



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jul 13, 2021 – 10:02 AM EDT

PDB ID : 1JRI  
Title : The Crystal Structure of an Sm-like Archaeal Protein with Two Heptamers in the Asymmetric Unit.  
Authors : Mura, C.; Eisenberg, D.  
Deposited on : 2001-08-13  
Resolution : 2.80 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](mailto:validation@mail.wwpdb.org)

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

---

The 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	:	2.22
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)	:	2.22

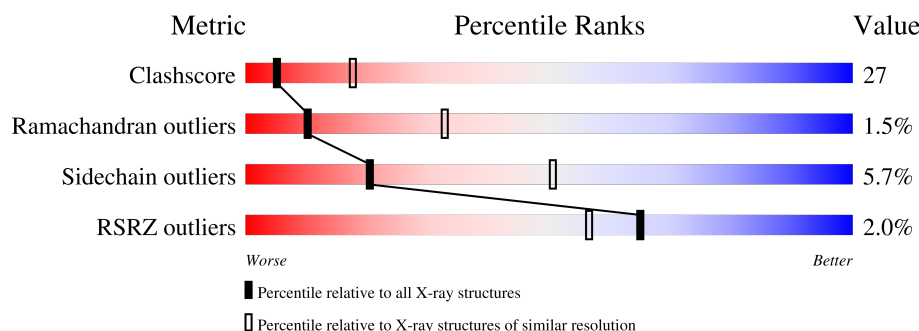
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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.







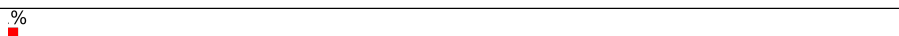
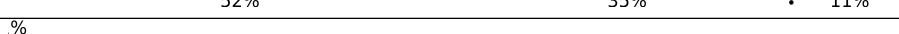

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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

Mol	Chain	Length	Quality of chain
1	A	85	<div> <div>2%</div> <div>36% 42% 6% 15%</div> </div>
1	B	85	<div> <div>53% 29% 5% 13%</div> </div>
1	C	85	<div> <div>2%</div> <div>44% 42% • 12%</div> </div>
1	D	85	<div> <div>46% 39% • 14%</div> </div>
1	E	85	<div> <div>55% 31% • 13%</div> </div>
1	F	85	<div> <div>%</div> <div>47% 32% 7% 14%</div> </div>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	G	85	
1	H	85	
1	I	85	
1	J	85	
1	K	85	
1	L	85	
1	M	85	
1	N	85	

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	F	109	-	-	-	X
3	EDO	H	112	-	-	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8226 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 Sm-like Archaeal Protein 1 (SmAP1).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	72	Total	C	N	O	S	0	0	0
			567	354	105	107	1			
1	B	74	Total	C	N	O	S	0	0	0
			581	363	107	110	1			
1	C	75	Total	C	N	O	S	0	0	0
			589	368	109	111	1			
1	D	73	Total	C	N	O	S	0	0	0
			574	359	106	108	1			
1	E	74	Total	C	N	O	S	0	0	0
			580	362	107	110	1			
1	F	73	Total	C	N	O	S	0	0	0
			574	359	106	108	1			
1	G	76	Total	C	N	O	S	0	0	0
			597	374	110	112	1			
1	H	73	Total	C	N	O	S	0	0	0
			573	357	106	109	1			
1	I	72	Total	C	N	O	S	0	0	0
			567	354	105	107	1			
1	J	71	Total	C	N	O	S	0	0	0
			558	349	101	107	1			
1	K	71	Total	C	N	O	S	0	0	0
			561	351	103	106	1			
1	L	76	Total	C	N	O	S	0	0	0
			594	371	110	112	1			
1	M	74	Total	C	N	O	S	0	0	0
			581	363	107	110	1			
1	N	75	Total	C	N	O	S	0	0	0
			589	368	109	111	1			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	81	ARG	PRO	conflict	UNP O26745

