



# wwPDB X-ray Structure Validation Summary Report ⓘ

Nov 15, 2023 – 02:52 PM JST

PDB ID : 6JE4  
Title : Crystal structure of Nme1Cas9-sgRNA-dsDNA dimer mediated by double protein inhibitor AcrIIC3 monomers  
Authors : Sun, W.; Yang, J.; Cheng, Z.; Liu, C.; Wang, K.; Huang, X.; Wang, Y.  
Deposited on : 2019-02-04  
Resolution : 3.07 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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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

MolProbity : 4.02b-467  
Mogul : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.36  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.36

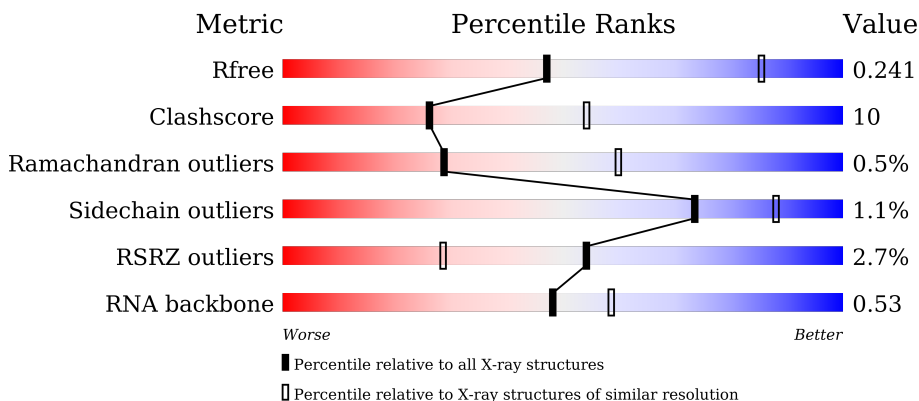
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.07 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	1754 (3.10-3.02)
Clashscore	141614	1864 (3.10-3.02)
Ramachandran outliers	138981	1794 (3.10-3.02)
Sidechain outliers	138945	1793 (3.10-3.02)
RSRZ outliers	127900	1713 (3.10-3.02)
RNA backbone	3102	1036 (3.32-2.80)

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	1083	 5% 76% 20% ..
1	E	1083	 % 77% 19% ..
1	K	1083	 % 77% 20% ..
1	O	1083	 5% 75% 21% ..

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	C	35	37% 63%
2	G	35	43% 54% .
2	M	35	46% 51% .
2	Q	35	40% 57% .
3	D	35	34% 9% 57%
3	H	35	31% 20% 49%
3	N	35	6% 26% 26% 49%
3	R	35	31% 20% 49%
4	B	135	35% 48% 11% . .
4	F	135	40% 40% 17% . .
4	L	135	41% 41% 16% . .
4	P	135	36% 45% 13% . .
5	I	117	3% 78% 21% .
5	J	117	% 82% 17% .
5	S	117	2% 79% 20% .
5	T	117	8% 83% 16% .

## 2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 52523 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 CRISPR-associated endonuclease Cas9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	1049	8261	5222	1509	1506	24	0	0	0
1	E	1056	8360	5283	1528	1525	24	0	0	0
1	K	1060	8384	5300	1535	1525	24	0	0	0
1	O	1049	8223	5200	1506	1493	24	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP C9X1G5
A	16	ALA	ASP	engineered mutation	UNP C9X1G5
A	588	ALA	HIS	engineered mutation	UNP C9X1G5
E	0	SER	-	expression tag	UNP C9X1G5
E	16	ALA	ASP	engineered mutation	UNP C9X1G5
E	588	ALA	HIS	engineered mutation	UNP C9X1G5
K	0	SER	-	expression tag	UNP C9X1G5
K	16	ALA	ASP	engineered mutation	UNP C9X1G5
K	588	ALA	HIS	engineered mutation	UNP C9X1G5
O	0	SER	-	expression tag	UNP C9X1G5
O	16	ALA	ASP	engineered mutation	UNP C9X1G5
O	588	ALA	HIS	engineered mutation	UNP C9X1G5

- Molecule 2 is a DNA chain called target DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	C	35	721	346	140	201	34	0	0	0
2	G	35	721	346	140	201	34	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	M	35	Total	C	N	O	P	0	0	0
			721	346	140	201	34			
2	Q	35	Total	C	N	O	P	0	0	0
			721	346	140	201	34			

- Molecule 3 is a DNA chain called non-target DNA strand.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	15	Total	C	N	O	P	0	0	0
			310	150	54	91	15			
3	H	18	Total	C	N	O	P	0	0	0
			373	180	66	109	18			
3	N	18	Total	C	N	O	P	0	0	0
			373	180	66	109	18			
3	R	18	Total	C	N	O	P	0	0	0
			373	180	66	109	18			

- Molecule 4 is a RNA chain called sgRNA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	132	Total	C	N	O	P	0	0	0
			2786	1247	474	933	132			
4	F	132	Total	C	N	O	P	0	0	0
			2786	1247	474	933	132			
4	L	133	Total	C	N	O	P	0	0	0
			2806	1256	477	940	133			
4	P	130	Total	C	N	O	P	0	0	0
			2746	1229	470	917	130			

