



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 04:45 AM UTC

PDB ID : 7Z4O / pdb_00007z4o
Title : Influenza A/H7N9 polymerase core dimer with Pol II pSer5 CTD peptide mimic bound in site 2A
Authors : Cusack, S.; Pflug, A.
Deposited on : 2022-03-04
Resolution : 3.41 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:

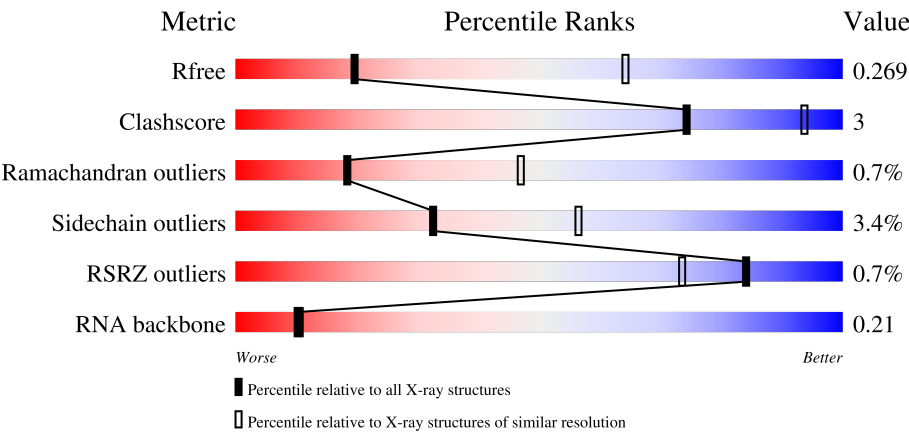
MolProbity : 4-5-2 with Phenix2.0
Mogul : 2022.3.0, CSD as543be (2022)
Xtriage (Phenix) : 2.0
EDS : 3.0
Percentile statistics : 20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4 : 9.0.010 (Gargrove)
Density-Fitness : 1.0.12
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (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 3.41 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	1210 (3.48-3.36)
Clashscore	190562	1234 (3.48-3.36)
Ramachandran outliers	187476	1222 (3.48-3.36)
Sidechain outliers	187428	1222 (3.48-3.36)
RSRZ outliers	180081	1210 (3.48-3.36)
RNA backbone	3983	1162 (3.84-3.00)

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	517	89%10%.
1	DDD	517	% 88%10%..
2	BBB	757	% 72%9%.18%
2	EEE	757	72%8%.19%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	CCC	147	43% 54%
3	FFF	147	45% 52%
4	JJJ	28	32% 64%
4	KKK	28	29% 11% 61%
5	UUU	12	42% 58%
5	VVV	12	42% 42% 17%

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 19969 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 Polymerase acidic protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	AAA	514	Total	C	N	O	S	0	0	0
			4133	2617	702	784	30			
1	DDD	514	Total	C	N	O	S	0	0	0
			4133	2617	702	784	30			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	200	MET	-	initiating methionine	UNP M9TI86
DDD	200	MET	-	initiating methionine	UNP M9TI86

- Molecule 2 is a protein called RNA-directed RNA polymerase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	BBB	622	Total	C	N	O	S	0	8	0
			5010	3162	866	944	38			
2	EEE	615	Total	C	N	O	S	0	0	0
			4892	3088	847	921	36			

- Molecule 3 is a protein called Polymerase basic protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	CCC	68	Total	C	N	O	S	0	0	0
			545	341	99	98	7			
3	FFF	71	Total	C	N	O	S	0	0	0
			564	353	102	102	7			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	129	SER	-	expression tag	UNP X5F427
CCC	130	GLY	-	expression tag	UNP X5F427

