

2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4657 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 SINGLE-STRAND SELECTIVE MONOFUNCTIONAL URACIL DNA GLYCOSYLASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	244	Total	C	N	O	S	0	2	1
			1934	1241	338	342	13			
1	B	245	Total	C	N	O	S	0	0	0
			1939	1244	339	344	12			

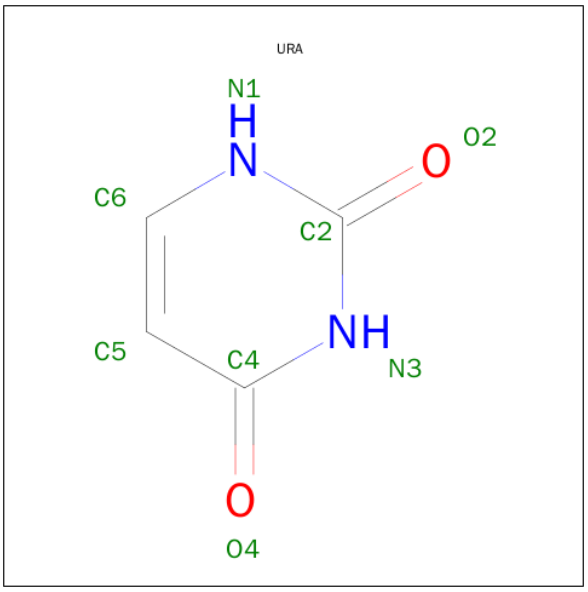
- Molecule 2 is a DNA chain called 5'-D(*CP*CP*CP*GP*TP*GP*AP*GP*TP*CP*CP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	12	Total	C	N	O	P	0	0	0
			241	115	44	71	11			

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

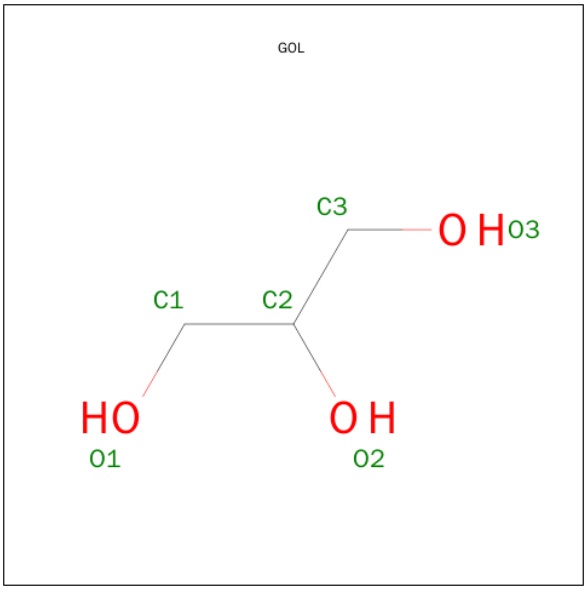
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	F	13	Total	C	N	O	P	0	0	0
			248	117	46	73	12			

- Molecule 4 is URACIL (three-letter code: URA) (formula: $C_4H_4N_2O_2$).



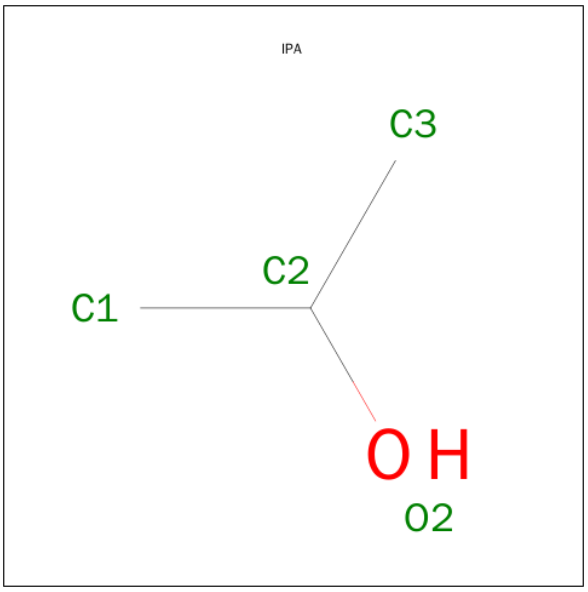
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 8 4 2 2	0	0

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



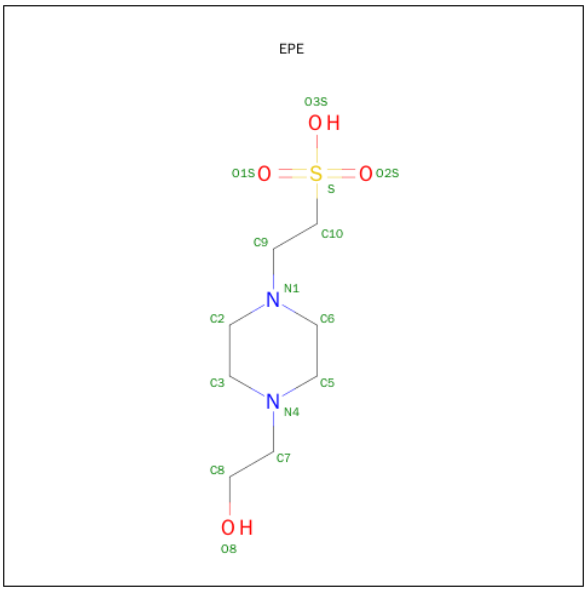
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 6 3 3	0	0

- Molecule 6 is ISOPROPYL ALCOHOL (three-letter code: IPA) (formula: C_3H_8O).



