



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 07:56 PM UTC

PDB ID : 7B6C / pdb_00007b6c
Title : BK Polyomavirus VP1 pentamer fusion with long C-terminal extended arm
Authors : Osipov, E.M.; Beelen, S.; Strelkov, S.V.
Deposited on : 2020-12-07
Resolution : 2.48 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	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

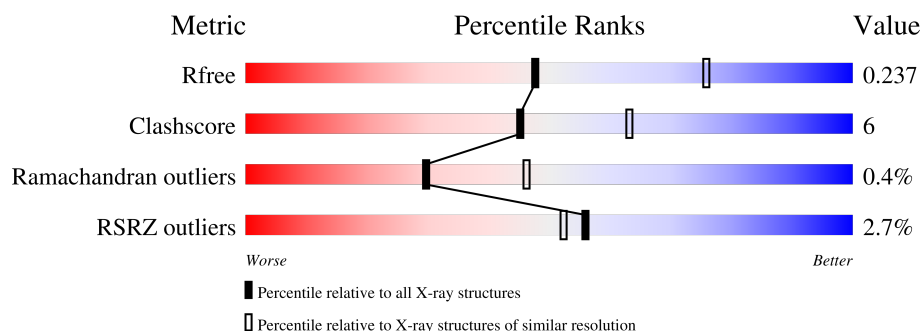
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.48 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}	180053	7589 (2.50-2.46)
Clashscore	190562	8295 (2.50-2.46)
Ramachandran outliers	187476	8164 (2.50-2.46)
RSRZ outliers	180081	7593 (2.50-2.46)

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	AAA	283	3% 85%13% .
1	BBB	283	2% 85%11% .
1	CCC	283	2% 84%14% ..
1	DDD	283	2% 82%16% ..
1	EEE	283	2% 86%13% .
1	FFF	283	3% 86%13% ..

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	GGG	283	3% 85% 13% ..
1	HHH	283	2% 83% 12% .
1	III	283	3% 81% 14% ..
1	JJJ	283	4% 84% 13% ..

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 22259 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 Major capsid protein VP1, Major capsid protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	279	Total	C	N	O	S	0	0	0
			2154	1347	371	421	15			
1	BBB	271	Total	C	N	O	S	0	0	0
			2092	1313	360	405	14			
1	CCC	279	Total	C	N	O	S	0	0	0
			2157	1347	373	422	15			
1	DDD	280	Total	C	N	O	S	0	0	0
			2155	1347	371	422	15			
1	EEE	279	Total	C	N	O	S	0	0	0
			2153	1347	373	418	15			
1	FFF	280	Total	C	N	O	S	0	0	0
			2174	1361	373	425	15			
1	GGG	279	Total	C	N	O	S	0	0	0
			2164	1353	375	421	15			
1	HHH	271	Total	C	N	O	S	0	0	0
			2102	1316	364	408	14			
1	III	272	Total	C	N	O	S	0	0	0
			2108	1321	362	411	14			
1	JJJ	276	Total	C	N	O	S	0	0	0
			2126	1331	367	414	14			

There are 10 discrepancies between the modelled and reference sequences:

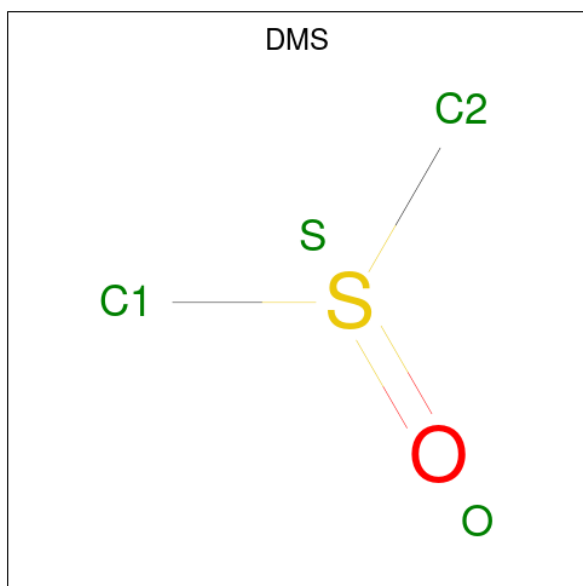
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	17	GLY	-	expression tag	UNP P03088
BBB	17	GLY	-	expression tag	UNP P03088
CCC	17	GLY	-	expression tag	UNP P03088
DDD	17	GLY	-	expression tag	UNP P03088
EEE	17	GLY	-	expression tag	UNP P03088
FFF	17	GLY	-	expression tag	UNP P03088
GGG	17	GLY	-	expression tag	UNP P03088
HHH	17	GLY	-	expression tag	UNP P03088
III	17	GLY	-	expression tag	UNP P03088