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	131	SER	-	expression tag	UNP X5F427
CCC	132	GLU	-	expression tag	UNP X5F427
CCC	133	ASN	-	expression tag	UNP X5F427
CCC	134	LEU	-	expression tag	UNP X5F427
CCC	135	TYR	-	expression tag	UNP X5F427
CCC	136	PHE	-	expression tag	UNP X5F427
CCC	137	GLN	-	expression tag	UNP X5F427
CCC	138	GLY	-	expression tag	UNP X5F427
CCC	139	SER	-	expression tag	UNP X5F427
CCC	140	HIS	-	expression tag	UNP X5F427
CCC	141	HIS	-	expression tag	UNP X5F427
CCC	142	HIS	-	expression tag	UNP X5F427
CCC	143	HIS	-	expression tag	UNP X5F427
CCC	144	HIS	-	expression tag	UNP X5F427
CCC	145	HIS	-	expression tag	UNP X5F427
CCC	146	HIS	-	expression tag	UNP X5F427
CCC	147	HIS	-	expression tag	UNP X5F427
FFF	129	SER	-	expression tag	UNP X5F427
FFF	130	GLY	-	expression tag	UNP X5F427
FFF	131	SER	-	expression tag	UNP X5F427
FFF	132	GLU	-	expression tag	UNP X5F427
FFF	133	ASN	-	expression tag	UNP X5F427
FFF	134	LEU	-	expression tag	UNP X5F427
FFF	135	TYR	-	expression tag	UNP X5F427
FFF	136	PHE	-	expression tag	UNP X5F427
FFF	137	GLN	-	expression tag	UNP X5F427
FFF	138	GLY	-	expression tag	UNP X5F427
FFF	139	SER	-	expression tag	UNP X5F427
FFF	140	HIS	-	expression tag	UNP X5F427
FFF	141	HIS	-	expression tag	UNP X5F427
FFF	142	HIS	-	expression tag	UNP X5F427
FFF	143	HIS	-	expression tag	UNP X5F427
FFF	144	HIS	-	expression tag	UNP X5F427
FFF	145	HIS	-	expression tag	UNP X5F427
FFF	146	HIS	-	expression tag	UNP X5F427
FFF	147	HIS	-	expression tag	UNP X5F427

- Molecule 4 is a protein called SER-TYR-SER-PRO-THR-SEP-PRO-SER-TYR-SER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	JJJ	10	Total	C	N	O	P	0	0	0
			79	47	10	21	1			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	KKK	11	Total	C	N	O	P	0	0	0
			87	53	11	22	1			

- Molecule 5 is a RNA chain called RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	UUU	12	Total	C	N	O	P	0	0	0
			262	117	52	81	12			
5	VVV	12	Total	C	N	O	P	0	0	0
			262	117	52	81	12			

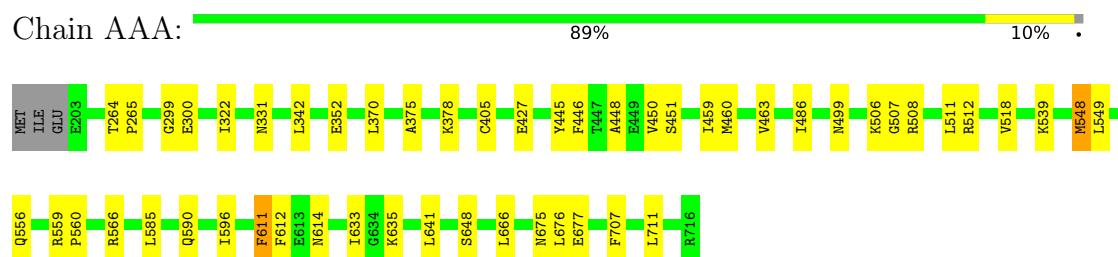
- Molecule 6 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	BBB	1	Total	Mg	0	0
			1	1		
6	EEE	1	Total	Mg	0	0
			1	1		