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

- Molecule 7 is 4-(2-HYDROXYETHYL)-1-PIPERAZINE ETHANESULFONIC ACID (three-letter code: EPE) (formula: $C_8H_{18}N_2O_4S$).



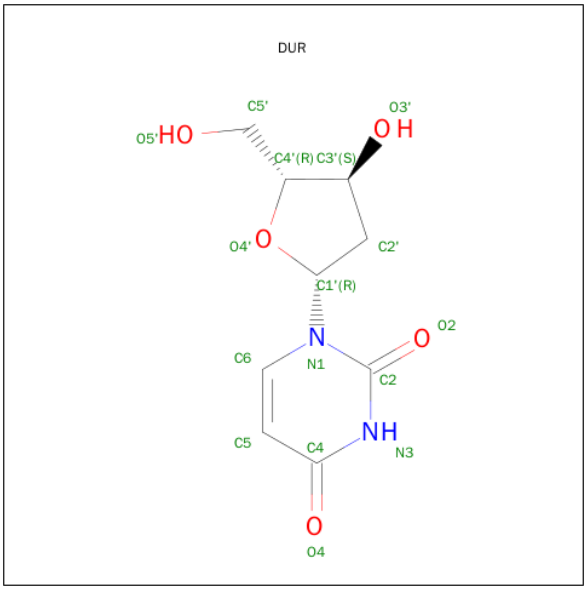
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C N O S 15 8 2 4 1	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
7	B	1	Total	C	N	O	S	0	0
			15	8	2	4	1		

- Molecule 8 is 2'-DEOXYURIDINE (three-letter code: DUR) (formula: $C_9H_{12}N_2O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
8	F	1	Total	C	N	O	0	0
			16	9	2	5		

- Molecule 9 is water.

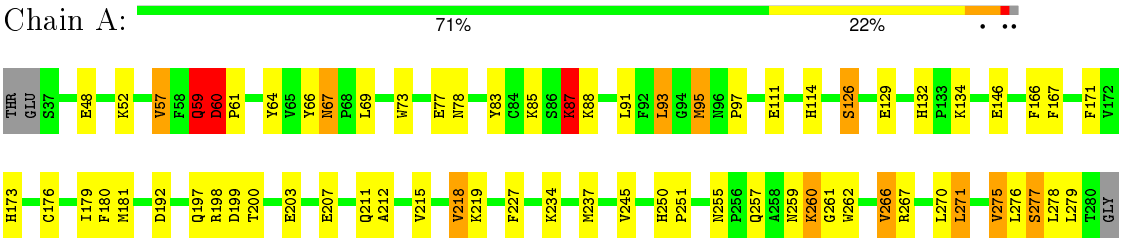
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	123	Total	O	0	0
			123	123		
9	B	82	Total	O	0	0
			82	82		
9	E	12	Total	O	0	0
			12	12		
9	F	10	Total	O	0	0
			10	10		

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 errors 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.

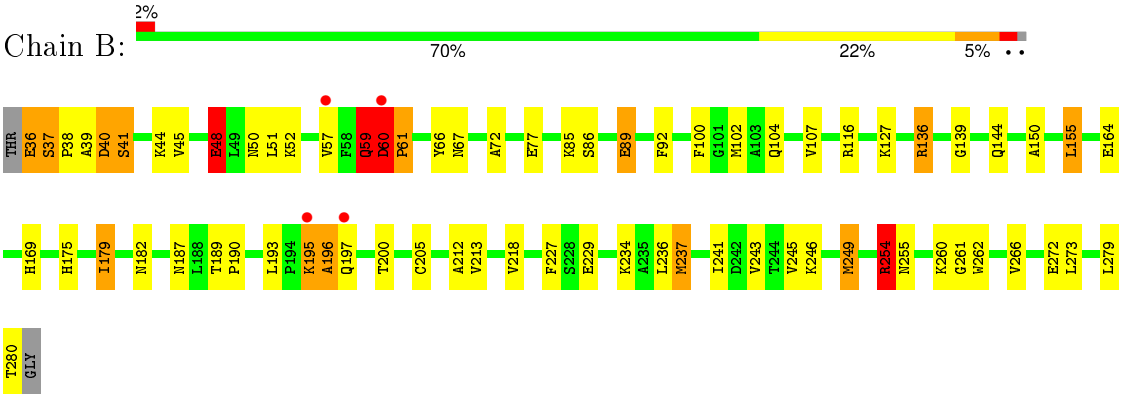
• Molecule 1: SINGLE-STRAND SELECTIVE MONOFUNCTIONAL URACIL DNA GLYCOSYLASE

Chain A:



• Molecule 1: SINGLE-STRAND SELECTIVE MONOFUNCTIONAL URACIL DNA GLYCOSYLASE

Chain B:

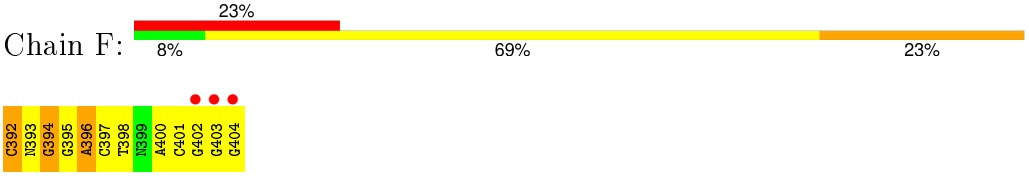


• Molecule 2: 5'-D(*CP*CP*CP*GP*TP*GP*AP*GP*TP*CP*CP*G)-3'