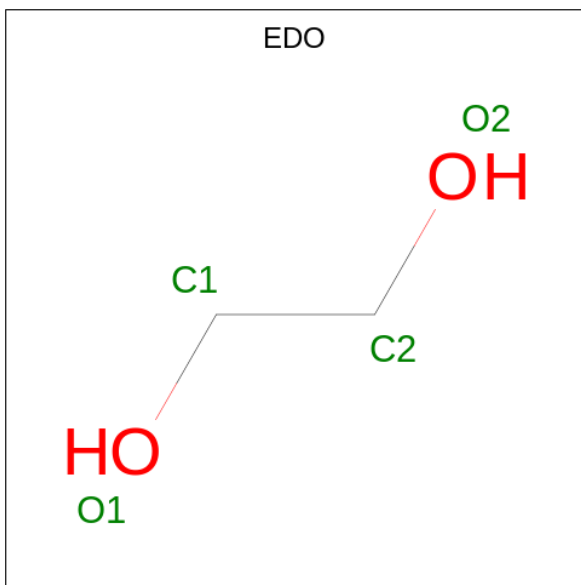
- Molecule 5 is a protein called AcrIIC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	I	117	Total	C	N	O	S	0	0	0
			911	584	154	170	3			
5	J	117	Total	C	N	O	S	0	0	0
			930	596	162	169	3			
5	S	117	Total	C	N	O	S	0	0	0
			940	601	162	174	3			
5	T	117	Total	C	N	O	S	0	0	0
			921	591	158	169	3			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
I	0	SER	-	expression tag	UNP A0A3E2QDI5
J	0	SER	-	expression tag	UNP A0A3E2QDI5
S	0	SER	-	expression tag	UNP A0A3E2QDI5
T	0	SER	-	expression tag	UNP A0A3E2QDI5

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



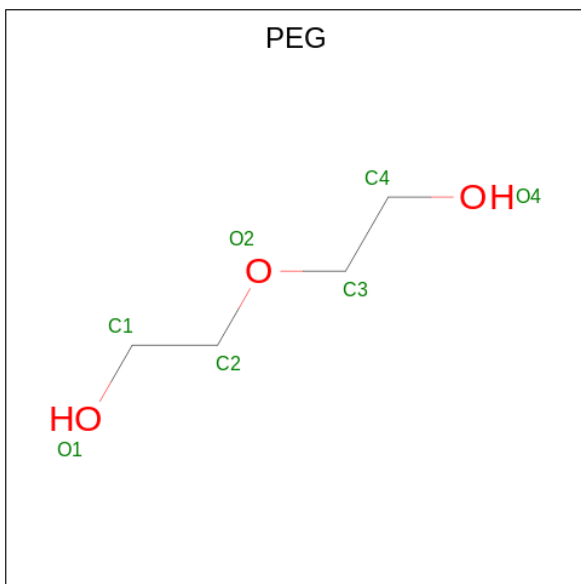
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C O 4 2 2	0	0
6	E	1	Total C O 4 2 2	0	0
6	E	1	Total C O 4 2 2	0	0
6	E	1	Total C O 4 2 2	0	0
6	F	1	Total C O 4 2 2	0	0
6	F	1	Total C O 4 2 2	0	0
6	J	1	Total C O 4 2 2	0	0
6	K	1	Total C O 4 2 2	0	0
6	K	1	Total C O 4 2 2	0	0
6	K	1	Total C O 4 2 2	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	K	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	L	1	Total	C	O	0	0
			4	2	2		
6	O	1	Total	C	O	0	0
			4	2	2		
6	O	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	B	1	Total	C	O	0	0
			7	4	3		
7	F	1	Total	C	O	0	0
			7	4	3		
7	K	1	Total	C	O	0	0
			7	4	3		

- Molecule 8 is water.