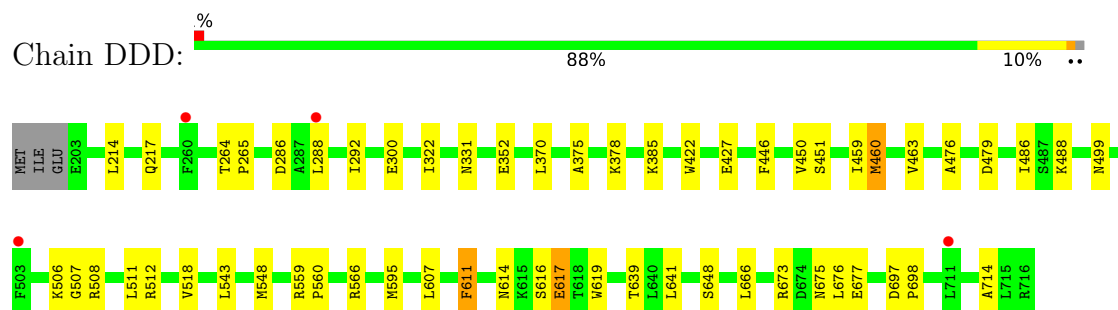
3 Residue-property plots

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.

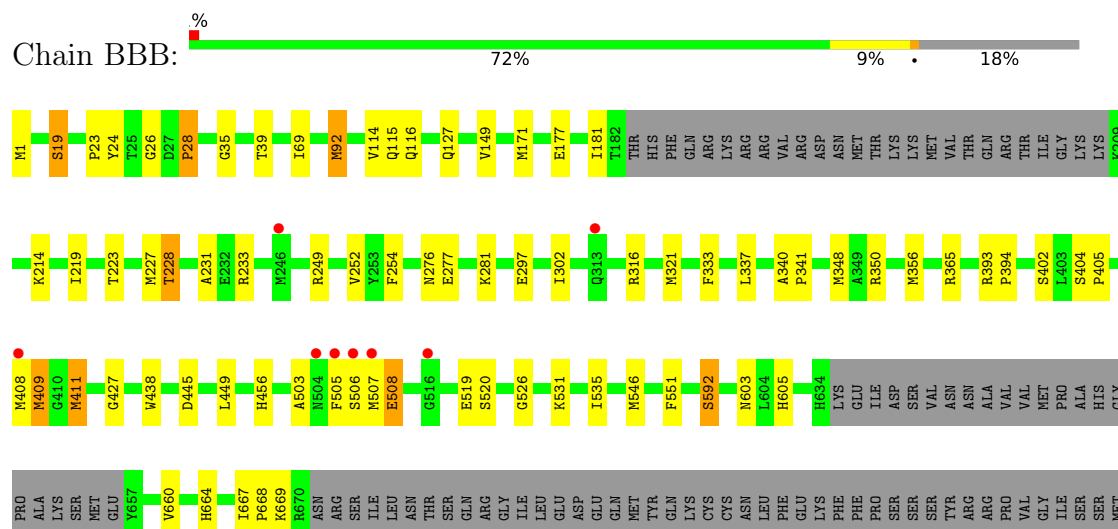
• Molecule 1: Polymerase acidic protein



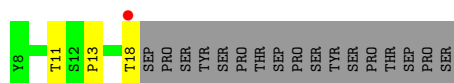
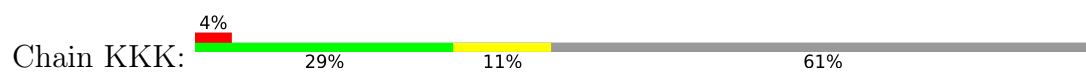
• Molecule 1: Polymerase acidic protein



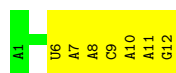
• Molecule 2: RNA-directed RNA polymerase catalytic subunit



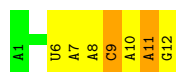




- Molecule 5: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*G)-3')



