



wwPDB X-ray Structure Validation Summary Report ⓘ

Oct 9, 2023 – 11:34 AM EDT

PDB ID : 8CTY
Title : 12-mer DNA structure of ExBIM bound to RNase-H
Authors : Pallan, P.S.; Egli, M.
Deposited on : 2022-05-16
Resolution : 2.30 Å(reported)

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

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

A user guide is available at

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

with specific help available everywhere you see the ⓘ 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 : 2.35.1
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) : 2.35.1

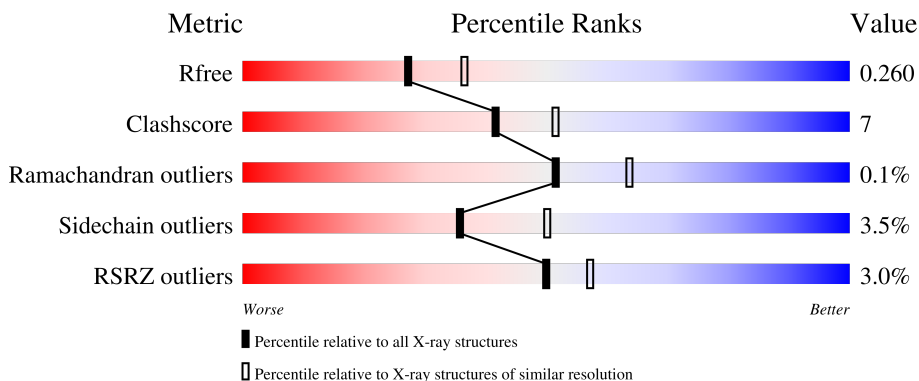
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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}	130704	5042 (2.30-2.30)
Clashscore	141614	5643 (2.30-2.30)
Ramachandran outliers	138981	5575 (2.30-2.30)
Sidechain outliers	138945	5575 (2.30-2.30)
RSRZ outliers	127900	4938 (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 of the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions $\leq 5\%$. 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	142	2% 81% 12% • 6%
1	B	142	% 82% 11% • 6%
1	C	142	71% 19% • 8%
1	D	142	77% 13% • 8%
1	E	142	75% 15% • 9%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	142	
1	G	142	
1	H	142	
2	I	12	
2	J	12	
2	K	12	
2	L	12	
2	M	12	
2	N	12	

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
3	EDO	J	101	-	-	X	-
5	GOL	E	205	-	-	X	-

2 Entry composition [i](#)

There are 9 unique types of molecules in this entry. The entry contains 10263 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 Ribonuclease H.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	133	1080	691	181	207	1	6	0	0
1	B	133	1080	691	181	207	1	13	0	0
1	C	130	1056	676	178	201	1	0	0	0
1	D	130	1056	676	178	201	1	2	0	0
1	E	129	1051	673	177	200	1	7	0	0
1	F	132	1076	689	180	206	1	13	0	0
1	G	132	1076	689	180	206	1	39	0	0
1	H	128	1042	667	175	199	1	28	0	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	55	GLY	-	expression tag	UNP Q9KEI9
A	56	SER	-	expression tag	UNP Q9KEI9
A	57	HIS	-	expression tag	UNP Q9KEI9
A	58	MET	-	expression tag	UNP Q9KEI9
A	132	ASN	ASP	engineered mutation	UNP Q9KEI9
B	55	GLY	-	expression tag	UNP Q9KEI9
B	56	SER	-	expression tag	UNP Q9KEI9
B	57	HIS	-	expression tag	UNP Q9KEI9
B	58	MET	-	expression tag	UNP Q9KEI9
B	132	ASN	ASP	engineered mutation	UNP Q9KEI9
C	55	GLY	-	expression tag	UNP Q9KEI9
C	56	SER	-	expression tag	UNP Q9KEI9
C	57	HIS	-	expression tag	UNP Q9KEI9