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	GLY	-	insertion	UNP O26745
A	83	LYS	-	insertion	UNP O26745
A	84	LEU	-	insertion	UNP O26745
A	85	ALA	-	insertion	UNP O26745
B	81	ARG	PRO	conflict	UNP O26745
B	82	GLY	-	insertion	UNP O26745
B	83	LYS	-	insertion	UNP O26745
B	84	LEU	-	insertion	UNP O26745
B	85	ALA	-	insertion	UNP O26745
C	81	ARG	PRO	conflict	UNP O26745
C	82	GLY	-	insertion	UNP O26745
C	83	LYS	-	insertion	UNP O26745
C	84	LEU	-	insertion	UNP O26745
C	85	ALA	-	insertion	UNP O26745
D	81	ARG	PRO	conflict	UNP O26745
D	82	GLY	-	insertion	UNP O26745
D	83	LYS	-	insertion	UNP O26745
D	84	LEU	-	insertion	UNP O26745
D	85	ALA	-	insertion	UNP O26745
E	81	ARG	PRO	conflict	UNP O26745
E	82	GLY	-	insertion	UNP O26745
E	83	LYS	-	insertion	UNP O26745
E	84	LEU	-	insertion	UNP O26745
E	85	ALA	-	insertion	UNP O26745
F	81	ARG	PRO	conflict	UNP O26745
F	82	GLY	-	insertion	UNP O26745
F	83	LYS	-	insertion	UNP O26745
F	84	LEU	-	insertion	UNP O26745
F	85	ALA	-	insertion	UNP O26745
G	81	ARG	PRO	conflict	UNP O26745
G	82	GLY	-	insertion	UNP O26745
G	83	LYS	-	insertion	UNP O26745
G	84	LEU	-	insertion	UNP O26745
G	85	ALA	-	insertion	UNP O26745
H	81	ARG	PRO	conflict	UNP O26745
H	82	GLY	-	insertion	UNP O26745
H	83	LYS	-	insertion	UNP O26745
H	84	LEU	-	insertion	UNP O26745
H	85	ALA	-	insertion	UNP O26745
I	81	ARG	PRO	conflict	UNP O26745
I	82	GLY	-	insertion	UNP O26745
I	83	LYS	-	insertion	UNP O26745

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
I	84	LEU	-	insertion	UNP O26745
I	85	ALA	-	insertion	UNP O26745
J	81	ARG	PRO	conflict	UNP O26745
J	82	GLY	-	insertion	UNP O26745
J	83	LYS	-	insertion	UNP O26745
J	84	LEU	-	insertion	UNP O26745
J	85	ALA	-	insertion	UNP O26745
K	81	ARG	PRO	conflict	UNP O26745
K	82	GLY	-	insertion	UNP O26745
K	83	LYS	-	insertion	UNP O26745
K	84	LEU	-	insertion	UNP O26745
K	85	ALA	-	insertion	UNP O26745
L	81	ARG	PRO	conflict	UNP O26745
L	82	GLY	-	insertion	UNP O26745
L	83	LYS	-	insertion	UNP O26745
L	84	LEU	-	insertion	UNP O26745
L	85	ALA	-	insertion	UNP O26745
M	81	ARG	PRO	conflict	UNP O26745
M	82	GLY	-	insertion	UNP O26745
M	83	LYS	-	insertion	UNP O26745
M	84	LEU	-	insertion	UNP O26745
M	85	ALA	-	insertion	UNP O26745
N	81	ARG	PRO	conflict	UNP O26745
N	82	GLY	-	insertion	UNP O26745
N	83	LYS	-	insertion	UNP O26745
N	84	LEU	-	insertion	UNP O26745
N	85	ALA	-	insertion	UNP O26745

- Molecule 2 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total Cl 1 1	0	0
2	G	1	Total Cl 1 1	0	0
2	L	1	Total Cl 1 1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			4	2	2		
3	D	1	Total	C	O	0	0
			4	2	2		
3	E	1	Total	C	O	0	0
			4	2	2		
3	F	1	Total	C	O	0	0
			4	2	2		
3	G	1	Total	C	O	0	0
			4	2	2		
3	H	1	Total	C	O	0	0
			4	2	2		
3	H	1	Total	C	O	0	0
			4	2	2		
3	I	1	Total	C	O	0	0
			4	2	2		
3	L	1	Total	C	O	0	0
			4	2	2		
3	M	1	Total	C	O	0	0
			4	2	2		
3	N	1	Total	C	O	0	0
			4	2	2		
3	N	1	Total	C	O	0	0
			4	2	2		
3	N	1	Total	C	O	0	0
			4	2	2		

- Molecule 4 is water.