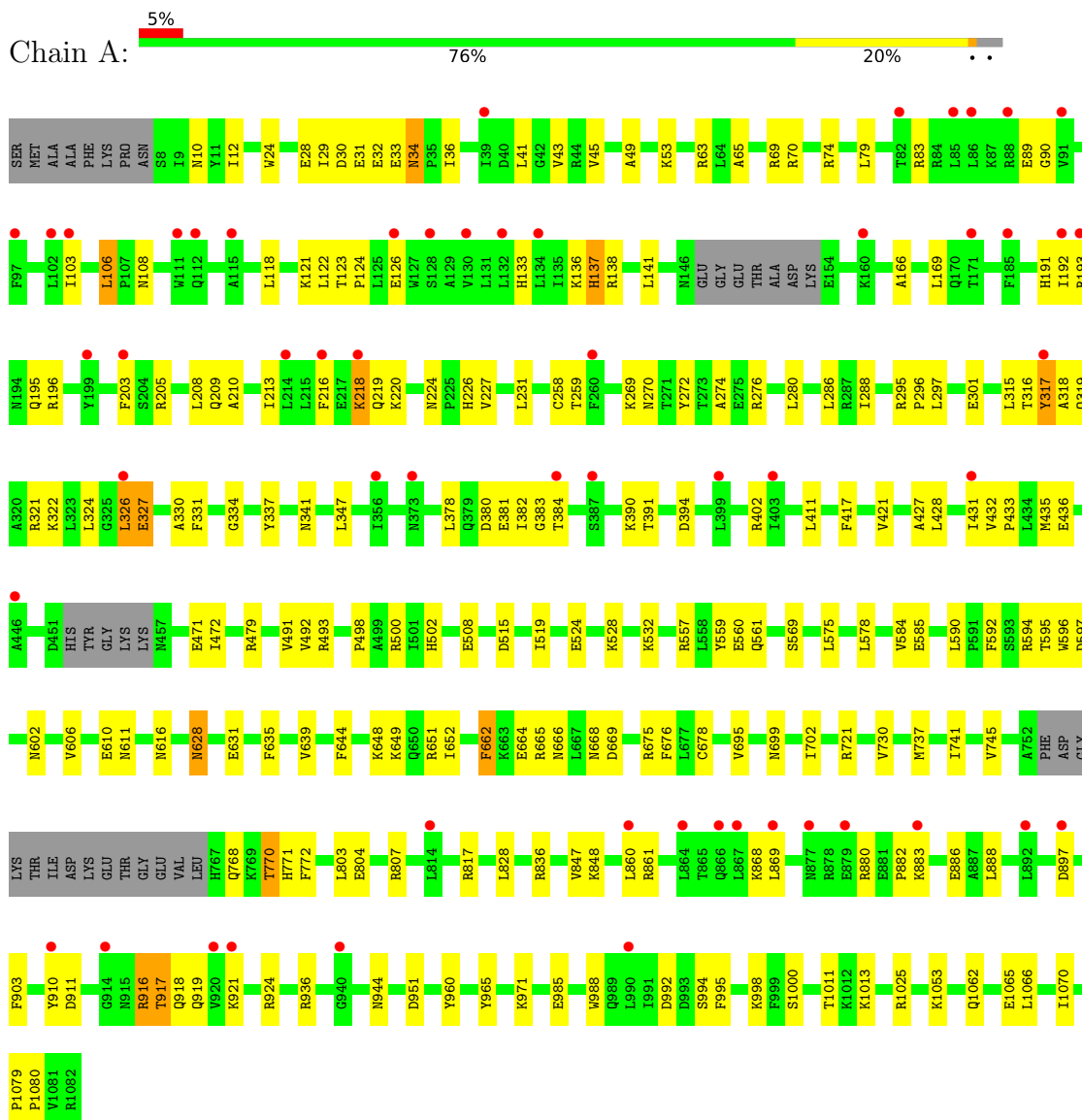
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	3	Total O 3 3	0	0
8	B	2	Total O 2 2	0	0
8	E	9	Total O 9 9	0	0
8	G	2	Total O 2 2	0	0
8	H	2	Total O 2 2	0	0
8	F	7	Total O 7 7	0	0
8	J	1	Total O 1 1	0	0
8	K	14	Total O 14 14	0	0
8	M	1	Total O 1 1	0	0
8	N	1	Total O 1 1	0	0
8	L	5	Total O 5 5	0	0
8	O	8	Total O 8 8	0	0
8	Q	1	Total O 1 1	0	0
8	P	4	Total O 4 4	0	0
8	S	3	Total O 3 3	0	0



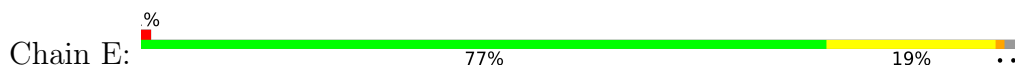
### 3 Residue-property plots [i](#)

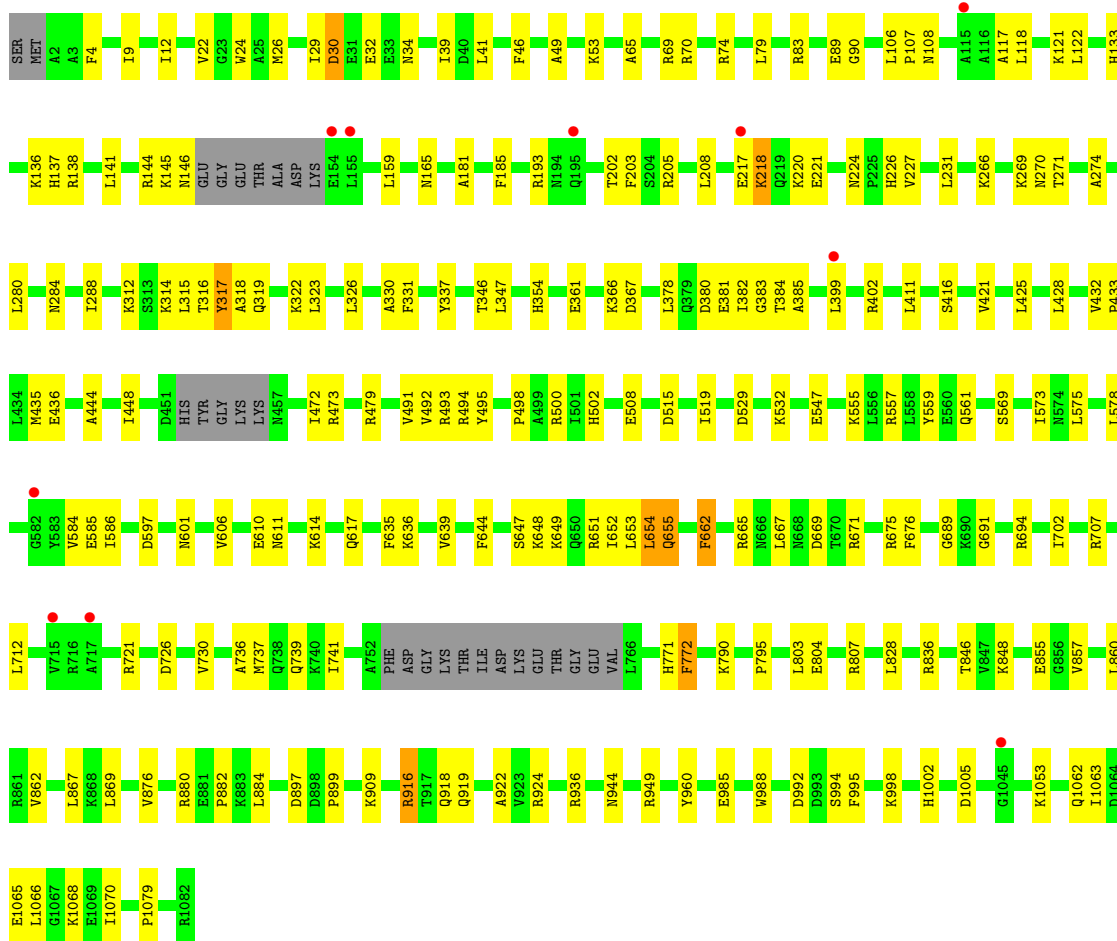
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: CRISPR-associated endonuclease Cas9