- Molecule 5: RNA (5'-R(P*AP*GP*UP*AP*GP*UP*AP*AP*CP*AP*AP*G)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	76.47Å 144.13Å 336.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.21 – 3.41 49.21 – 3.41	Depositor EDS
% Data completeness (in resolution range)	73.2 (49.21-3.41) 73.2 (49.21-3.41)	Depositor EDS
R_{merge}	0.28	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.44 (at 3.40Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.220 , 0.273 0.221 , 0.269	Depositor DCC
R_{free} test set	1852 reflections (2.84%)	wwPDB-VP
Wilson B-factor (Å ²)	73.0	Xtriage
Anisotropy	0.100	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 64.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	19969	wwPDB-VP
Average B, all atoms (Å ²)	92.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.79% 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: SEP, MG

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	0/4222	1.47	2/5700 (0.0%)
1	DDD	1.02	0/4222	1.48	2/5700 (0.0%)
2	BBB	1.02	0/5111	1.46	0/6905
2	EEE	1.02	0/4993	1.48	0/6747
3	CCC	1.03	0/558	1.46	0/755
3	FFF	1.02	0/577	1.44	0/782
4	JJJ	0.90	0/71	1.00	0/95
4	KKK	0.93	0/80	1.02	0/109
5	UUU	0.80	0/294	1.08	1/455 (0.2%)
5	VVV	0.80	0/294	1.06	1/455 (0.2%)
All	All	1.01	0/20422	1.46	6/27703 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	UUU	9	C	P-O3'-C3'	-5.73	111.61	120.20
5	VVV	9	C	P-O3'-C3'	-5.43	112.05	120.20
1	AAA	300	GLU	CA-C-N	5.11	124.83	119.92
1	AAA	300	GLU	C-N-CA	5.11	124.83	119.92
1	DDD	300	GLU	CA-C-N	5.01	124.86	120.10

There are no chirality outliers.

There are no planarity outliers.

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	4133	0	4103	26	0
1	DDD	4133	0	4103	28	0
2	BBB	5010	0	4984	43	0
2	EEE	4892	0	4862	34	0
3	CCC	545	0	543	3	0
3	FFF	564	0	561	3	0
4	JJJ	79	0	61	0	0
4	KKK	87	0	70	2	0
5	UUU	262	0	131	1	0
5	VVV	262	0	131	3	0
6	BBB	1	0	0	0	0
6	EEE	1	0	0	0	0
All	All	19969	0	19549	121	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 3.

The worst 5 of 121 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:BBB:507[B]:MET:SD	2:BBB:508[B]:GLU:N	2.54	0.81
2:BBB:252:VAL:HG11	2:BBB:411:MET:HG2	1.64	0.78
1:DDD:607:LEU:O	1:DDD:611:PHE:HB2	1.84	0.76
2:BBB:503[B]:ALA:HB1	2:BBB:505[B]:PHE:CE2	2.23	0.73
2:BBB:411:MET:SD	2:BBB:411:MET:N	2.64	0.70

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	AAA	512/517 (99%)	477 (93%)	33 (6%)	2 (0%)	30	59
1	DDD	512/517 (99%)	479 (94%)	31 (6%)	2 (0%)	30	59
2	BBB	624/757 (82%)	571 (92%)	46 (7%)	7 (1%)	11	37
2	EEE	609/757 (80%)	560 (92%)	44 (7%)	5 (1%)	16	44
3	CCC	66/147 (45%)	65 (98%)	1 (2%)	0	100	100
3	FFF	69/147 (47%)	67 (97%)	1 (1%)	1 (1%)	9	31
4	JJJ	7/28 (25%)	4 (57%)	3 (43%)	0	100	100
4	KKK	8/28 (29%)	8 (100%)	0	0	100	100
All	All	2407/2898 (83%)	2231 (93%)	159 (7%)	17 (1%)	18	47

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	614	ASN
1	DDD	614	ASN
1	DDD	617	GLU
3	FFF	107	SER
2	BBB	19	SER