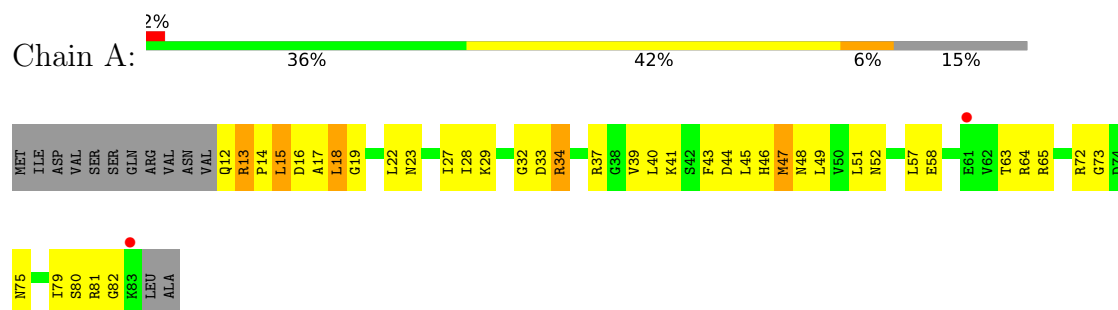
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	3	Total O 3 3	0	0
4	B	6	Total O 6 6	0	0
4	C	9	Total O 9 9	0	0
4	D	6	Total O 6 6	0	0
4	E	9	Total O 9 9	0	0
4	F	7	Total O 7 7	0	0
4	G	4	Total O 4 4	0	0
4	H	3	Total O 3 3	0	0
4	I	3	Total O 3 3	0	0
4	J	2	Total O 2 2	0	0
4	K	2	Total O 2 2	0	0
4	L	13	Total O 13 13	0	0
4	M	10	Total O 10 10	0	0
4	N	9	Total O 9 9	0	0



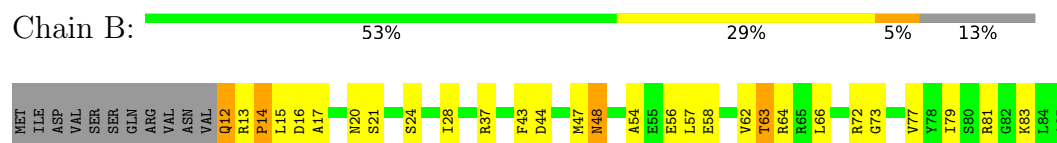
### 3 Residue-property plots [i](#)

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.

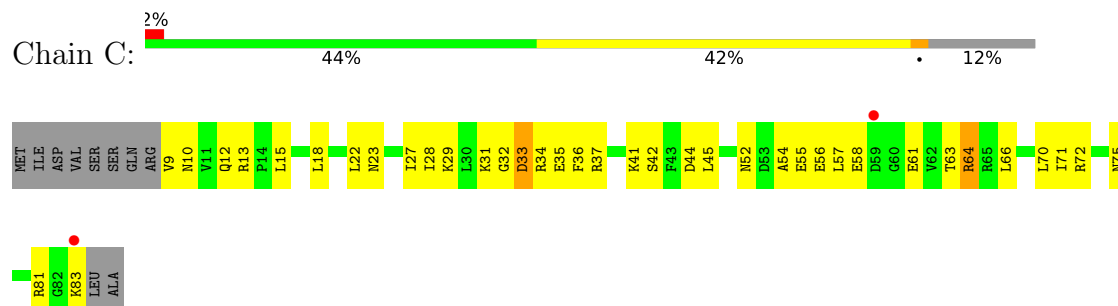
#### • Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



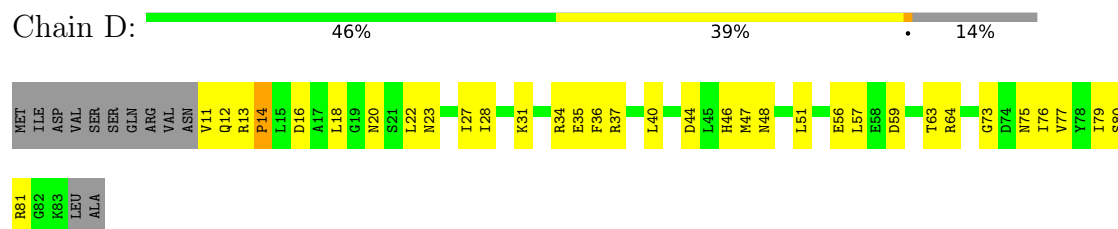
#### • Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



#### • Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

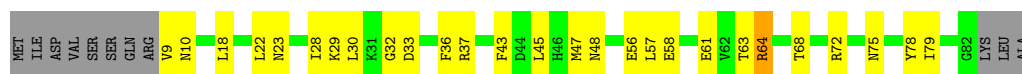


#### • Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



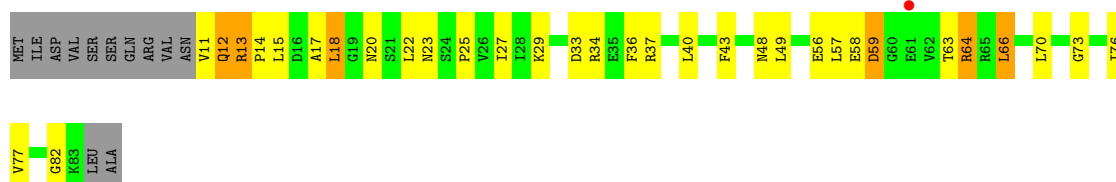
- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

Chain E: 



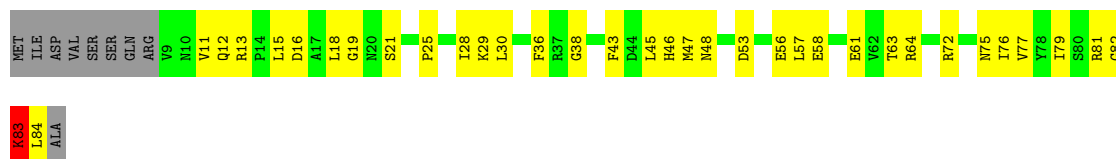
- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

Chain F: 



- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

Chain G: 



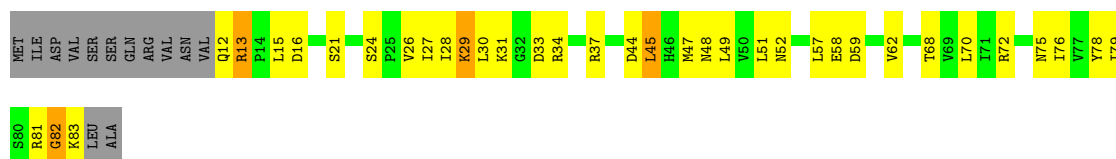
- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

Chain H: 



- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

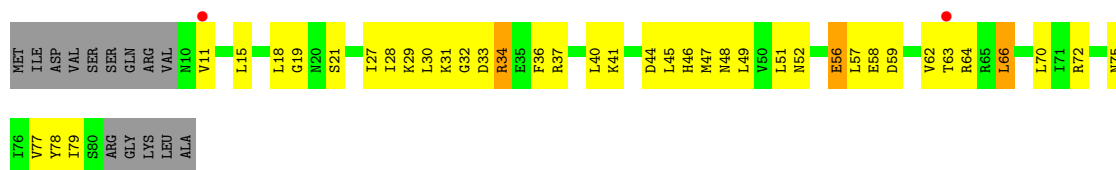
Chain I: 



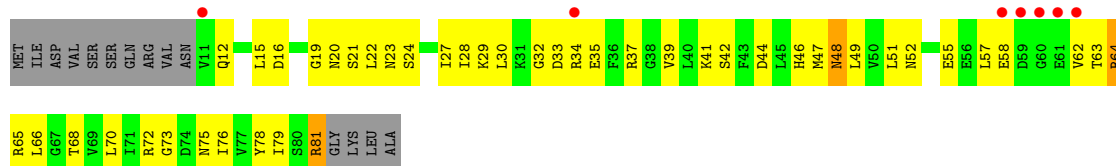
- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)

Chain J: 





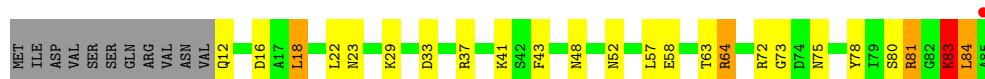
- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