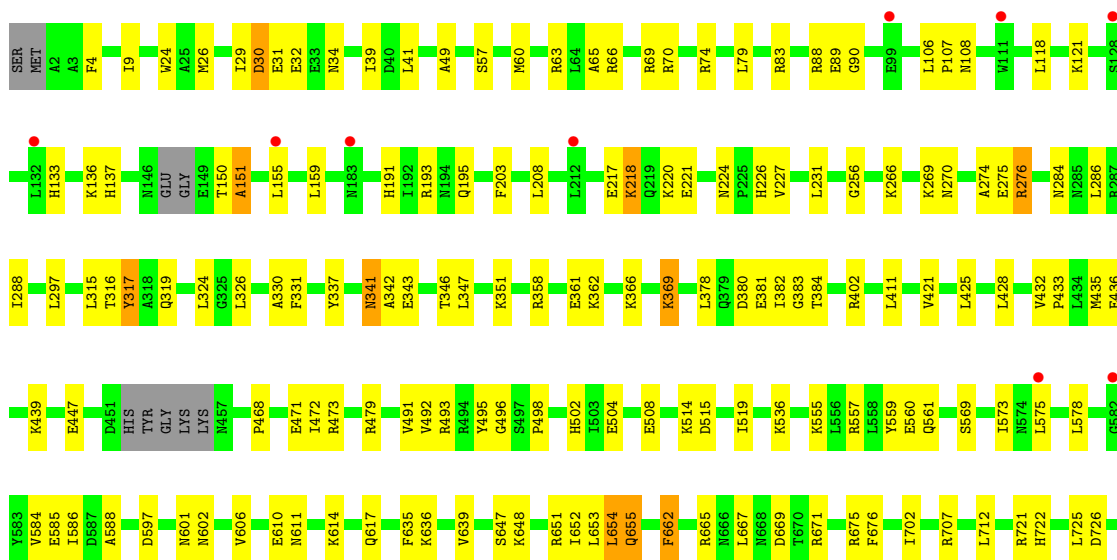
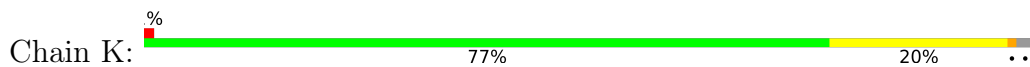


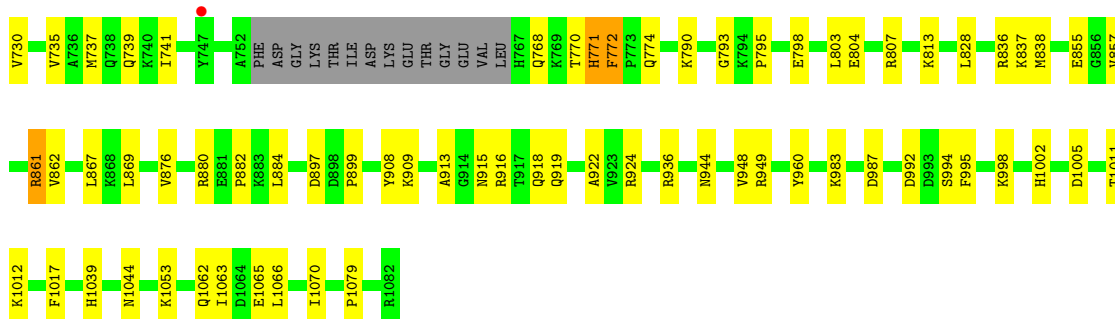
- Molecule 1: CRISPR-associated endonuclease Cas9