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
JJJ	17	GLY	-	expression tag	UNP P03088

- Molecule 2 is DIMETHYL SULFOXIDE (CCD ID: DMS) (formula: C₂H₆OS).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	CCC	1	Total	C	O	S	0	0
			4	2	1	1		
2	HHH	1	Total	C	O	S	0	0
			4	2	1	1		

- Molecule 3 is CALCIUM ION (CCD ID: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	DDD	1	Total	Ca	0	0
			1	1		
3	FFF	1	Total	Ca	0	0
			1	1		
3	III	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	AAA	93	Total	O	0	0
			93	93		

Continued on next page...

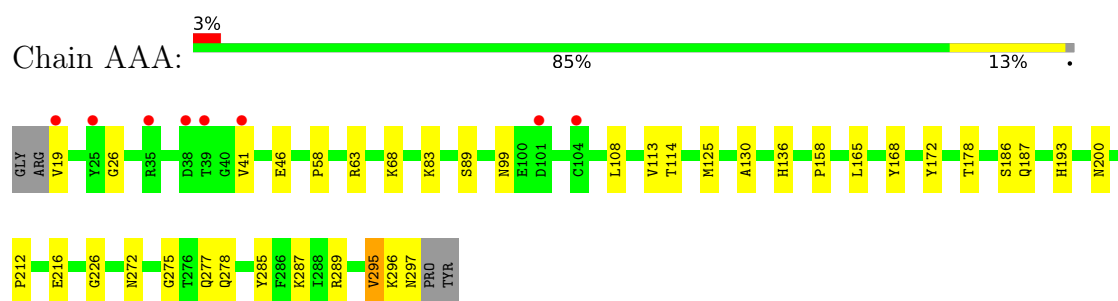
Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	BBB	89	Total 89	O 89	0	0
4	CCC	92	Total 92	O 92	0	0
4	DDD	88	Total 88	O 88	0	0
4	EEE	84	Total 84	O 84	0	0
4	FFF	98	Total 98	O 98	0	0
4	GGG	80	Total 80	O 80	0	0
4	HHH	64	Total 64	O 64	0	0
4	III	84	Total 84	O 84	0	0
4	JJJ	91	Total 91	O 91	0	0

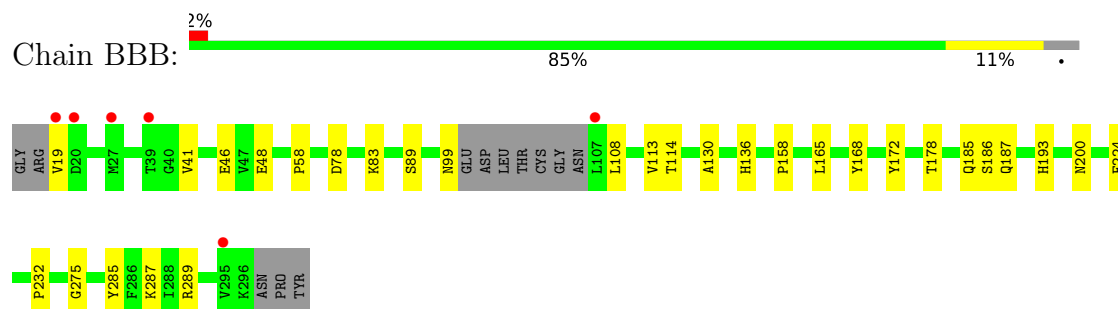
3 Residue-property plots [i](#)

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.