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	58	MET	-	expression tag	UNP Q9KEI9
C	132	ASN	ASP	engineered mutation	UNP Q9KEI9
D	55	GLY	-	expression tag	UNP Q9KEI9
D	56	SER	-	expression tag	UNP Q9KEI9
D	57	HIS	-	expression tag	UNP Q9KEI9
D	58	MET	-	expression tag	UNP Q9KEI9
D	132	ASN	ASP	engineered mutation	UNP Q9KEI9
E	55	GLY	-	expression tag	UNP Q9KEI9
E	56	SER	-	expression tag	UNP Q9KEI9
E	57	HIS	-	expression tag	UNP Q9KEI9
E	58	MET	-	expression tag	UNP Q9KEI9
E	132	ASN	ASP	engineered mutation	UNP Q9KEI9
F	55	GLY	-	expression tag	UNP Q9KEI9
F	56	SER	-	expression tag	UNP Q9KEI9
F	57	HIS	-	expression tag	UNP Q9KEI9
F	58	MET	-	expression tag	UNP Q9KEI9
F	132	ASN	ASP	engineered mutation	UNP Q9KEI9
G	55	GLY	-	expression tag	UNP Q9KEI9
G	56	SER	-	expression tag	UNP Q9KEI9
G	57	HIS	-	expression tag	UNP Q9KEI9
G	58	MET	-	expression tag	UNP Q9KEI9
G	132	ASN	ASP	engineered mutation	UNP Q9KEI9
H	55	GLY	-	expression tag	UNP Q9KEI9
H	56	SER	-	expression tag	UNP Q9KEI9
H	57	HIS	-	expression tag	UNP Q9KEI9
H	58	MET	-	expression tag	UNP Q9KEI9
H	132	ASN	ASP	engineered mutation	UNP Q9KEI9

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

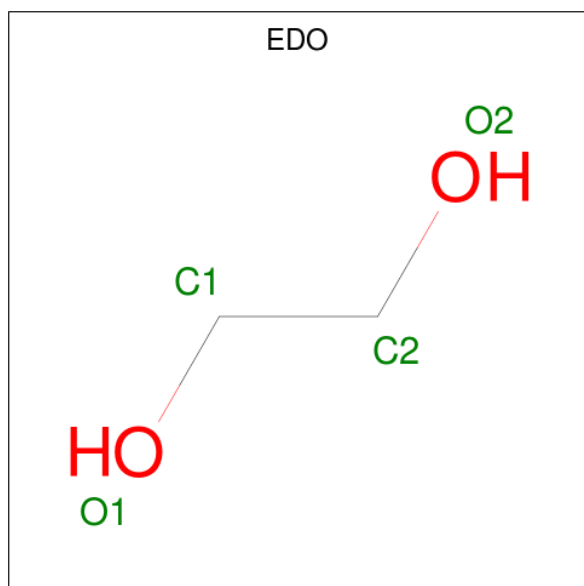
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	I	12	248	123	45	69	11	0	0	0
2	J	12	248	123	45	69	11	0	0	0
2	K	12	248	123	45	69	11	0	0	0
2	L	12	248	123	45	69	11	0	0	0
2	M	9	186	93	33	52	8	9	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	N	10	204	97	41	57	9	29	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C₂H₆O₂).

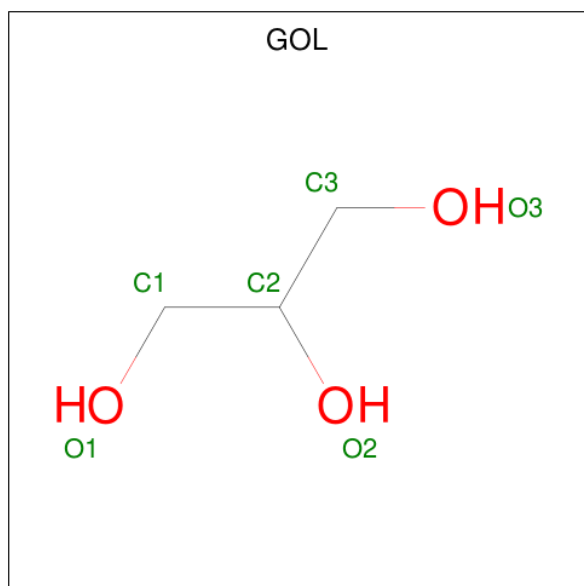


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	C	O		
3	A	1	4	2	2	0	0
3	A	1	4	2	2	0	0
3	B	1	4	2	2	0	0
3	B	1	4	2	2	0	0
3	C	1	4	2	2	0	0
3	C	1	4	2	2	0	0
3	E	1	4	2	2	0	0
3	E	1	4	2	2	0	0
3	J	1	4	2	2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