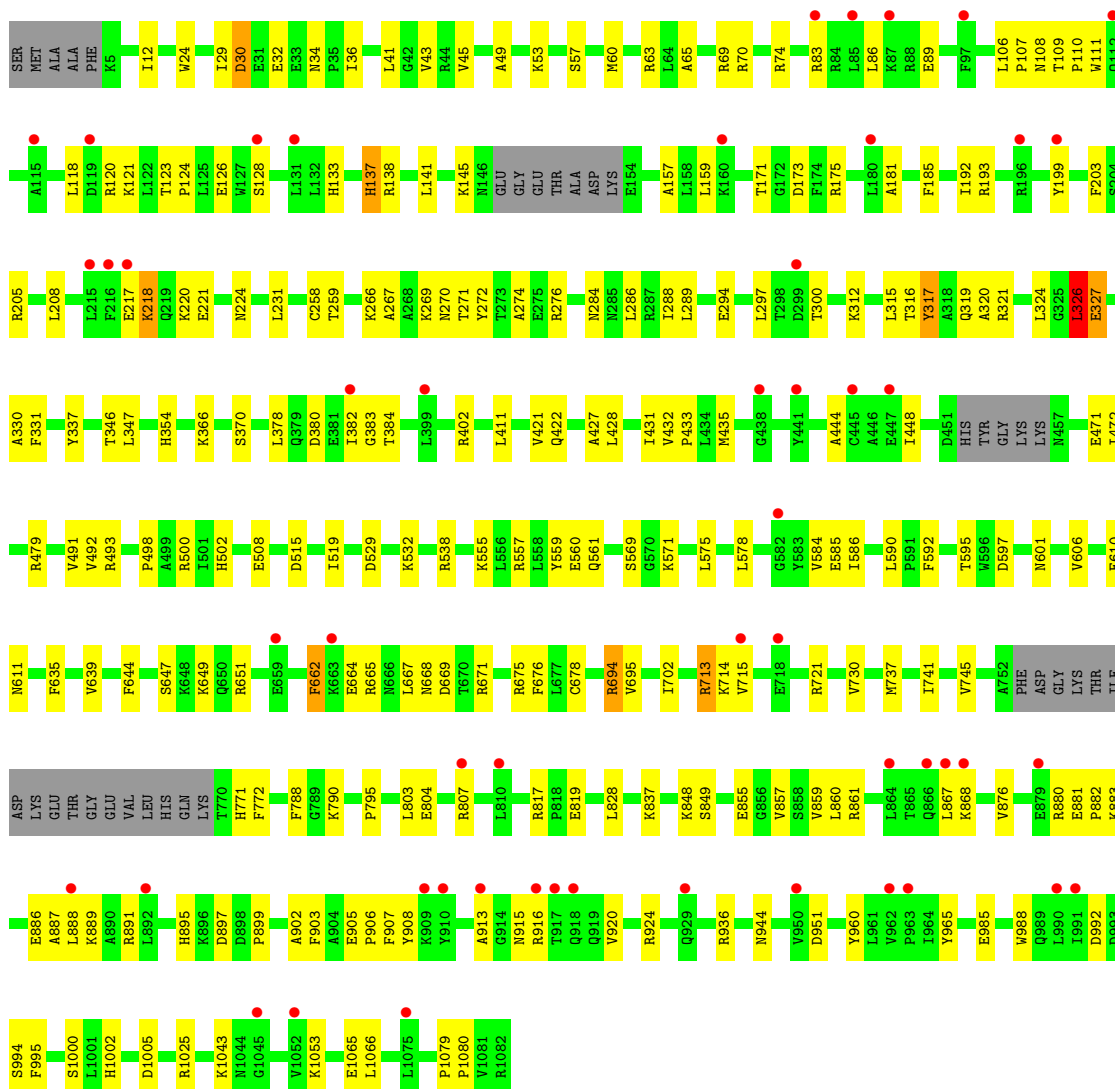
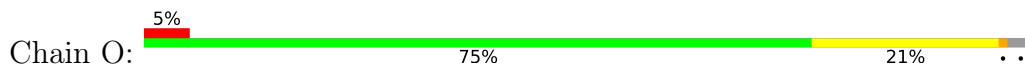