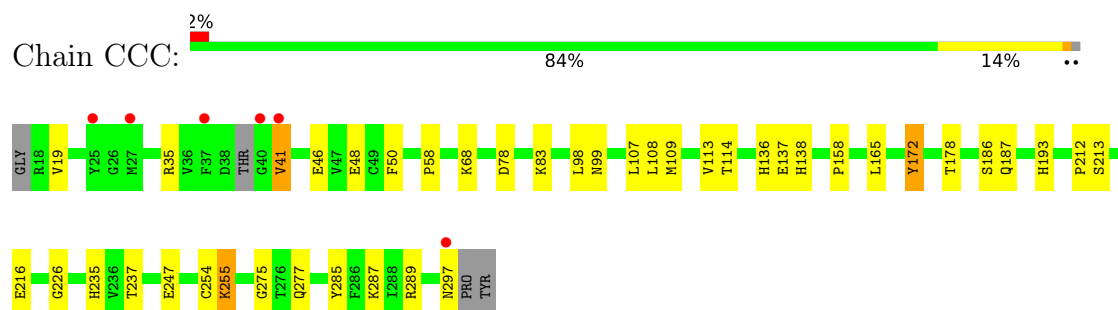
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



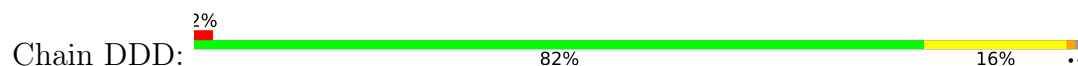
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1

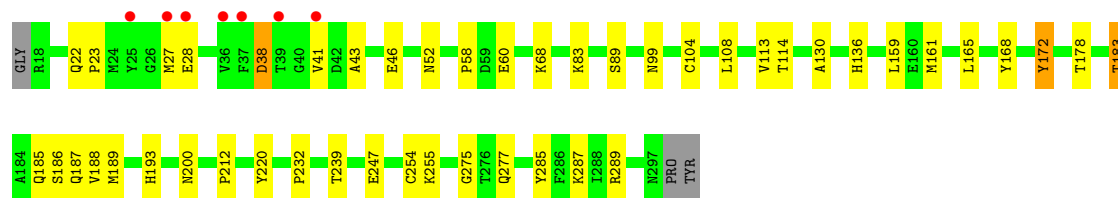


- Molecule 1: Major capsid protein VP1, Major capsid protein VP1

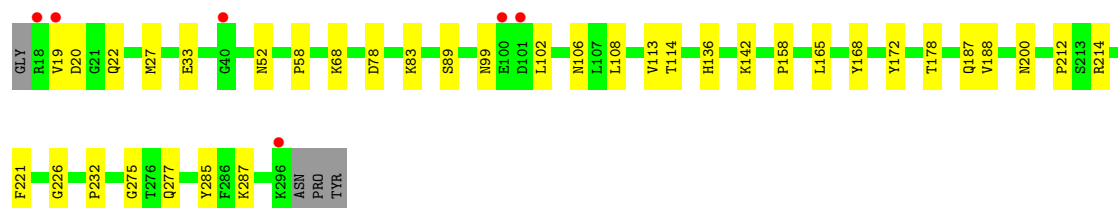
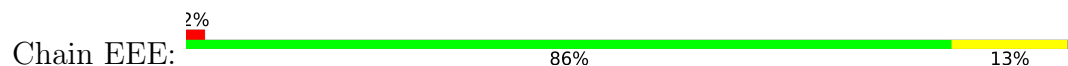


- Molecule 1: Major capsid protein VP1, Major capsid protein VP1

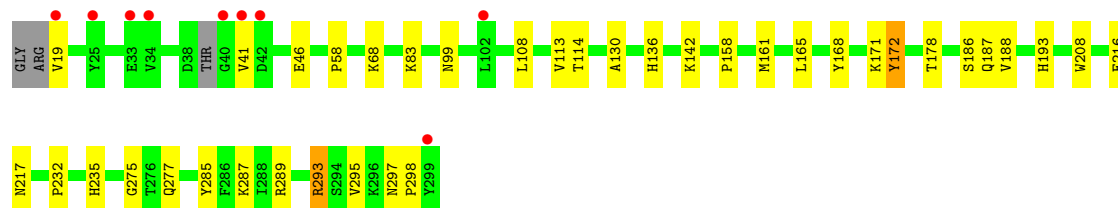
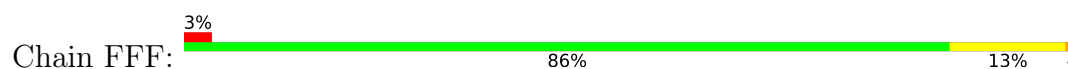