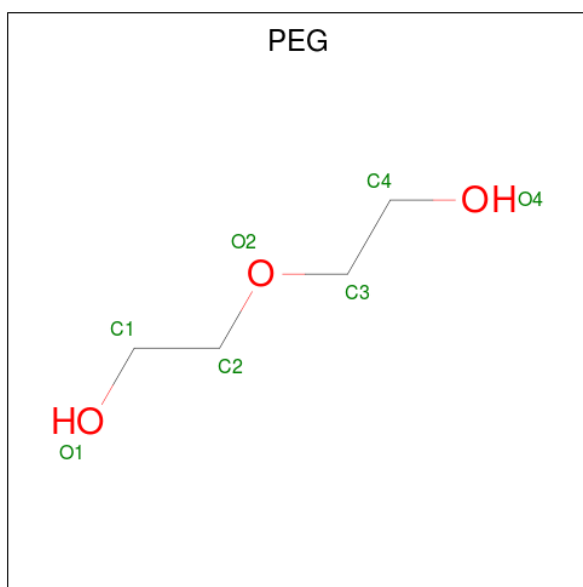
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	1	Total	Cl	0	0
			1	1		

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	C	O	0	0
			6	3	3		
5	B	1	Total	C	O	0	0
			6	3	3		
5	C	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	E	1	Total	C	O	0	0
			6	3	3		
5	F	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: $C_4H_{10}O_3$).

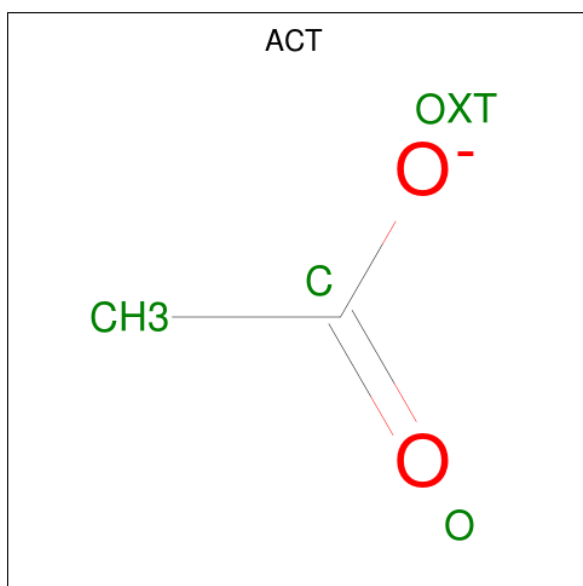


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	C	1	Total C O 7 4 3	0	0

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	F	1	Total Na 1 1	0	0

- Molecule 8 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	F	1	Total C O 4 2 2	0	0
8	G	1	Total C O 4 2 2	0	0
8	M	1	Total C O 4 2 2	0	0

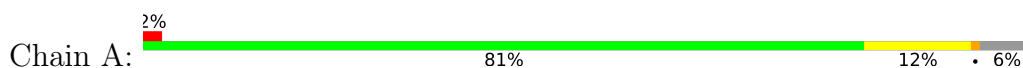
- Molecule 9 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	45	Total O 45 45	0	0
9	B	31	Total O 31 31	0	0
9	C	30	Total O 30 30	0	0
9	D	39	Total O 39 39	0	0
9	E	35	Total O 35 35	0	0
9	F	31	Total O 31 31	0	0
9	G	11	Total O 11 11	0	0
9	H	15	Total O 15 15	0	0
9	I	3	Total O 3 3	0	0
9	J	6	Total O 6 6	0	0
9	K	7	Total O 7 7	0	0
9	L	8	Total O 8 8	0	0
9	M	1	Total O 1 1	0	0
9	N	3	Total O 3 3	0	0