- Molecule 1: Sm-like Archaeal Protein 1 (SmAP1)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	40.37Å 114.70Å 238.60Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	14.95 – 2.80 82.68 – 2.79	Depositor EDS
% Data completeness (in resolution range)	94.9 (14.95-2.80) 94.5 (82.68-2.79)	Depositor EDS
$R_{merge}$	0.11	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.82 (at 2.77Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.199 , 0.291 0.198 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	54.1	Xtriage
Anisotropy	0.124	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.35 , 72.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	8226	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	50.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.42% of the height of the origin peak. No significant pseudotranslation is detected.*

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

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: EDO, CL

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	0.34	0/572	0.72	0/768
1	B	0.37	0/586	0.67	1/786 (0.1%)
1	C	0.34	0/594	0.67	0/799
1	D	0.35	0/579	0.71	0/778
1	E	0.38	0/585	0.69	0/788
1	F	0.38	0/579	0.71	1/778 (0.1%)
1	G	0.36	0/602	0.68	0/810
1	H	0.37	0/578	0.68	0/778
1	I	0.35	0/572	0.70	0/768
1	J	0.37	0/563	0.66	0/759
1	K	0.34	0/566	0.64	1/762 (0.1%)
1	L	0.37	0/599	0.69	0/806
1	M	0.39	0/586	0.70	0/786
1	N	0.34	0/594	0.69	0/799
All	All	0.36	0/8155	0.69	3/10965 (0.0%)

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	48	ASN	N-CA-C	-5.79	95.37	111.00
1	F	48	ASN	N-CA-C	-5.65	95.75	111.00
1	K	48	ASN	N-CA-C	-5.07	97.31	111.00

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	567	0	586	45	0
1	B	581	0	602	30	0
1	C	589	0	610	36	0
1	D	574	0	595	32	0
1	E	580	0	597	26	0
1	F	574	0	595	35	0
1	G	597	0	621	36	0
1	H	573	0	588	30	0
1	I	567	0	586	46	0
1	J	558	0	572	37	0
1	K	561	0	579	51	0
1	L	594	0	612	34	0
1	M	581	0	602	31	0
1	N	589	0	610	33	0
2	B	1	0	0	0	0
2	G	1	0	0	0	0
2	L	1	0	0	0	0
3	C	4	0	5	0	0
3	D	4	0	5	0	0
3	E	4	0	5	0	0
3	F	4	0	5	2	0
3	G	4	0	5	2	0
3	H	8	0	10	1	0
3	I	4	0	5	0	0
3	L	4	0	5	1	0
3	M	4	0	5	1	0
3	N	12	0	15	1	0
4	A	3	0	0	0	0
4	B	6	0	0	0	0
4	C	9	0	0	1	0
4	D	6	0	0	0	0
4	E	9	0	0	0	0
4	F	7	0	0	0	0
4	G	4	0	0	1	0
4	H	3	0	0	0	0
4	I	3	0	0	1	0
4	J	2	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	K	2	0	0	0	0
4	L	13	0	0	0	0
4	M	10	0	0	1	0
4	N	9	0	0	0	0
All	All	8226	0	8420	444	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 27.

The worst 5 of 444 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:K:81:ARG:HH21	1:L:67:GLY:HA3	1.00	1.12
1:K:41:LYS:HE3	1:K:52:ASN:HD22	1.16	1.02
1:I:12:GLN:HG3	1:I:13:ARG:HH12	1.26	1.01
1:C:9:VAL:HG12	1:C:10:ASN:H	1.27	0.99
1:K:81:ARG:NH2	1:L:67:GLY:HA3	1.78	0.99

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	70/85 (82%)	63 (90%)	4 (6%)	3 (4%)	2	8
1	B	72/85 (85%)	67 (93%)	4 (6%)	1 (1%)	11	34
1	C	73/85 (86%)	70 (96%)	3 (4%)	0	100	100
1	D	71/85 (84%)	69 (97%)	2 (3%)	0	100	100
1	E	72/85 (85%)	65 (90%)	7 (10%)	0	100	100
1	F	71/85 (84%)	64 (90%)	5 (7%)	2 (3%)	5	17