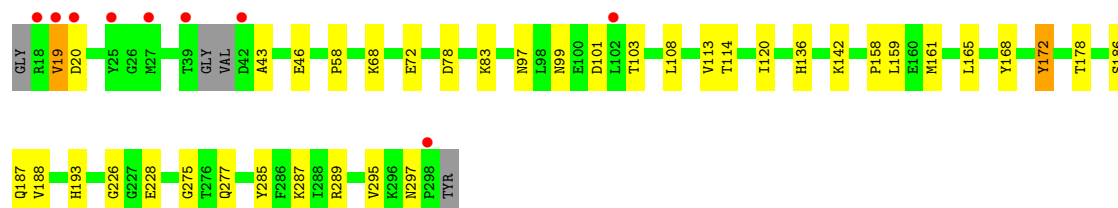
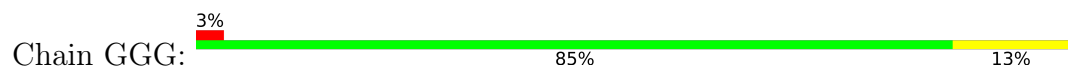
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



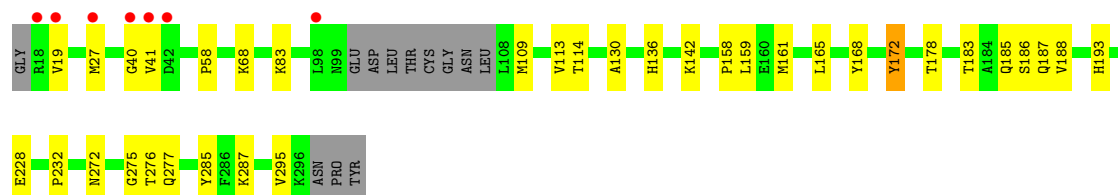
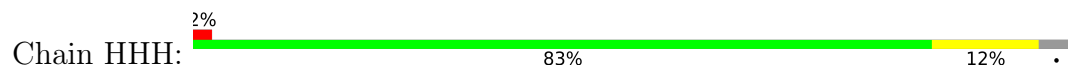
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



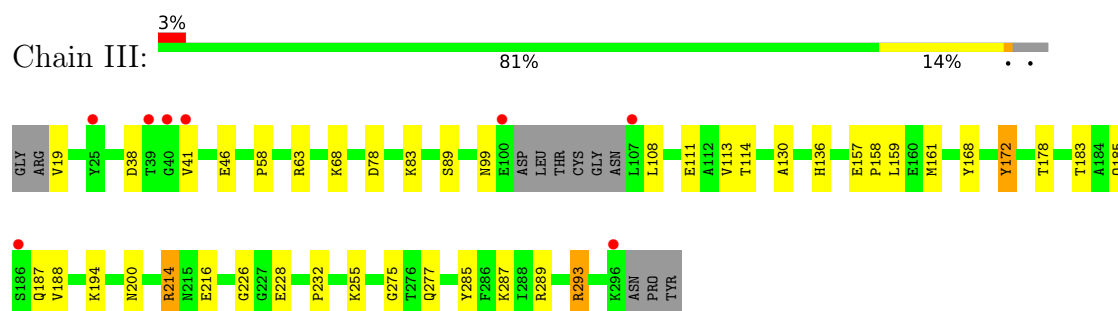
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