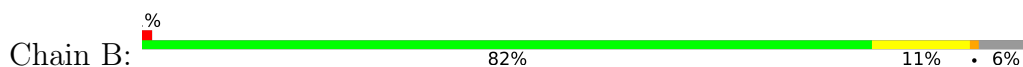
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

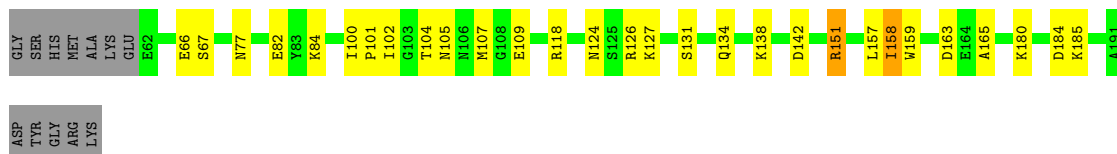
- Molecule 1: Ribonuclease H



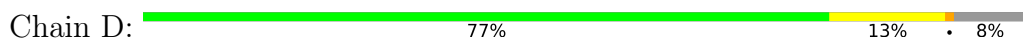
- Molecule 1: Ribonuclease H



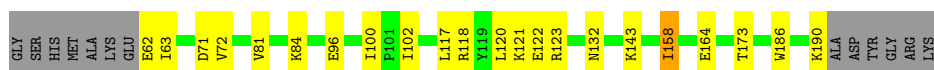
- Molecule 1: Ribonuclease H



- Molecule 1: Ribonuclease H



- Molecule 1: Ribonuclease H




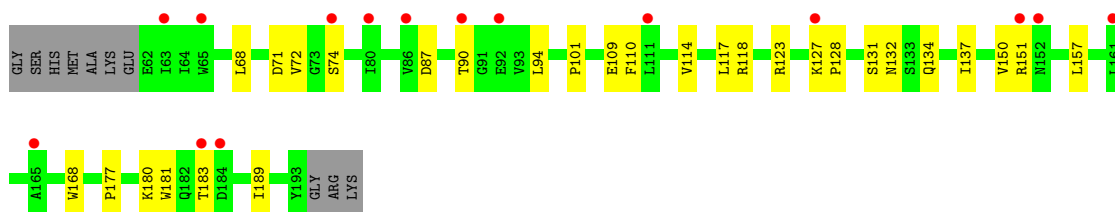
- Molecule 1: Ribonuclease H

Chain F:  74% 18% 7%




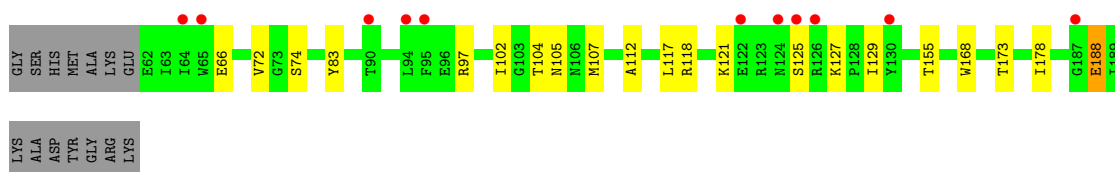
• Molecule 1: Ribonuclease H

Chain G:  11% 73% 20% 7%



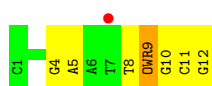
• Molecule 1: Ribonuclease H

Chain H:  8% 75% 14% 10%



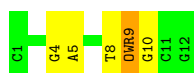
• Molecule 2: DNA (5'-D(*CP*GP*CP*GP*AP*AP*TP*TP*(OWR)P*GP*CP*G)-3')