Chain E:

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

Chain F:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	120.60Å 85.65Å 78.47Å 90.00° 118.56° 90.00°	Depositor
Resolution (Å)	69.01 – 2.30 42.83 – 2.30	Depositor EDS
% Data completeness (in resolution range)	95.8 (69.01-2.30) 95.8 (42.83-2.30)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.97 (at 2.29Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.163 , 0.229 0.164 , 0.224	Depositor DCC
R_{free} test set	1512 reflections (5.32%)	DCC
Wilson B-factor (Å^2)	31.2	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 52.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 29919 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4657	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.06% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, IPA, URA, 3DR, DUR, EPE

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	1.48	16/1993 (0.8%)	1.20	13/2701 (0.5%)
1	B	1.36	8/1990 (0.4%)	1.12	7/2698 (0.3%)
2	E	2.49	17/269 (6.3%)	3.47	53/413 (12.8%)
3	F	2.52	12/252 (4.8%)	3.18	45/383 (11.7%)
All	All	1.59	53/4504 (1.2%)	1.61	118/6195 (1.9%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	2
1	B	0	1
All	All	0	3

All (53) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	212	ALA	CA-CB	7.99	1.69	1.52
3	F	402	DG	C3'-O3'	7.95	1.54	1.44
3	F	394	DG	C8-N7	7.69	1.35	1.30
1	A	245	VAL	CB-CG1	7.42	1.68	1.52
2	E	290	DC	P-O5'	7.34	1.67	1.59
2	E	282	DC	P-O5'	7.33	1.67	1.59
1	B	254	ARG	CG-CD	7.23	1.70	1.51
3	F	392	DC	N3-C4	7.20	1.39	1.33
2	E	282	DC	O5'-C5'	7.17	1.60	1.42
1	B	212	ALA	CA-CB	6.91	1.67	1.52
1	A	95	MET	CG-SD	6.44	1.97	1.81
3	F	395	DG	C6-N1	6.35	1.44	1.39

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	398	DT	C1'-N1	6.22	1.57	1.49
2	E	291	DC	O5'-C5'	-6.18	1.26	1.42
1	A	60	ASP	C-O	6.09	1.34	1.23
2	E	282	DC	C2-N3	6.08	1.40	1.35
2	E	292	DG	C5-C4	-6.07	1.34	1.38
3	F	395	DG	C8-N7	6.06	1.34	1.30
2	E	282	DC	C5'-C4'	6.05	1.58	1.51
1	A	60	ASP	CB-CG	6.02	1.64	1.51
1	B	89	GLU	CD-OE2	6.00	1.32	1.25
1	A	83	TYR	CE1-CZ	5.87	1.46	1.38
2	E	292	DG	P-O5'	5.87	1.65	1.59
3	F	400	DA	C6-N6	5.84	1.38	1.33
1	A	73	TRP	CG-CD1	5.82	1.44	1.36
3	F	397	DC	C5-C6	5.82	1.39	1.34
3	F	398	DT	N1-C6	5.79	1.42	1.38
2	E	289	DT	C1'-N1	5.70	1.56	1.49
3	F	394	DG	C3'-O3'	-5.68	1.36	1.44
1	A	203	GLU	CD-OE1	5.67	1.31	1.25
3	F	392	DC	C1'-N1	5.65	1.56	1.49
1	B	72	ALA	CA-CB	5.60	1.64	1.52
2	E	284	DG	N3-C4	5.59	1.39	1.35
1	A	87	LYS	CD-CE	5.59	1.65	1.51
1	A	166	PHE	CB-CG	5.55	1.60	1.51
1	A	180	PHE	CE1-CZ	5.54	1.47	1.37
2	E	281	DC	N3-C4	5.53	1.37	1.33
1	B	164	GLU	CD-OE1	5.49	1.31	1.25
1	A	167	PHE	CG-CD2	5.36	1.46	1.38
1	B	150	ALA	CA-CB	5.31	1.63	1.52
1	A	260	LYS	CB-CG	5.29	1.66	1.52
3	F	404	DG	P-O5'	5.25	1.65	1.59
2	E	282	DC	C1'-N1	5.21	1.56	1.49
1	B	48	GLU	CD-OE2	5.17	1.31	1.25
1	B	92	PHE	CE2-CZ	5.16	1.47	1.37
2	E	289	DT	C5-C6	5.14	1.38	1.34
2	E	291	DC	C4-N4	5.13	1.38	1.33
2	E	289	DT	N1-C6	5.11	1.41	1.38
2	E	291	DC	C2-O2	5.09	1.29	1.24
1	A	266	VAL	CA-CB	5.02	1.65	1.54
1	A	57	VAL	CB-CG1	5.01	1.63	1.52
1	A	275	VAL	CB-CG2	-5.00	1.42	1.52
2	E	281	DC	C2-O2	5.00	1.28	1.24