• Molecule 1: CRISPR-associated endonuclease Cas9





● Molecule 1: CRISPR-associated endonuclease Cas9



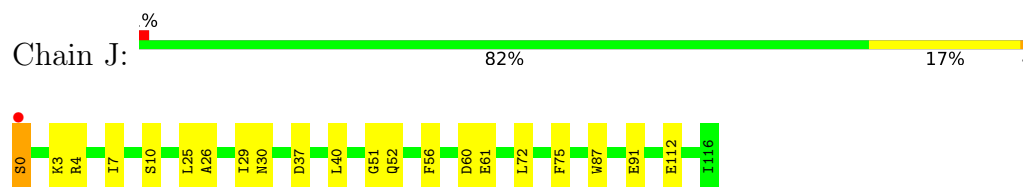
● Molecule 2: target DNA strand



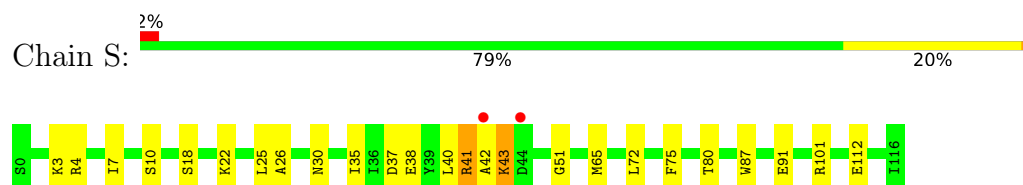




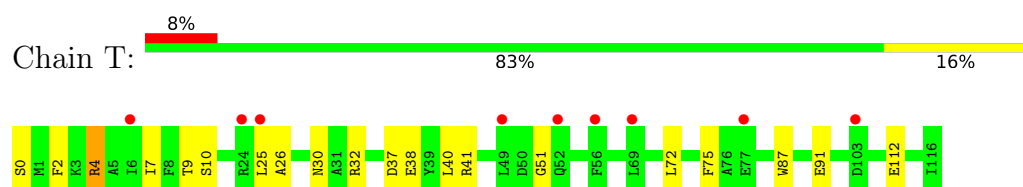
- Molecule 5: AcrIIC3



- Molecule 5: AcrIIC3



- Molecule 5: AcrIIC3



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	129.20Å 118.01Å 323.99Å 90.00° 89.92° 90.00°	Depositor
Resolution (Å)	46.28 – 3.07 46.28 – 3.07	Depositor EDS
% Data completeness (in resolution range)	87.7 (46.28-3.07) 87.4 (46.28-3.07)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.12 (at 3.06Å)	Xtrriage
Refinement program	PHENIX (1.14_3247: ???)	Depositor
R, $R_{free}$	0.210 , 0.241 0.210 , 0.241	Depositor DCC
$R_{free}$ test set	8047 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	70.1	Xtrriage
Anisotropy	0.236	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 32.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.439 for h,-k,-l	Xtrriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	52523	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	71.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.65% of the height of the origin peak. No significant pseudotranslation is detected.*

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

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

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.32	0/8418	0.51	0/11365
1	E	0.27	0/8520	0.51	0/11494
1	K	0.28	0/8545	0.51	0/11525
1	O	0.27	0/8380	0.50	0/11315
2	C	0.58	0/812	0.96	0/1252
2	G	0.61	0/812	1.00	2/1252 (0.2%)
2	M	0.66	0/812	1.06	3/1252 (0.2%)
2	Q	0.58	0/812	1.00	2/1252 (0.2%)
3	D	0.64	0/347	1.03	0/534
3	H	0.64	0/418	1.06	0/644
3	N	0.71	0/418	1.08	0/644
3	R	0.63	0/418	1.10	0/644
4	B	0.34	0/3107	1.01	7/4831 (0.1%)
4	F	0.38	0/3107	1.01	5/4831 (0.1%)
4	L	0.35	0/3129	0.98	4/4865 (0.1%)
4	P	0.35	0/3063	1.02	8/4763 (0.2%)
5	I	0.30	0/927	0.48	0/1250
5	J	0.28	0/946	0.47	0/1271
5	S	0.29	0/956	0.53	1/1283 (0.1%)
5	T	0.27	0/937	0.44	0/1261
All	All	0.35	0/54884	0.72	32/77528 (0.0%)

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	1
1	E	0	1
1	K	0	2

*Continued on next page...*



Continued from previous page...

Mol	Chain	#Chirality outliers	#Planarity outliers
1	O	0	1
5	I	0	1
All	All	0	6

There are no bond length outliers.

The worst 5 of 32 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	M	34	DC	P-O3'-C3'	8.84	130.31	119.70
2	Q	16	DA	O4'-C4'-C3'	-6.81	101.78	104.50
4	B	74	U	C5-C6-N1	6.39	125.89	122.70
2	Q	16	DA	O4'-C1'-N9	6.36	112.45	108.00
2	M	16	DA	O4'-C1'-N9	6.33	112.43	108.00

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	770	THR	Peptide
1	E	691	GLY	Peptide
5	I	4	ARG	Sidechain
1	K	150	THR	Peptide
1	K	151	ALA	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	8261	0	8139	162	0
1	E	8360	0	8269	145	0
1	K	8384	0	8303	159	1
1	O	8223	0	8083	160	1
2	C	721	0	396	26	0
2	G	721	0	396	24	0
2	M	721	0	396	23	0
2	Q	721	0	396	24	0

Continued on next page...

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	D	310	0	173	5	0
3	H	373	0	207	5	0
3	N	373	0	207	5	0
3	R	373	0	207	7	0
4	B	2786	0	1414	79	0
4	F	2786	0	1414	69	0
4	L	2806	0	1425	71	0
4	P	2746	0	1394	75	0
5	I	911	0	887	17	0
5	J	930	0	928	15	0
5	S	940	0	941	18	0
5	T	921	0	908	16	0
6	B	4	0	6	0	0
6	E	12	0	18	0	0
6	F	8	0	12	2	0
6	J	4	0	6	1	0
6	K	16	0	24	2	0
6	L	20	0	30	0	0
6	O	8	0	12	0	0
7	B	7	0	10	0	0
7	F	7	0	10	2	0
7	K	7	0	10	0	0
8	A	3	0	0	0	0
8	B	2	0	0	0	0
8	E	9	0	0	0	0
8	F	7	0	0	0	0
8	G	2	0	0	0	0
8	H	2	0	0	0	0
8	J	1	0	0	0	0
8	K	14	0	0	0	0
8	L	5	0	0	0	0
8	M	1	0	0	0	0
8	N	1	0	0	0	0
8	O	8	0	0	0	0
8	P	4	0	0	0	0
8	Q	1	0	0	0	0
8	S	3	0	0	0	0
All	All	52523	0	44621	1004	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 10.