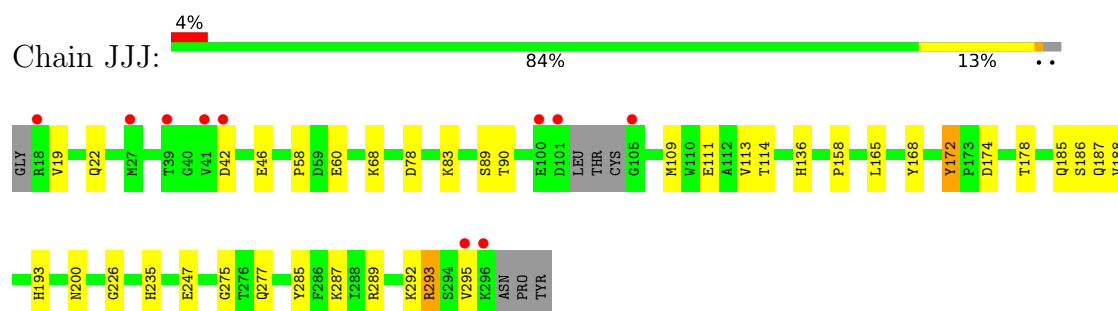
- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



- Molecule 1: Major capsid protein VP1, Major capsid protein VP1



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	234.30Å 97.20Å 146.17Å 90.00° 98.38° 90.00°	Depositor
Resolution (Å)	90.00 – 2.48 90.00 – 2.48	Depositor EDS
% Data completeness (in resolution range)	99.1 (90.00-2.48) 99.1 (90.00-2.48)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.86 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.205 , 0.233 (Not available) , 0.237	Depositor DCC
R_{free} test set	5586 reflections (4.87%)	wwPDB-VP
Wilson B-factor (Å ²)	39.2	Xtriage
Anisotropy	0.394	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 41.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	22259	wwPDB-VP
Average B, all atoms (Å ²)	43.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.31% 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: DMS, CA

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	AAA	1.01	0/2203	1.29	7/2994 (0.2%)
1	BBB	1.00	0/2140	1.24	0/2907
1	CCC	1.05	2/2205 (0.1%)	1.27	4/2995 (0.1%)
1	DDD	1.05	1/2204 (0.0%)	1.32	7/2997 (0.2%)
1	EEE	1.04	0/2202	1.27	3/2992 (0.1%)
1	FFF	1.04	2/2224 (0.1%)	1.28	6/3019 (0.2%)
1	GGG	1.02	1/2213 (0.0%)	1.27	3/3005 (0.1%)
1	HHH	1.02	0/2150	1.25	2/2919 (0.1%)
1	III	1.02	0/2156	1.26	4/2928 (0.1%)
1	JJJ	1.03	1/2174 (0.0%)	1.29	6/2953 (0.2%)
All	All	1.03	7/21871 (0.0%)	1.28	42/29709 (0.1%)

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	HHH	0	1

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	CCC	254	CYS	C-O	7.62	1.34	1.23
1	GGG	19	VAL	C-O	6.91	1.31	1.24
1	FFF	297	ASN	C-N	6.43	1.38	1.33
1	DDD	254	CYS	C-O	5.92	1.31	1.23
1	CCC	235	HIS	CE1-NE2	5.48	1.38	1.32

The worst 5 of 42 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	DDD	38	ASP	CB-CA-C	10.83	126.60	111.23
1	DDD	38	ASP	CA-CB-CG	8.95	121.55	112.60
1	EEE	214	ARG	CB-CG-CD	8.88	131.74	111.30
1	AAA	216	GLU	CB-CG-CD	8.55	127.13	112.60
1	III	216	GLU	CB-CG-CD	8.53	127.11	112.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	HHH	40	GLY	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	AAA	2154	0	2072	25	0
1	BBB	2092	0	2017	22	0
1	CCC	2157	0	2068	30	0
1	DDD	2155	0	2063	43	0
1	EEE	2153	0	2075	26	0
1	FFF	2174	0	2093	29	0
1	GGG	2164	0	2088	31	0
1	HHH	2102	0	2027	31	0
1	III	2108	0	2031	31	0
1	JJJ	2126	0	2040	27	0
2	CCC	4	0	6	0	0
2	HHH	4	0	6	0	0
3	DDD	1	0	0	0	0
3	FFF	1	0	0	0	0
3	III	1	0	0	0	0
4	AAA	93	0	0	1	0
4	BBB	89	0	0	0	0
4	CCC	92	0	0	2	0
4	DDD	88	0	0	1	0
4	EEE	84	0	0	3	0
4	FFF	98	0	0	3	0
4	GGG	80	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	HHH	64	0	0	1	0
4	III	84	0	0	0	0
4	JJJ	91	0	0	2	0
All	All	22259	0	20586	247	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 6.