All (118) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	281	DC	C4-C5-C6	-15.01	109.89	117.40
2	E	291	DC	O5'-P-OP1	-14.60	92.56	105.70
3	F	392	DC	O4'-C1'-N1	12.61	116.83	108.00
2	E	281	DC	C5-C6-N1	11.70	126.85	121.00
2	E	281	DC	N3-C4-C5	10.99	126.30	121.90
2	E	291	DC	C5'-C4'-C3'	-10.81	94.65	114.10
3	F	396	DA	C5-N7-C8	10.57	109.19	103.90
2	E	291	DC	O4'-C1'-N1	10.42	115.29	108.00
3	F	402	DG	P-O3'-C3'	10.39	132.17	119.70
2	E	284	DG	N3-C2-N2	9.85	126.80	119.90
2	E	291	DC	O4'-C1'-C2'	9.66	113.63	105.90
3	F	402	DG	O4'-C1'-N9	9.62	114.74	108.00
3	F	398	DT	O4'-C1'-N1	9.51	114.66	108.00
2	E	286	DG	O4'-C1'-N9	9.48	114.64	108.00
3	F	400	DA	P-O3'-C3'	9.27	130.83	119.70
3	F	397	DC	C5-C4-N4	9.25	126.68	120.20
2	E	282	DC	O4'-C1'-N1	9.25	114.47	108.00
2	E	291	DC	OP1-P-OP2	9.20	133.39	119.60
3	F	396	DA	C4-C5-N7	-8.81	106.29	110.70
2	E	281	DC	C5-C4-N4	-8.81	114.04	120.20
2	E	285	DT	C5-C4-O4	-8.79	118.75	124.90
1	A	192	ASP	CB-CG-OD2	8.78	126.20	118.30
3	F	397	DC	N3-C4-N4	-8.78	111.86	118.00
3	F	394	DG	C6-N1-C2	-8.77	119.84	125.10
2	E	282	DC	O4'-C4'-C3'	-8.76	100.74	106.00
2	E	290	DC	O4'-C1'-N1	-8.75	101.88	108.00
2	E	285	DT	O4'-C1'-N1	8.73	114.11	108.00
3	F	403	DG	P-O3'-C3'	8.70	130.13	119.70
3	F	394	DG	N9-C4-C5	8.63	108.85	105.40
3	F	396	DA	N7-C8-N9	-8.55	109.53	113.80
3	F	400	DA	N9-C4-C5	8.54	109.22	105.80
2	E	282	DC	O5'-P-OP1	-8.45	98.09	105.70
2	E	291	DC	O5'-P-OP2	8.35	120.72	110.70
3	F	392	DC	N3-C4-N4	8.20	123.74	118.00
3	F	396	DA	N1-C6-N6	-8.17	113.70	118.60
2	E	291	DC	O4'-C4'-C3'	8.09	110.85	106.00
2	E	292	DG	N1-C2-N2	8.03	123.43	116.20
2	E	282	DC	O5'-P-OP2	8.02	120.33	110.70
3	F	392	DC	C5-C4-N4	-8.02	114.58	120.20
1	A	198	ARG	NE-CZ-NH2	-7.80	116.40	120.30
3	F	400	DA	C5-N7-C8	7.22	107.51	103.90
2	E	282	DC	OP1-P-OP2	7.17	130.36	119.60
2	E	282	DC	OP2-P-O3'	7.16	120.95	105.20

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	396	DA	N9-C4-C5	7.10	108.64	105.80
3	F	397	DC	O4'-C1'-N1	7.10	112.97	108.00
2	E	283	DC	O4'-C1'-C2'	-7.08	100.23	105.90
1	B	136	ARG	NE-CZ-NH1	7.05	123.83	120.30
2	E	292	DG	N3-C2-N2	-7.05	114.96	119.90
2	E	292	DG	C5-C6-O6	-7.04	124.38	128.60
2	E	285	DT	N3-C4-O4	7.03	124.11	119.90
3	F	400	DA	C4-C5-N7	-6.98	107.21	110.70
3	F	394	DG	C5-C6-N1	6.96	114.98	111.50
2	E	289	DT	C6-C5-C7	-6.73	118.86	122.90
2	E	286	DG	C5-C6-N1	6.58	114.79	111.50
3	F	401	DC	N3-C4-C5	6.52	124.51	121.90
3	F	396	DA	C5-C6-N6	6.46	128.87	123.70
3	F	397	DC	C2-N1-C1'	-6.46	111.70	118.80
2	E	281	DC	N1-C2-N3	-6.42	114.70	119.20
1	A	93	LEU	CB-CG-CD2	6.40	121.88	111.00
2	E	282	DC	N1-C2-O2	-6.34	115.09	118.90
3	F	397	DC	C6-N1-C1'	6.30	128.35	120.80
3	F	395	DG	C6-N1-C2	-6.29	121.32	125.10
1	A	198	ARG	NE-CZ-NH1	6.27	123.43	120.30
1	B	260	LYS	CD-CE-NZ	-6.26	97.31	111.70
3	F	396	DA	O4'-C1'-C2'	-6.25	100.90	105.90
2	E	291	DC	C6-N1-C2	6.24	122.80	120.30
2	E	288	DG	O4'-C1'-N9	-6.24	103.64	108.00
3	F	394	DG	N3-C4-C5	-6.09	125.55	128.60
2	E	281	DC	P-O3'-C3'	-6.06	112.43	119.70
3	F	398	DT	OP1-P-OP2	6.02	128.63	119.60
2	E	289	DT	C5-C6-N1	-5.95	120.13	123.70
1	B	255	ASN	N-CA-CB	-5.95	99.90	110.60
2	E	284	DG	N1-C2-N2	-5.95	110.85	116.20
3	F	394	DG	C2-N3-C4	5.94	114.87	111.90
3	F	401	DC	O4'-C1'-N1	5.91	112.13	108.00
1	B	249	MET	CG-SD-CE	-5.82	90.89	100.20
2	E	291	DC	P-O5'-C5'	-5.81	111.61	120.90
3	F	400	DA	C4-C5-C6	5.74	119.87	117.00
1	A	199	ASP	CB-CG-OD2	5.74	123.46	118.30
2	E	282	DC	P-O3'-C3'	-5.72	112.83	119.70
1	A	278	LEU	CA-CB-CG	5.71	128.44	115.30
3	F	394	DG	C4-C5-N7	-5.63	108.55	110.80
2	E	283	DC	O4'-C1'-N1	5.61	111.92	108.00
3	F	400	DA	N3-C4-C5	-5.60	122.88	126.80
1	A	126	SER	CB-CA-C	5.59	120.71	110.10