Chain I:  8% 42% 50% 8%




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

Chain J:  58% 33% 8%




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

Chain K:  83% 17%



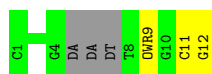
• Molecule 2: DNA (5'-D(*CP*GP*CP*GP*AP*AP*TP*TP*(OWR)P*GP*CP*G)-3')

Chain L:  83% 17%



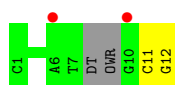
- Molecule 2: DNA (5'-D(*CP*GP*CP*GP*AP*AP*TP*TP*(OWR)P*GP*CP*G)-3')

Chain M:  50% 25% 25%



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

Chain N:  17% 67% 17% 17%



4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	64.98Å 65.83Å 95.72Å 84.54° 88.20° 62.49°	Depositor
Resolution (Å)	32.73 – 2.30 38.61 – 2.30	Depositor EDS
% Data completeness (in resolution range)	97.1 (32.73-2.30) 97.9 (38.61-2.30)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.11 (at 2.29Å)	Xtrriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.196 , 0.261 0.201 , 0.260	Depositor DCC
R_{free} test set	3146 reflections (5.14%)	wwPDB-VP
Wilson B-factor (Å ²)	38.6	Xtrriage
Anisotropy	0.086	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.36 , 47.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.34$	Xtrriage
Estimated twinning fraction	0.018 for h,h-k,-l	Xtrriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10263	wwPDB-VP
Average B, all atoms (Å ²)	42.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.99% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, NA, GOL, CL, OWR, PEG, ACT

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	0.44	0/1105	0.59	0/1498
1	B	0.42	0/1105	0.58	0/1498
1	C	0.49	0/1080	0.61	0/1464
1	D	0.46	0/1080	0.62	0/1464
1	E	0.45	0/1075	0.64	0/1457
1	F	0.43	0/1101	0.61	0/1493
1	G	0.36	0/1101	0.54	0/1493
1	H	0.37	0/1066	0.58	0/1446
2	I	0.84	0/250	0.91	0/382
2	J	0.93	0/250	1.02	1/382 (0.3%)
2	K	1.09	0/250	1.17	1/382 (0.3%)
2	L	1.03	0/250	1.15	0/382
2	M	0.88	0/179	0.94	0/270
2	N	0.83	0/228	0.91	0/348
All	All	0.53	0/10120	0.68	2/13959 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	10	DG	O4'-C1'-N9	-5.61	104.08	108.00
2	K	1	DC	O4'-C4'-C3'	-5.58	102.27	104.50

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	1080	0	1071	8	1
1	B	1080	0	1071	10	0
1	C	1056	0	1055	21	0
1	D	1056	0	1055	13	0
1	E	1051	0	1050	17	0
1	F	1076	0	1068	18	1
1	G	1076	0	1068	17	0
1	H	1042	0	1037	11	0
2	I	248	0	125	7	0
2	J	248	0	125	7	0
2	K	248	0	125	0	0
2	L	248	0	125	1	0
2	M	186	0	91	1	0
2	N	204	0	114	1	0
3	A	8	0	12	1	0
3	B	8	0	12	1	0
3	C	8	0	12	1	0
3	E	8	0	12	1	0
3	J	4	0	6	8	0
4	B	1	0	0	0	0
5	B	12	0	16	3	0
5	C	6	0	8	2	0
5	E	18	0	24	5	0
5	F	6	0	8	2	0
6	C	7	0	10	1	0
7	F	1	0	0	0	0
8	F	4	0	3	0	0
8	G	4	0	3	0	0
8	M	4	0	3	0	0
9	A	45	0	0	1	0
9	B	31	0	0	1	0
9	C	30	0	0	2	0
9	D	39	0	0	0	0
9	E	35	0	0	1	0
9	F	31	0	0	2	0
9	G	11	0	0	2	0
9	H	15	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
9	I	3	0	0	0	0
9	J	6	0	0	0	0
9	K	7	0	0	0	0
9	L	8	0	0	0	0
9	M	1	0	0	0	0
9	N	3	0	0	0	0
All	All	10263	0	9309	130	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 7.