The worst 5 of 1004 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:276:ARG:NH1	1:K:436:GLU:OE2	1.92	1.02
1:O:320:ALA:O	1:O:324:LEU:HD12	1.66	0.95
5:T:4:ARG:NH1	5:T:91:GLU:OE2	2.00	0.93
1:A:883:LYS:NZ	4:B:36:C:OP1	2.03	0.92
1:A:321:ARG:NH1	1:A:327:GLU:O	2.04	0.90

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:K:798:GLU:OE1	1:O:538:ARG:NH2[1_455]	2.13	0.07

## 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	1041/1083 (96%)	1001 (96%)	35 (3%)	5 (0%)	29	60
1	E	1048/1083 (97%)	1010 (96%)	33 (3%)	5 (0%)	29	60
1	K	1052/1083 (97%)	1011 (96%)	34 (3%)	7 (1%)	22	52
1	O	1041/1083 (96%)	1004 (96%)	32 (3%)	5 (0%)	29	60
5	I	115/117 (98%)	108 (94%)	7 (6%)	0	100	100
5	J	115/117 (98%)	110 (96%)	5 (4%)	0	100	100
5	S	115/117 (98%)	109 (95%)	5 (4%)	1 (1%)	17	47
5	T	115/117 (98%)	111 (96%)	4 (4%)	0	100	100
All	All	4642/4800 (97%)	4464 (96%)	155 (3%)	23 (0%)	29	60

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	771	HIS

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	E	655	GLN
1	K	342	ALA
1	K	655	GLN
1	K	772	PHE

### 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	840/931 (90%)	830 (99%)	10 (1%)	71	87
1	E	856/931 (92%)	848 (99%)	8 (1%)	78	90
1	K	857/931 (92%)	847 (99%)	10 (1%)	71	87
1	O	829/931 (89%)	819 (99%)	10 (1%)	71	87
5	I	91/99 (92%)	88 (97%)	3 (3%)	38	67
5	J	94/99 (95%)	93 (99%)	1 (1%)	73	88
5	S	97/99 (98%)	97 (100%)	0	100	100
5	T	92/99 (93%)	91 (99%)	1 (1%)	73	88
All	All	3756/4120 (91%)	3713 (99%)	43 (1%)	73	88

5 of 43 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	369	LYS
1	O	218	LYS
1	K	662	PHE
1	O	30	ASP
1	O	662	PHE

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

Mol	Chain	Res	Type
1	A	341	ASN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
4	B	130/135 (96%)	32 (24%)	3 (2%)
4	F	130/135 (96%)	34 (26%)	3 (2%)
4	L	131/135 (97%)	32 (24%)	6 (4%)
4	P	128/135 (94%)	35 (27%)	3 (2%)
All	All	519/540 (96%)	133 (25%)	15 (2%)

5 of 133 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
4	B	2	G
4	B	21	U
4	B	24	C
4	B	30	A
4	B	38	U

5 of 15 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
4	L	70	C
4	P	70	C
4	L	78	A
4	P	133	A
4	L	134	U

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

21 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
7	PEG	B	202	-	6,6,6	0.51	0	5,5,5	0.26	0
6	EDO	K	1104	-	3,3,3	0.46	0	2,2,2	0.38	0
6	EDO	L	201	-	3,3,3	0.47	0	2,2,2	0.36	0
6	EDO	E	1103	-	3,3,3	0.51	0	2,2,2	0.22	0
6	EDO	O	1101	-	3,3,3	0.46	0	2,2,2	0.35	0
6	EDO	K	1103	-	3,3,3	0.49	0	2,2,2	0.26	0
6	EDO	O	1102	-	3,3,3	0.47	0	2,2,2	0.32	0
6	EDO	L	204	-	3,3,3	0.49	0	2,2,2	0.28	0
6	EDO	K	1101	-	3,3,3	0.51	0	2,2,2	0.22	0
6	EDO	L	203	-	3,3,3	0.53	0	2,2,2	0.18	0
6	EDO	E	1101	-	3,3,3	0.50	0	2,2,2	0.25	0
6	EDO	K	1102	-	3,3,3	0.54	0	2,2,2	0.07	0
6	EDO	J	201	-	3,3,3	0.52	0	2,2,2	0.17	0
6	EDO	F	202	-	3,3,3	0.48	0	2,2,2	0.22	0
6	EDO	B	201	-	3,3,3	0.46	0	2,2,2	0.33	0
6	EDO	E	1102	-	3,3,3	0.48	0	2,2,2	0.30	0
6	EDO	L	202	-	3,3,3	0.51	0	2,2,2	0.24	0
6	EDO	F	201	-	3,3,3	0.46	0	2,2,2	0.35	0
6	EDO	L	205	-	3,3,3	0.45	0	2,2,2	0.28	0
7	PEG	F	203	-	6,6,6	0.48	0	5,5,5	0.26	0
7	PEG	K	1105	-	6,6,6	0.50	0	5,5,5	0.27	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
7	PEG	B	202	-	-	3/4/4/4	-
6	EDO	K	1104	-	-	1/1/1/1	-
6	EDO	L	201	-	-	0/1/1/1	-
6	EDO	E	1103	-	-	1/1/1/1	-
6	EDO	O	1101	-	-	0/1/1/1	-
6	EDO	K	1103	-	-	1/1/1/1	-
6	EDO	O	1102	-	-	0/1/1/1	-
6	EDO	L	204	-	-	1/1/1/1	-
6	EDO	K	1101	-	-	0/1/1/1	-
6	EDO	L	203	-	-	0/1/1/1	-