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	394	DG	O4'-C4'-C3'	5.58	109.35	106.00
1	A	60	ASP	CB-CG-OD2	5.56	123.30	118.30
2	E	289	DT	N1-C2-N3	5.56	117.94	114.60
3	F	400	DA	C6-N1-C2	-5.54	115.28	118.60
3	F	396	DA	O4'-C1'-N9	-5.53	104.13	108.00
3	F	398	DT	C4-C5-C7	5.52	122.31	119.00
2	E	288	DG	O3'-P-O5'	-5.51	93.53	104.00
1	A	95	MET	CG-SD-CE	5.50	109.00	100.20
3	F	400	DA	C2-N3-C4	5.45	113.33	110.60
2	E	286	DG	C6-N1-C2	-5.45	121.83	125.10
2	E	292	DG	C5-N7-C8	5.43	107.01	104.30
3	F	392	DC	C5-C6-N1	5.41	123.70	121.00
1	B	155	LEU	CA-CB-CG	5.40	127.72	115.30
2	E	284	DG	N3-C4-N9	5.40	129.24	126.00
2	E	284	DG	N9-C4-C5	-5.40	103.24	105.40
3	F	401	DC	C6-N1-C2	5.36	122.44	120.30
1	B	36	GLU	CB-CA-C	5.32	121.04	110.40
1	A	237[A]	MET	CG-SD-CE	5.31	108.69	100.20
1	A	237[B]	MET	CG-SD-CE	5.31	108.69	100.20
2	E	284	DG	N1-C6-O6	-5.29	116.73	119.90
2	E	282	DC	N3-C2-O2	5.21	125.55	121.90
2	E	287	DA	N7-C8-N9	5.21	116.41	113.80
1	A	267	ARG	CG-CD-NE	5.19	122.71	111.80
2	E	292	DG	C2-N3-C4	5.16	114.48	111.90
3	F	394	DG	C8-N9-C4	-5.15	104.34	106.40
2	E	292	DG	N7-C8-N9	-5.13	110.54	113.10
2	E	286	DG	N3-C4-C5	-5.11	126.05	128.60
1	B	40	ASP	CB-CG-OD2	5.09	122.89	118.30
2	E	288	DG	C5-C6-N1	5.09	114.05	111.50
2	E	286	DG	C2-N3-C4	5.09	114.44	111.90
1	A	260	LYS	N-CA-CB	5.07	119.72	110.60
3	F	400	DA	C8-N9-C4	-5.06	103.78	105.80
3	F	401	DC	O3'-P-O5'	-5.02	94.45	104.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	259	ASN	Peptide
1	A	59	GLN	Peptide
1	B	59	GLN	Peptide

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	1934	0	1940	39	0
1	B	1939	0	1942	49	0
2	E	241	0	136	5	0
3	F	248	0	140	6	0
4	A	8	0	3	1	0
5	A	6	0	8	2	0
6	A	4	0	8	0	0
6	B	4	0	8	0	0
7	A	15	0	18	0	0
7	B	15	0	18	0	0
8	F	16	0	12	6	0
9	A	123	0	0	8	0
9	B	82	0	0	13	0
9	E	12	0	0	1	0
9	F	10	0	0	2	0
All	All	4657	0	4233	102	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 12.

All (102) 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:60:ASP:HB3	1:A:61:PRO:CD	1.18	1.48
1:B:249:MET:CE	1:B:249:MET:SD	2.02	1.47
1:B:237:MET:CE	1:B:237:MET:SD	2.02	1.46
1:A:60:ASP:CB	1:A:61:PRO:CD	2.13	1.26
1:A:77:GLU:HG2	9:A:2019:HOH:O	1.33	1.20
1:A:60:ASP:CB	1:A:61:PRO:HD2	1.75	1.11
1:A:60:ASP:HB3	1:A:61:PRO:HD3	1.16	1.08
1:B:195:LYS:H	1:B:195:LYS:HD2	1.18	1.07
1:B:60:ASP:HB3	1:B:61:PRO:HD3	1.37	1.05
3:F:392:DC:H1'	8:F:1405:DUR:H5'1	1.43	1.01
1:A:60:ASP:HB3	1:A:61:PRO:HD2	1.03	1.00
8:F:1405:DUR:H5'2	8:F:1405:DUR:H6	1.41	0.98