The worst 5 of 130 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)
2:I:4:DG:H1	3:J:101:EDO:H22	1.30	0.97
2:J:8:DT:H2'	3:J:101:EDO:H12	1.52	0.89
1:C:134:GLN:HE21	1:C:180:LYS:HG3	1.46	0.80
1:G:68:LEU:HD21	1:G:94:LEU:HD12	1.68	0.75
5:E:204:GOL:O3	9:E:301:HOH:O	2.05	0.75

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:A:185:LYS:O	1:F:97:ARG:NH2[1_544]	2.19	0.01

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	131/142 (92%)	127 (97%)	4 (3%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	131/142 (92%)	128 (98%)	3 (2%)	0	100	100
1	C	128/142 (90%)	128 (100%)	0	0	100	100
1	D	128/142 (90%)	127 (99%)	1 (1%)	0	100	100
1	E	127/142 (89%)	123 (97%)	4 (3%)	0	100	100
1	F	130/142 (92%)	130 (100%)	0	0	100	100
1	G	130/142 (92%)	125 (96%)	5 (4%)	0	100	100
1	H	126/142 (89%)	119 (94%)	6 (5%)	1 (1%)	19	23
All	All	1031/1136 (91%)	1007 (98%)	23 (2%)	1 (0%)	51	64

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	H	188	GLU

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	117/124 (94%)	111 (95%)	6 (5%)	24	33
1	B	117/124 (94%)	115 (98%)	2 (2%)	60	76
1	C	115/124 (93%)	110 (96%)	5 (4%)	29	40
1	D	115/124 (93%)	112 (97%)	3 (3%)	46	63
1	E	115/124 (93%)	111 (96%)	4 (4%)	36	50
1	F	117/124 (94%)	112 (96%)	5 (4%)	29	40
1	G	117/124 (94%)	114 (97%)	3 (3%)	46	63
1	H	114/124 (92%)	110 (96%)	4 (4%)	36	50
All	All	927/992 (93%)	895 (96%)	32 (4%)	36	50

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	72	VAL
1	H	97	ARG
1	C	158	ILE
1	C	151	ARG
1	H	125	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	134	GLN
1	F	105	ASN
1	G	134	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](#)

5 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	OWR	K	9	2	22,27,28	1.42	3 (13%)	28,39,42	0.97	1 (3%)
2	OWR	L	9	2	22,27,28	1.42	3 (13%)	28,39,42	0.79	1 (3%)
2	OWR	J	9	2	22,27,28	1.26	3 (13%)	28,39,42	0.68	0
2	OWR	M	9	2	22,27,28	1.31	3 (13%)	28,39,42	0.68	0
2	OWR	I	9	2	22,27,28	1.30	3 (13%)	28,39,42	0.74	1 (3%)

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	OWR	K	9	2	-	0/3/21/22	0/4/4/4
2	OWR	L	9	2	-	0/3/21/22	0/4/4/4
2	OWR	J	9	2	-	0/3/21/22	0/4/4/4
2	OWR	M	9	2	-	0/3/21/22	0/4/4/4
2	OWR	I	9	2	-	0/3/21/22	0/4/4/4

The worst 5 of 15 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	K	9	OWR	C6-C5	-3.65	1.33	1.39
2	L	9	OWR	C6-C5	-3.44	1.34	1.39
2	I	9	OWR	C6-C5	-3.18	1.34	1.39
2	M	9	OWR	C6-C5	-3.18	1.34	1.39
2	L	9	OWR	C4-C5	-3.07	1.33	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	K	9	OWR	O3'-C3'-C4'	-2.61	100.11	110.10
2	L	9	OWR	O3'-C3'-C4'	-2.46	100.68	110.10
2	I	9	OWR	O4'-C4'-C5'	2.01	115.97	109.37