The worst 5 of 247 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:CCC:109:MET:CE	1:CCC:255:LYS:HA	1.79	1.10
1:BBB:41:VAL:HG12	1:HHH:295:VAL:HG12	1.39	1.03
1:CCC:109:MET:HE3	1:CCC:255:LYS:HA	1.53	0.88
1:III:157:GLU:OE2	1:III:255:LYS:HE3	1.73	0.87
1:HHH:68:LYS:HB2	1:HHH:276:THR:HG23	1.55	0.85

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	CCC	14/283 (5%)	14 (100%)	0	0	100	100
1	DDD	278/283 (98%)	265 (95%)	12 (4%)	1 (0%)	30	46
1	EEE	277/283 (98%)	266 (96%)	10 (4%)	1 (0%)	30	46
1	FFF	276/283 (98%)	268 (97%)	7 (2%)	1 (0%)	30	46
1	GGG	275/283 (97%)	265 (96%)	8 (3%)	2 (1%)	18	32
1	HHH	267/283 (94%)	255 (96%)	11 (4%)	1 (0%)	30	46
1	III	268/283 (95%)	257 (96%)	10 (4%)	1 (0%)	30	46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	JJJ	272/283 (96%)	262 (96%)	9 (3%)	1 (0%)	30	46
All	All	1927/2264 (85%)	1852 (96%)	67 (4%)	8 (0%)	30	46

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	GGG	43	ALA
1	EEE	188	VAL
1	GGG	188	VAL
1	HHH	188	VAL
1	III	188	VAL

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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 oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 3 are monoatomic - leaving 2 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	$\# Z > 2$
2	DMS	HHH	301	-	3,3,3	0.42	0	3,3,3	0.36	0
2	DMS	CCC	301	-	3,3,3	0.51	0	3,3,3	0.32	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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 ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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

Mol	Chain	Analysed	<RSRZ>	#RSRZ > 2			OWAB(Å ²)	Q < 0.9
1	AAA	279/283 (98%)	-0.04	8 (2%)	53	50	28, 39, 72, 101	0
1	BBB	271/283 (95%)	-0.01	6 (2%)	62	59	25, 37, 64, 91	0
1	CCC	279/283 (98%)	-0.06	6 (2%)	62	59	27, 37, 62, 88	0
1	DDD	280/283 (98%)	0.02	7 (2%)	58	55	29, 39, 64, 95	0
1	EEE	279/283 (98%)	0.02	6 (2%)	62	59	29, 40, 75, 101	0
1	FFF	280/283 (98%)	0.07	9 (3%)	50	47	28, 39, 72, 89	0
1	GGG	279/283 (98%)	0.13	9 (3%)	50	47	32, 42, 72, 96	0
1	HHH	271/283 (95%)	0.25	7 (2%)	57	53	32, 45, 74, 99	0
1	III	272/283 (96%)	0.09	8 (2%)	53	50	32, 44, 68, 96	0
1	JJJ	276/283 (97%)	0.07	10 (3%)	46	42	28, 39, 68, 89	0
All	All	2766/2830 (97%)	0.05	76 (2%)	56	52	25, 40, 70, 101	0

The worst 5 of 76 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	JJJ	105	GLY	5.6
1	JJJ	101	ASP	4.8
1	CCC	40	GLY	4.4
1	EEE	101	ASP	4.4
1	GGG	39	THR	4.3

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

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

6.3 Carbohydrates [i](#)

There are no oligosaccharides 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($\text{\AA}^2$)	Q<0.9
2	DMS	HHH	301	4/4	0.79	0.27	72,74,74,75	0
3	CA	DDD	301	1/1	0.86	0.08	82,82,82,82	0
2	DMS	CCC	301	4/4	0.88	0.26	67,74,78,80	0
3	CA	FFF	301	1/1	0.90	0.08	90,90,90,90	0
3	CA	III	301	1/1	0.94	0.12	91,91,91,91	0

6.5 Other polymers [i](#)

There are no such residues in this entry.