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:195:LYS:HD2	1:B:195:LYS:N	1.91	0.86
1:B:60:ASP:HB3	1:B:61:PRO:CD	2.07	0.84
3:F:394:DG:OP1	8:F:1405:DUR:H2'2	1.82	0.80
5:A:1283:GOL:H12	9:A:2099:HOH:O	1.84	0.78
2:E:282:DC:H2"	2:E:283:DC:OP2	1.82	0.78
4:A:1282:URA:H5	9:A:2035:HOH:O	1.70	0.74
1:B:60:ASP:CB	1:B:61:PRO:HD3	2.19	0.71
1:B:241:ILE:HG22	1:B:243:VAL:HG23	1.73	0.69
1:B:60:ASP:CB	1:B:61:PRO:CD	2.70	0.69
1:B:249:MET:CG	1:B:249:MET:CE	2.71	0.68
1:A:146:GLU:OE2	5:A:1283:GOL:H11	1.96	0.66
1:A:173:HIS:HD2	9:A:2004:HOH:O	1.77	0.66
1:B:187:ASN:HB2	9:B:2029:HOH:O	1.94	0.66
1:A:60:ASP:CB	1:A:61:PRO:HD3	2.03	0.65
1:B:48:GLU:HG3	9:B:2007:HOH:O	1.95	0.65
8:F:1405:DUR:C5'	8:F:1405:DUR:H6	2.22	0.65
1:B:40:ASP:O	1:B:44:LYS:HG3	1.98	0.64
1:B:59:GLN:O	1:B:61:PRO:HD2	1.97	0.63
1:A:173:HIS:HE1	9:A:2085:HOH:O	1.81	0.63
1:B:100:PHE:HE1	9:B:2030:HOH:O	1.82	0.62
1:B:50:ASN:ND2	9:B:2009:HOH:O	2.33	0.61
1:A:60:ASP:HB2	1:A:61:PRO:HD2	1.82	0.58
1:A:59:GLN:O	1:A:60:ASP:HB2	2.05	0.57
1:A:67:ASN:C	1:A:67:ASN:HD22	2.09	0.56
1:B:190:PRO:O	1:B:193:LEU:HB2	2.06	0.55
1:A:67:ASN:ND2	1:A:69:LEU:H	2.04	0.55
1:A:78[B]:ASN:ND2	9:A:2018:HOH:O	2.39	0.55
1:B:52:LYS:HE2	9:B:2007:HOH:O	2.06	0.55
1:B:102:MET:HE1	9:B:2046:HOH:O	2.06	0.55
2:E:291:DC:H5'	2:E:292:DG:H5"	1.89	0.54
1:B:254:ARG:NE	1:B:254:ARG:O	2.36	0.54
1:A:87:LYS:H	1:A:87:LYS:NZ	2.04	0.54
1:B:262:TRP:CE2	1:B:266:VAL:HG21	2.42	0.54
1:B:195:LYS:HA	9:B:2056:HOH:O	2.08	0.53
1:B:195:LYS:CD	1:B:195:LYS:N	2.68	0.53
1:A:111:GLU:OE1	1:A:114:HIS:HD2	1.90	0.53
1:B:227:PHE:CD1	1:B:227:PHE:C	2.83	0.51
1:B:229:GLU:HG3	1:B:246:LYS:O	2.12	0.50
1:B:189:THR:HG23	9:B:2054:HOH:O	2.11	0.50
1:B:41:SER:O	1:B:45:VAL:HG23	2.10	0.50
1:B:77:GLU:HG3	1:B:77:GLU:O	2.11	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:255:ASN:HB2	2:E:281:DC:H1'	1.95	0.49
3:F:396:DA:H2"	9:F:2008:HOH:O	2.11	0.49
1:B:279:LEU:O	1:B:280:THR:C	2.50	0.49
1:B:246:LYS:HG3	1:B:273:LEU:HD22	1.96	0.48
1:B:102:MET:HB2	1:B:107:VAL:O	2.14	0.48
1:A:279:LEU:HD23	1:A:279:LEU:HA	1.72	0.48
1:B:213:VAL:HG13	1:B:218:VAL:CG2	2.44	0.47
1:B:104:GLN:O	1:B:139:GLY:HA2	2.14	0.47
1:A:270:LEU:HD22	1:A:275:VAL:HG21	1.97	0.47
2:E:290:DC:H2"	2:E:291:DC:H5"	1.96	0.47
1:A:88:LYS:HE3	1:A:171:PHE:HB2	1.98	0.46
1:A:60:ASP:HB2	1:A:197:GLN:HE22	1.79	0.46
1:B:37:SER:HB2	1:B:85:LYS:O	2.16	0.46
1:B:85:LYS:NZ	9:B:2014:HOH:O	2.45	0.45
1:B:52:LYS:CE	9:B:2007:HOH:O	2.64	0.45
1:A:48:GLU:OE2	1:A:52:LYS:HE3	2.17	0.45
1:A:87:LYS:HB2	1:A:87:LYS:HE2	1.38	0.45
1:B:116:ARG:HD3	9:B:2034:HOH:O	2.18	0.44
1:B:196:ALA:N	9:B:2056:HOH:O	2.51	0.44
1:B:236:LEU:HD12	1:B:245:VAL:CG2	2.47	0.44
1:A:207:GLU:O	1:A:211:GLN:HG3	2.17	0.44
1:B:195:LYS:O	1:B:196:ALA:HB2	2.17	0.44
1:A:91:LEU:HB2	1:A:218:VAL:HG21	1.99	0.44
1:A:173:HIS:CD2	9:A:2004:HOH:O	2.61	0.44
1:A:64:TYR:CE1	1:A:132:HIS:HD2	2.36	0.44
1:A:66:TYR:HB2	1:A:179:ILE:HG23	2.00	0.44
1:B:182:ASN:OD1	1:B:182:ASN:C	2.57	0.43
1:B:175:HIS:O	1:B:205:CYS:HB3	2.19	0.43
1:A:95:MET:HG3	1:A:227:PHE:CE2	2.54	0.43
1:A:262:TRP:CE2	1:A:266:VAL:HG21	2.54	0.43
1:A:200:THR:HA	9:A:2081:HOH:O	2.18	0.43
1:A:87:LYS:H	1:A:87:LYS:HZ2	1.66	0.42
1:B:213:VAL:HG13	1:B:218:VAL:HG22	2.01	0.42
2:E:282:DC:P	9:E:2003:HOH:O	2.77	0.42
1:B:50:ASN:ND2	1:B:67:ASN:HD21	2.17	0.42
1:B:169:HIS:ND1	9:B:2044:HOH:O	2.37	0.42
1:A:211:GLN:O	1:A:215:VAL:HG13	2.20	0.42
1:B:89:GLU:HG3	1:B:169:HIS:HB3	2.00	0.42
1:A:97:PRO:HD3	1:A:176:CYS:O	2.20	0.42
1:A:250:HIS:CG	1:A:251:PRO:HD2	2.55	0.42
1:B:279:LEU:O	1:B:280:THR:O	2.38	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:38:PRO:O	1:B:39:ALA:C	2.57	0.41
1:A:271:LEU:HG	1:A:276:LEU:HD12	2.02	0.41
1:A:260:LYS:HG2	1:A:261:GLY:N	2.34	0.41
1:A:67:ASN:HD22	1:A:69:LEU:H	1.68	0.41
3:F:393:3DR:OP2	9:F:2002:HOH:O	2.22	0.41
3:F:393:3DR:OP1	8:F:1405:DUR:H4'	2.22	0.40
3:F:394:DG:OP1	8:F:1405:DUR:C2'	2.63	0.40
1:B:66:TYR:HB2	1:B:179:ILE:HG23	2.04	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	244/247 (99%)	235 (96%)	7 (3%)	2 (1%)	24	27
1	B	243/247 (98%)	227 (93%)	12 (5%)	4 (2%)	12	11
All	All	487/494 (99%)	462 (95%)	19 (4%)	6 (1%)	16	16