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	J	9	OWR	3	0
2	I	9	OWR	1	0

5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 22 ligands modelled in this entry, 2 are monoatomic - leaving 20 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
6	PEG	C	203	-	6,6,6	0.25	0	5,5,5	0.08	0
3	EDO	C	201	-	3,3,3	0.54	0	2,2,2	0.20	0
5	GOL	B	204	-	5,5,5	1.18	0	5,5,5	1.04	0
3	EDO	A	201	-	3,3,3	0.34	0	2,2,2	0.75	0
3	EDO	B	202	-	3,3,3	0.60	0	2,2,2	0.25	0
5	GOL	C	204	-	5,5,5	1.48	1 (20%)	5,5,5	0.69	0
5	GOL	B	205	-	5,5,5	0.91	0	5,5,5	1.03	0
3	EDO	B	203	-	3,3,3	0.50	0	2,2,2	0.22	0
3	EDO	J	101	-	3,3,3	0.51	0	2,2,2	0.15	0
5	GOL	E	205	-	5,5,5	1.47	1 (20%)	5,5,5	1.03	0
3	EDO	C	202	-	3,3,3	0.55	0	2,2,2	0.17	0
8	ACT	G	201	-	3,3,3	1.49	1 (33%)	3,3,3	1.16	0
3	EDO	E	201	-	3,3,3	0.53	0	2,2,2	0.29	0
3	EDO	A	202	-	3,3,3	0.50	0	2,2,2	0.30	0
5	GOL	F	203	-	5,5,5	1.42	1 (20%)	5,5,5	1.14	0
3	EDO	E	202	-	3,3,3	0.49	0	2,2,2	0.28	0
5	GOL	E	204	-	5,5,5	1.07	0	5,5,5	0.89	0
8	ACT	M	101	-	3,3,3	1.50	1 (33%)	3,3,3	1.46	0
5	GOL	E	203	-	5,5,5	1.14	0	5,5,5	1.06	0
8	ACT	F	202	-	3,3,3	1.45	1 (33%)	3,3,3	1.60	1 (33%)

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
5	GOL	E	203	-	-	0/4/4/4	-
6	PEG	C	203	-	-	1/4/4/4	-
5	GOL	C	204	-	-	4/4/4/4	-
3	EDO	C	202	-	-	1/1/1/1	-
3	EDO	E	202	-	-	1/1/1/1	-
5	GOL	E	204	-	-	0/4/4/4	-
3	EDO	C	201	-	-	1/1/1/1	-
5	GOL	B	205	-	-	0/4/4/4	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	E	201	-	-	1/1/1/1	-
3	EDO	A	202	-	-	1/1/1/1	-
5	GOL	B	204	-	-	1/4/4/4	-
3	EDO	B	203	-	-	0/1/1/1	-
3	EDO	A	201	-	-	1/1/1/1	-
3	EDO	B	202	-	-	0/1/1/1	-
5	GOL	F	203	-	-	1/4/4/4	-
3	EDO	J	101	-	-	1/1/1/1	-
5	GOL	E	205	-	-	4/4/4/4	-

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	F	203	GOL	C3-C2	2.78	1.63	1.51
5	E	205	GOL	C1-C2	2.61	1.62	1.51
8	G	201	ACT	CH3-C	2.36	1.59	1.49
8	F	202	ACT	CH3-C	2.29	1.58	1.49
8	M	101	ACT	CH3-C	2.26	1.58	1.49

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
8	F	202	ACT	O-C-CH3	-2.14	114.02	122.33

There are no chirality outliers.

5 of 18 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	E	205	GOL	O1-C1-C2-C3
5	E	205	GOL	C1-C2-C3-O3
6	C	203	PEG	C4-C3-O2-C2
5	C	204	GOL	O1-C1-C2-C3
5	C	204	GOL	C1-C2-C3-O3

There are no ring outliers.