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	74/85 (87%)	66 (89%)	7 (10%)	1 (1%)	11	34
1	H	71/85 (84%)	61 (86%)	9 (13%)	1 (1%)	11	34
1	I	70/85 (82%)	63 (90%)	5 (7%)	2 (3%)	4	15
1	J	69/85 (81%)	60 (87%)	9 (13%)	0	100	100
1	K	69/85 (81%)	60 (87%)	8 (12%)	1 (1%)	11	34
1	L	74/85 (87%)	67 (90%)	6 (8%)	1 (1%)	11	34
1	M	72/85 (85%)	65 (90%)	5 (7%)	2 (3%)	5	17
1	N	73/85 (86%)	68 (93%)	4 (6%)	1 (1%)	11	34
All	All	1001/1190 (84%)	908 (91%)	78 (8%)	15 (2%)	10	33

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	12	GLN
1	G	83	LYS
1	I	82	GLY
1	M	83	LYS
1	N	11	VAL

### 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	63/75 (84%)	57 (90%)	6 (10%)	8	25
1	B	64/75 (85%)	60 (94%)	4 (6%)	18	46
1	C	66/75 (88%)	63 (96%)	3 (4%)	27	60
1	D	64/75 (85%)	61 (95%)	3 (5%)	26	59
1	E	65/75 (87%)	63 (97%)	2 (3%)	40	74
1	F	64/75 (85%)	59 (92%)	5 (8%)	12	35
1	G	67/75 (89%)	65 (97%)	2 (3%)	41	75
1	H	64/75 (85%)	58 (91%)	6 (9%)	8	26

Continued on next page...



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	I	63/75 (84%)	59 (94%)	4 (6%)	18	46
1	J	63/75 (84%)	59 (94%)	4 (6%)	18	46
1	K	63/75 (84%)	61 (97%)	2 (3%)	39	73
1	L	66/75 (88%)	63 (96%)	3 (4%)	27	60
1	M	64/75 (85%)	60 (94%)	4 (6%)	18	46
1	N	66/75 (88%)	63 (96%)	3 (4%)	27	60
All	All	902/1050 (86%)	851 (94%)	51 (6%)	20	50

5 of 51 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	H	37	ARG
1	J	34	ARG
1	N	61	GLU
1	H	47	MET
1	I	45	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 24 such sidechains are listed below:

Mol	Chain	Res	Type
1	G	75	ASN
1	J	52	ASN
1	I	75	ASN
1	K	20	ASN
1	D	23	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no monosaccharides in this entry.

## 5.6 Ligand geometry

Of 16 ligands modelled in this entry, 3 are monoatomic - leaving 13 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
3	EDO	M	108	-	3,3,3	1.78	1 (33%)	2,2,2	0.43	0
3	EDO	H	112	-	3,3,3	1.68	1 (33%)	2,2,2	0.50	0
3	EDO	H	111	-	3,3,3	1.78	1 (33%)	2,2,2	0.44	0
3	EDO	F	109	-	3,3,3	1.70	1 (33%)	2,2,2	0.55	0
3	EDO	N	106	-	3,3,3	1.75	1 (33%)	2,2,2	0.54	0
3	EDO	N	105	-	3,3,3	1.81	1 (33%)	2,2,2	0.49	0
3	EDO	G	104	-	3,3,3	1.71	1 (33%)	2,2,2	0.52	0
3	EDO	L	102	-	3,3,3	1.69	1 (33%)	2,2,2	0.49	0
3	EDO	I	113	-	3,3,3	1.68	1 (33%)	2,2,2	0.53	0
3	EDO	E	103	-	3,3,3	1.81	1 (33%)	2,2,2	0.56	0
3	EDO	C	101	-	3,3,3	1.68	1 (33%)	2,2,2	0.57	0
3	EDO	N	110	-	3,3,3	1.73	1 (33%)	2,2,2	0.47	0
3	EDO	D	107	-	3,3,3	1.71	1 (33%)	2,2,2	0.59	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	M	108	-	-	1/1/1/1	-
3	EDO	H	112	-	-	1/1/1/1	-
3	EDO	H	111	-	-	1/1/1/1	-
3	EDO	F	109	-	-	1/1/1/1	-
3	EDO	N	106	-	-	1/1/1/1	-
3	EDO	N	105	-	-	1/1/1/1	-
3	EDO	G	104	-	-	1/1/1/1	-
3	EDO	L	102	-	-	1/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	EDO	I	113	-	-	1/1/1/1	-
3	EDO	E	103	-	-	1/1/1/1	-
3	EDO	C	101	-	-	1/1/1/1	-
3	EDO	N	110	-	-	1/1/1/1	-
3	EDO	D	107	-	-	1/1/1/1	-