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	60	ASP
1	B	60	ASP
1	B	196	ALA
1	A	277	SER
1	B	61	PRO
1	B	261	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	214/215 (100%)	197 (92%)	17 (8%)	15 19
1	B	214/215 (100%)	193 (90%)	21 (10%)	10 11
All	All	428/430 (100%)	390 (91%)	38 (9%)	12 14

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	VAL
1	A	59	GLN
1	A	60	ASP
1	A	67	ASN
1	A	85	LYS
1	A	87	LYS
1	A	93	LEU
1	A	126	SER
1	A	129	GLU
1	A	134	LYS
1	A	181	MET
1	A	218	VAL
1	A	219	LYS
1	A	234	LYS
1	A	257	GLN
1	A	271	LEU
1	A	277	SER
1	B	36	GLU
1	B	37	SER
1	B	41	SER
1	B	48	GLU
1	B	51	LEU
1	B	57	VAL
1	B	59	GLN
1	B	60	ASP
1	B	86	SER
1	B	127	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	136	ARG
1	B	144	GLN
1	B	155	LEU
1	B	179	ILE
1	B	195	LYS
1	B	197	GLN
1	B	200	THR
1	B	234	LYS
1	B	237	MET
1	B	254	ARG
1	B	272	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	50	ASN
1	A	67	ASN
1	A	81	GLN
1	A	114	HIS
1	A	173	HIS
1	A	183	HIS
1	A	230	GLN
1	A	257	GLN
1	B	50	ASN
1	B	78	ASN
1	B	120	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)

2 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	$\#\|Z\| > 2$
3	3DR	F	393	3	7,11,12	0.79	0	8,14,17	1.23	1 (12%)
3	3DR	F	399	3	7,11,12	1.69	2 (28%)	8,14,17	1.78	2 (25%)

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	3DR	F	393	3	-	0/3/15/16	0/1/1/1
3	3DR	F	399	3	-	0/3/15/16	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	F	399	3DR	O4'-C4'	2.01	1.48	1.44
3	F	399	3DR	O3'-C3'	3.80	1.51	1.43

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	F	399	3DR	C2'-C3'-C4'	-3.09	96.40	102.77
3	F	393	3DR	O4'-C4'-C5'	-2.37	104.34	109.53
3	F	399	3DR	O3'-C3'-C4'	2.60	120.56	110.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	F	393	3DR	2	0

5.5 Carbohydrates (i)

There are no carbohydrates 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	$\#\|Z\| > 2$
4	URA	A	1282	-	4,8,8	5.59	3 (75%)	6,10,10	11.74	5 (83%)
5	GOL	A	1283	-	5,5,5	0.49	0	5,5,5	1.04	0
6	IPA	A	1284	-	3,3,3	0.80	0	3,3,3	0.35	0
7	EPE	A	1285	-	14,15,15	0.96	0	18,20,20	2.79	11 (61%)
6	IPA	B	1281	-	3,3,3	1.02	0	3,3,3	0.55	0
7	EPE	B	1282	-	14,15,15	0.89	0	18,20,20	3.26	11 (61%)
8	DUR	F	1405	-	12,17,17	1.29	2 (16%)	17,24,24	3.60	3 (17%)