11 monomers are involved in 25 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	C	203	PEG	1	0
3	C	201	EDO	1	0
5	B	204	GOL	3	0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	201	EDO	1	0
3	B	202	EDO	1	0
5	C	204	GOL	2	0
3	J	101	EDO	8	0
5	E	205	GOL	4	0
5	F	203	GOL	2	0
3	E	202	EDO	1	0
5	E	204	GOL	1	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, 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	133/142 (93%)	0.03	3 (2%) 60 67	27, 39, 52, 70	2 (1%)
1	B	133/142 (93%)	-0.06	1 (0%) 86 89	29, 37, 50, 60	3 (2%)
1	C	130/142 (91%)	-0.04	0 100 100	24, 34, 48, 57	0
1	D	130/142 (91%)	-0.15	0 100 100	24, 33, 46, 52	1 (0%)
1	E	129/142 (90%)	-0.02	0 100 100	23, 34, 50, 59	2 (1%)
1	F	132/142 (92%)	-0.03	0 100 100	27, 37, 50, 58	4 (3%)
1	G	132/142 (92%)	0.70	15 (11%) 5 7	41, 58, 74, 78	9 (6%)
1	H	128/142 (90%)	0.59	11 (8%) 10 14	37, 49, 75, 81	5 (3%)
2	I	11/12 (91%)	0.39	1 (9%) 9 12	34, 50, 81, 81	0
2	J	11/12 (91%)	-0.32	0 100 100	35, 40, 64, 68	0
2	K	11/12 (91%)	-0.22	0 100 100	29, 34, 37, 39	0
2	L	11/12 (91%)	-0.25	0 100 100	30, 34, 46, 47	0
2	M	8/12 (66%)	0.31	0 100 100	46, 53, 70, 98	1 (12%)
2	N	9/12 (75%)	0.87	2 (22%) 0 1	38, 63, 85, 102	2 (22%)
All	All	1108/1208 (91%)	0.13	33 (2%) 50 57	23, 39, 68, 102	29 (2%)

The worst 5 of 33 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	H	65	TRP	4.0
2	N	6	DA	3.9
1	H	124	ASN	3.7
1	G	161	LEU	3.4
1	H	187	GLY	3.3

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	OWR	I	9	24/25	0.81	0.27	64,80,96,97	0
2	OWR	M	9	24/25	0.87	0.29	59,75,93,98	0
2	OWR	J	9	24/25	0.94	0.14	40,46,53,63	0
2	OWR	L	9	24/25	0.96	0.15	30,34,44,46	0
2	OWR	K	9	24/25	0.97	0.13	28,32,36,39	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, 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
5	GOL	E	204	6/6	0.77	0.18	43,47,57,64	0
6	PEG	C	203	7/7	0.77	0.20	43,54,60,62	0
3	EDO	J	101	4/4	0.78	0.22	46,50,55,57	0
3	EDO	C	201	4/4	0.79	0.18	54,57,57,58	0
5	GOL	B	204	6/6	0.79	0.20	28,33,37,39	6
3	EDO	C	202	4/4	0.81	0.16	46,47,47,49	0
8	ACT	G	201	4/4	0.82	0.17	41,48,51,53	0
8	ACT	M	101	4/4	0.82	0.16	57,69,71,72	0
3	EDO	A	201	4/4	0.84	0.24	33,35,42,49	0
3	EDO	A	202	4/4	0.86	0.16	55,59,60,60	0
4	CL	B	201	1/1	0.86	0.17	63,63,63,63	0
5	GOL	E	205	6/6	0.86	0.16	32,41,43,48	0
5	GOL	F	203	6/6	0.87	0.12	36,45,47,48	0
5	GOL	C	204	6/6	0.87	0.19	26,39,42,42	0
3	EDO	E	202	4/4	0.89	0.18	43,45,47,51	0
8	ACT	F	202	4/4	0.89	0.19	39,39,42,42	0
3	EDO	E	201	4/4	0.90	0.15	39,44,48,50	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	EDO	B	202	4/4	0.91	0.12	35,39,41,45	0
5	GOL	E	203	6/6	0.92	0.15	32,39,44,45	0
3	EDO	B	203	4/4	0.94	0.15	46,49,50,52	0
7	NA	F	201	1/1	0.95	0.05	42,42,42,42	0
5	GOL	B	205	6/6	0.95	0.12	35,44,46,48	0

6.5 Other polymers [i](#)

There are no such residues in this entry.