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	N	105	EDO	O2-C2	2.67	1.55	1.42
3	M	108	EDO	O2-C2	2.63	1.55	1.42
3	E	103	EDO	O2-C2	2.62	1.55	1.42
3	H	111	EDO	O2-C2	2.60	1.55	1.42
3	N	106	EDO	O2-C2	2.55	1.55	1.42

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	E	103	EDO	O1-C1-C2-O2
3	F	109	EDO	O1-C1-C2-O2
3	G	104	EDO	O1-C1-C2-O2
3	N	105	EDO	O1-C1-C2-O2
3	N	106	EDO	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	M	108	EDO	1	0
3	H	111	EDO	1	0
3	F	109	EDO	2	0
3	N	105	EDO	1	0
3	G	104	EDO	2	0
3	L	102	EDO	1	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<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	72/85 (84%)	0.20	2 (2%) 53 43	31, 51, 95, 100	0
1	B	74/85 (87%)	0.04	0 100 100	25, 43, 81, 93	0
1	C	75/85 (88%)	-0.03	2 (2%) 54 44	26, 43, 79, 109	0
1	D	73/85 (85%)	-0.01	0 100 100	29, 46, 86, 95	0
1	E	74/85 (87%)	-0.14	0 100 100	28, 42, 81, 97	0
1	F	73/85 (85%)	-0.12	1 (1%) 75 70	26, 38, 77, 94	0
1	G	76/85 (89%)	-0.06	0 100 100	31, 46, 86, 94	0
1	H	73/85 (85%)	0.10	2 (2%) 54 44	25, 47, 88, 97	0
1	I	72/85 (84%)	-0.05	0 100 100	27, 43, 77, 86	0
1	J	71/85 (83%)	0.28	2 (2%) 53 43	36, 54, 85, 97	0
1	K	71/85 (83%)	0.62	7 (9%) 7 4	35, 55, 92, 96	0
1	L	76/85 (89%)	-0.07	1 (1%) 77 72	25, 43, 63, 73	0
1	M	74/85 (87%)	-0.10	1 (1%) 75 70	24, 39, 87, 121	0
1	N	75/85 (88%)	-0.08	3 (4%) 38 28	25, 43, 77, 101	0
All	All	1029/1190 (86%)	0.04	21 (2%) 65 56	24, 46, 86, 121	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	K	61	GLU	7.3
1	H	13	ARG	4.0
1	N	12	GLN	3.5
1	K	60	GLY	3.5
1	A	61	GLU	3.4

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

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

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
3	EDO	F	109	4/4	0.60	0.49	51,65,65,71	0
3	EDO	H	112	4/4	0.62	0.43	62,63,71,72	0
3	EDO	D	107	4/4	0.68	0.27	71,72,74,77	0
3	EDO	M	108	4/4	0.71	0.22	50,67,69,76	0
3	EDO	G	104	4/4	0.73	0.33	46,53,53,61	0
3	EDO	H	111	4/4	0.79	0.28	65,73,74,77	0
3	EDO	N	105	4/4	0.82	0.38	48,62,65,71	0
3	EDO	L	102	4/4	0.87	0.39	40,52,66,66	0
3	EDO	N	106	4/4	0.87	0.36	62,68,69,76	0
3	EDO	E	103	4/4	0.90	0.41	45,48,49,50	0
3	EDO	I	113	4/4	0.90	0.33	49,49,59,62	0
3	EDO	N	110	4/4	0.91	0.37	52,59,69,74	0
2	CL	G	203	1/1	0.95	0.25	73,73,73,73	0
3	EDO	C	101	4/4	0.95	0.39	49,56,61,65	0
2	CL	L	201	1/1	0.97	0.20	51,51,51,51	0
2	CL	B	202	1/1	0.97	0.19	60,60,60,60	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.