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	URA	A	1282	-	-	0/0/0/0	0/1/1/1
5	GOL	A	1283	-	-	0/4/4/4	0/0/0/0
6	IPA	A	1284	-	-	0/0/0/0	0/0/0/0
7	EPE	A	1285	-	-	0/9/19/19	0/1/1/1
6	IPA	B	1281	-	-	0/0/0/0	0/0/0/0
7	EPE	B	1282	-	-	0/9/19/19	0/1/1/1
8	DUR	F	1405	-	-	0/2/18/18	0/2/2/2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	F	1405	DUR	C6-C5	-2.24	1.33	1.38
8	F	1405	DUR	C4-N3	2.76	1.38	1.33
4	A	1282	URA	O4-C4	3.80	1.33	1.24
4	A	1282	URA	C6-N1	4.89	1.45	1.34
4	A	1282	URA	C4-N3	9.30	1.50	1.33

All (30) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1282	URA	N1-C2-N3	-21.16	114.83	128.33
4	A	1282	URA	C5-C4-N3	-4.18	112.40	123.12
7	B	1282	EPE	O3S-S-O2S	-3.90	102.53	111.61
8	F	1405	DUR	C5-C4-N3	-3.14	115.07	123.12
7	A	1285	EPE	O2S-S-O1S	-2.74	103.49	113.48
7	A	1285	EPE	O3S-S-O1S	-2.43	105.96	111.61
7	A	1285	EPE	C6-N1-C2	2.04	113.31	108.90
7	B	1282	EPE	C5-C6-N1	2.06	114.31	110.63
7	B	1282	EPE	O3S-S-O1S	2.10	116.50	111.61
8	F	1405	DUR	O4'-C1'-N1	2.15	111.45	107.72
7	B	1282	EPE	C7-N4-C3	2.26	117.06	111.27
7	A	1285	EPE	C9-N1-C6	2.40	117.43	111.27
7	A	1285	EPE	O2S-S-C10	2.42	108.97	106.91
7	B	1282	EPE	C2-C3-N4	2.60	115.29	110.63
7	B	1282	EPE	C7-N4-C5	2.78	118.40	111.27
7	A	1285	EPE	C7-N4-C5	2.80	118.46	111.27
7	B	1282	EPE	C3-C2-N1	2.82	115.68	110.63
7	A	1285	EPE	C6-C5-N4	2.97	115.94	110.63
7	B	1282	EPE	C6-C5-N4	2.99	115.98	110.63
4	A	1282	URA	C6-C5-C4	3.20	123.27	117.28
7	A	1285	EPE	C7-N4-C3	3.21	119.50	111.27
7	B	1282	EPE	C6-N1-C2	3.72	116.95	108.90
7	A	1285	EPE	C2-C3-N4	4.45	118.60	110.63
7	A	1285	EPE	O1S-S-C10	4.76	110.97	106.91
7	B	1282	EPE	C5-N4-C3	5.36	120.51	108.90
7	A	1285	EPE	C5-N4-C3	6.07	122.04	108.90
7	B	1282	EPE	O2S-S-C10	9.07	114.64	106.91
4	A	1282	URA	C4-N3-C2	11.84	125.87	114.14
8	F	1405	DUR	C4-N3-C2	14.03	128.04	114.14
4	A	1282	URA	C6-N1-C2	14.39	121.48	114.40

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	1282	URA	1	0
5	A	1283	GOL	2	0
8	F	1405	DUR	6	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

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	244/247 (98%)	-0.39	0 100 100	16, 29, 47, 58	0
1	B	245/247 (99%)	-0.19	4 (1%) 74 80	19, 38, 62, 74	2 (0%)
2	E	12/12 (100%)	0.07	0 100 100	30, 55, 70, 71	0
3	F	11/13 (84%)	0.58	3 (27%) 1 1	36, 52, 83, 84	0
All	All	512/519 (98%)	-0.26	7 (1%) 78 83	16, 34, 61, 84	2 (0%)

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	195	LYS	3.2
1	B	60	ASP	3.2
3	F	403	DG	2.8
3	F	404	DG	2.5
3	F	402	DG	2.4
1	B	57	VAL	2.4
1	B	197	GLN	2.2

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

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
3	3DR	F	393	11/12	0.86	0.21	-	66,83,93,93	0
3	3DR	F	399	11/12	0.97	0.10	-	55,63,66,67	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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å^2)	Q<0.9
8	DUR	F	1405	16/16	0.68	0.83	16.13	73,76,81,81	16
6	IPA	B	1281	4/4	0.97	0.21	7.02	18,19,24,26	0
5	GOL	A	1283	6/6	0.69	0.34	6.17	77,83,84,84	0
6	IPA	A	1284	4/4	0.98	0.29	6.02	20,21,28,30	0
7	EPE	B	1282	15/15	0.94	0.27	4.19	46,86,97,97	0
7	EPE	A	1285	15/15	0.96	0.21	3.56	37,69,74,74	0
4	URA	A	1282	8/8	0.92	0.16	1.73	23,36,42,45	0

6.5 Other polymers (i)

There are no such residues in this entry.