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
6	EDO	E	1101	-	-	0/1/1/1	-
6	EDO	K	1102	-	-	1/1/1/1	-
6	EDO	J	201	-	-	1/1/1/1	-
6	EDO	F	202	-	-	0/1/1/1	-
6	EDO	B	201	-	-	1/1/1/1	-
6	EDO	E	1102	-	-	0/1/1/1	-
6	EDO	L	202	-	-	0/1/1/1	-
6	EDO	F	201	-	-	0/1/1/1	-
6	EDO	L	205	-	-	0/1/1/1	-
7	PEG	F	203	-	-	2/4/4/4	-
7	PEG	K	1105	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
7	B	202	PEG	O2-C3-C4-O4
7	K	1105	PEG	O1-C1-C2-O2
6	B	201	EDO	O1-C1-C2-O2
6	L	204	EDO	O1-C1-C2-O2
7	F	203	PEG	O1-C1-C2-O2

There are no ring outliers.

6 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	K	1103	EDO	1	0
6	K	1102	EDO	1	0
6	J	201	EDO	1	0
6	F	202	EDO	1	0
6	F	201	EDO	1	0
7	F	203	PEG	2	0

## 5.7 Other polymers [i](#)

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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

In the following table, the column labelled ‘#RSRZ> 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q< 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	1049/1083 (96%)	0.47	55 (5%) 27 11	38, 75, 119, 141	0
1	E	1056/1083 (97%)	0.19	10 (0%) 84 66	26, 55, 84, 110	0
1	K	1060/1083 (97%)	0.20	10 (0%) 84 66	24, 54, 85, 130	0
1	O	1049/1083 (96%)	0.45	52 (4%) 28 12	37, 73, 121, 142	0
2	C	35/35 (100%)	-0.00	0 100 100	60, 76, 93, 96	0
2	G	35/35 (100%)	-0.10	0 100 100	32, 47, 66, 79	0
2	M	35/35 (100%)	0.10	0 100 100	32, 48, 64, 82	0
2	Q	35/35 (100%)	-0.06	0 100 100	60, 77, 92, 94	0
3	D	15/35 (42%)	-0.07	0 100 100	70, 80, 144, 156	0
3	H	18/35 (51%)	0.13	0 100 100	44, 63, 188, 194	0
3	N	18/35 (51%)	0.45	2 (11%) 5 2	43, 62, 192, 194	0
3	R	18/35 (51%)	0.21	0 100 100	68, 80, 206, 206	0
4	B	132/135 (97%)	0.02	0 100 100	53, 87, 158, 178	0
4	F	132/135 (97%)	0.03	0 100 100	33, 55, 135, 171	0
4	L	133/135 (98%)	0.15	0 100 100	32, 54, 140, 168	0
4	P	130/135 (96%)	-0.02	0 100 100	54, 86, 148, 162	0
5	I	117/117 (100%)	0.54	3 (2%) 56 30	54, 76, 101, 116	0
5	J	117/117 (100%)	0.36	1 (0%) 84 66	37, 58, 80, 91	0
5	S	117/117 (100%)	0.35	2 (1%) 70 46	38, 59, 81, 93	0
5	T	117/117 (100%)	0.64	9 (7%) 13 4	59, 76, 100, 110	0
All	All	5418/5620 (96%)	0.30	144 (2%) 54 28	24, 66, 117, 206	0

The worst 5 of 144 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	920	VAL	4.9
1	A	214	LEU	4.8
1	A	111	TRP	4.6
1	O	718	GLU	4.4
1	O	913	ALA	4.4

## 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
6	EDO	O	1102	4/4	0.72	0.26	58,63,64,70	0
6	EDO	K	1104	4/4	0.79	0.35	49,50,51,55	0
6	EDO	E	1101	4/4	0.83	0.27	47,56,56,57	0
6	EDO	J	201	4/4	0.85	0.18	49,56,63,63	0
6	EDO	K	1102	4/4	0.85	0.20	32,36,40,48	0
7	PEG	B	202	7/7	0.85	0.24	83,86,93,93	0
7	PEG	F	203	7/7	0.86	0.25	47,56,66,70	0
7	PEG	K	1105	7/7	0.86	0.21	59,65,75,81	0
6	EDO	L	204	4/4	0.87	0.18	61,65,69,71	0
6	EDO	L	202	4/4	0.88	0.22	52,58,58,61	0
6	EDO	K	1101	4/4	0.88	0.21	44,55,59,61	0
6	EDO	O	1101	4/4	0.89	0.20	63,63,64,65	0
6	EDO	E	1103	4/4	0.90	0.24	41,44,47,49	0
6	EDO	F	202	4/4	0.90	0.21	43,45,45,47	0
6	EDO	L	205	4/4	0.90	0.31	44,46,47,48	0
6	EDO	E	1102	4/4	0.90	0.16	59,63,64,68	0
6	EDO	F	201	4/4	0.93	0.15	43,44,45,47	0
6	EDO	L	201	4/4	0.94	0.20	37,44,44,46	0
6	EDO	K	1103	4/4	0.94	0.24	39,43,45,47	0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
6	EDO	B	201	4/4	0.94	0.24	67,69,70,71	0
6	EDO	L	203	4/4	0.95	0.22	48,51,51,56	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.