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	AAA	461/464 (99%)	448 (97%)	13 (3%)	38	59
1	DDD	461/464 (99%)	447 (97%)	14 (3%)	36	59
2	BBB	551/668 (82%)	531 (96%)	20 (4%)	31	55
2	EEE	538/668 (80%)	518 (96%)	20 (4%)	30	54
3	CCC	59/134 (44%)	58 (98%)	1 (2%)	53	67
3	FFF	61/134 (46%)	60 (98%)	1 (2%)	55	68
4	JJJ	9/24 (38%)	8 (89%)	1 (11%)	6	21
4	KKK	10/24 (42%)	8 (80%)	2 (20%)	1	5

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	2150/2580 (83%)	2078 (97%)	72 (3%)	32 57

5 of 72 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	EEE	411	MET
4	KKK	18	THR
2	EEE	456	HIS
2	EEE	660	VAL
2	BBB	350	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
5	UUU	11/12 (91%)	5 (45%)	0
5	VVV	11/12 (91%)	5 (45%)	0
All	All	22/24 (91%)	10 (45%)	0

5 of 10 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
5	UUU	6	U
5	UUU	7	A
5	UUU	8	A
5	UUU	11	A
5	UUU	12	G

There are no RNA pucker outliers to report.

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

2 non-standard protein/DNA/RNA residues are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the

expected value. A bond length (or angle) with $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	SEP	JJJ	12	4	8,9,10	0.60	0	7,12,14	0.71	0
4	SEP	KKK	12	4	8,9,10	0.61	0	7,12,14	0.66	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
4	SEP	JJJ	12	4	-	3/6/8/10	-
4	SEP	KKK	12	4	-	2/6/8/10	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	KKK	12	SEP	N-CA-CB-OG
4	KKK	12	SEP	C-CA-CB-OG
4	JJJ	12	SEP	N-CA-CB-OG
4	JJJ	12	SEP	C-CA-CB-OG
4	JJJ	12	SEP	CA-CB-OG-P

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 2 ligands modelled in this entry, 2 are monoatomic - leaving 0 for Mogul analysis.

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 [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	AAA	514/517 (99%)	-0.31	0 100 100	28, 58, 111, 149	0
1	DDD	514/517 (99%)	0.11	4 (0%) 82 71	59, 99, 130, 160	0
2	BBB	622/757 (82%)	-0.04	8 (1%) 75 61	22, 88, 133, 167	8 (1%)
2	EEE	615/757 (81%)	0.09	2 (0%) 90 84	47, 107, 146, 179	0
3	CCC	68/147 (46%)	0.31	0 100 100	92, 125, 174, 178	0
3	FFF	71/147 (48%)	0.08	2 (2%) 55 41	58, 93, 146, 161	0
4	JJJ	9/28 (32%)	0.24	0 100 100	88, 96, 121, 130	0
4	KKK	10/28 (35%)	-0.10	1 (10%) 12 12	63, 71, 115, 121	0
5	UUU	12/12 (100%)	0.07	0 100 100	112, 119, 131, 161	0
5	VVV	12/12 (100%)	-0.67	0 100 100	58, 78, 123, 152	0
All	All	2447/2922 (83%)	-0.02	17 (0%) 84 73	22, 92, 137, 179	8 (0%)

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	BBB	506[A]	SER	7.1
2	BBB	504[A]	ASN	5.1
2	EEE	236	LEU	3.2
2	BBB	507[A]	MET	3.0
2	EEE	46	THR	2.6

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($\text{\AA}^2$)	Q<0.9
4	SEP	JJJ	12	10/11	0.94	0.06	85,88,97,97	0
4	SEP	KKK	12	10/11	0.97	0.05	52,56,62,62	0

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
6	MG	BBB	801	1/1	0.97	0.10	19,19,19,19	0
6	MG	EEE	801	1/1	0.99	0.02	18,18,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.