



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2018 – 10:44 pm GMT

PDB ID : 3N4P
Title : Human cytomegalovirus terminase nuclease domain
Authors : Nadal, M.; Mas, P.J.; Blanco, A.G.; Arnan, C.; Sola, M.; Hart, D.J.; Coll, M.
Deposited on : 2010-05-22
Resolution : 2.15 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 following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467
Xtriage (Phenix) : 1.13
EDS : trunk30967
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk30967

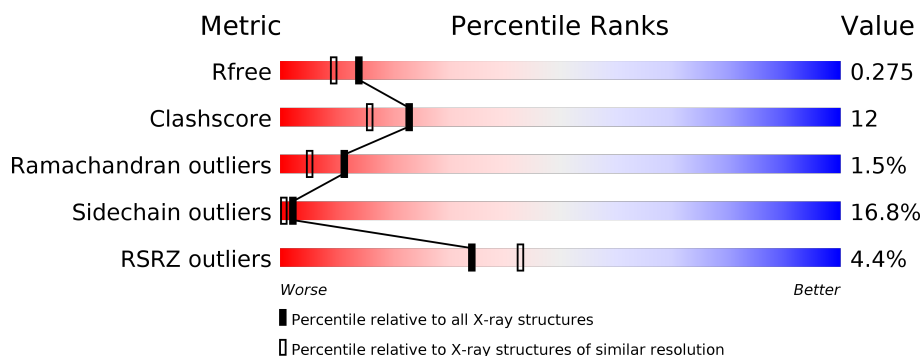
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.15 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}	111664	1287 (2.16-2.16)
Clashscore	122126	1390 (2.16-2.16)
Ramachandran outliers	120053	1368 (2.16-2.16)
Sidechain outliers	120020	1367 (2.16-2.16)
RSRZ outliers	108989	1262 (2.16-2.16)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 3 , 2, 1 and 0 types of geometric quality criteria. 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	279	<div> <div>4%</div> <div> <div>51%</div> <div>23%</div> <div>• •</div> <div>23%</div> </div> </div>
1	B	279	<div> <div>4%</div> <div> <div>53%</div> <div>20%</div> <div>8%</div> <div>19%</div> </div> </div>
1	C	279	<div> <div>2%</div> <div> <div>53%</div> <div>19%</div> <div>5%</div> <div>22%</div> </div> </div>
1	D	279	<div> <div>3%</div> <div> <div>49%</div> <div>20%</div> <div>6%</div> <div>•</div> <div>24%</div> </div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7146 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 Terminase subunit UL89 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	216	Total	C	N	O	S	0	0	0
			1736	1113	292	325	6			
1	B	227	Total	C	N	O	S	0	1	0
			1838	1176	315	341	6			
1	C	218	Total	C	N	O	S	0	0	0
			1747	1120	292	329	6			
1	D	212	Total	C	N	O	S	0	1	0
			1720	1103	290	321	6			

There are 88 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	396	MET	-	EXPRESSION TAG	UNP P16732
A	397	GLY	-	EXPRESSION TAG	UNP P16732
A	398	HIS	-	EXPRESSION TAG	UNP P16732
A	399	HIS	-	EXPRESSION TAG	UNP P16732
A	400	HIS	-	EXPRESSION TAG	UNP P16732
A	401	HIS	-	EXPRESSION TAG	UNP P16732
A	402	HIS	-	EXPRESSION TAG	UNP P16732
A	403	HIS	-	EXPRESSION TAG	UNP P16732
A	404	ASP	-	EXPRESSION TAG	UNP P16732
A	405	TYR	-	EXPRESSION TAG	UNP P16732
A	406	ASP	-	EXPRESSION TAG	UNP P16732
A	407	ILE	-	EXPRESSION TAG	UNP P16732
A	408	PRO	-	EXPRESSION TAG	UNP P16732
A	409	THR	-	EXPRESSION TAG	UNP P16732
A	410	THR	-	EXPRESSION TAG	UNP P16732
A	411	GLU	-	EXPRESSION TAG	UNP P16732
A	412	ASN	-	EXPRESSION TAG	UNP P16732
A	413	LEU	-	EXPRESSION TAG	UNP P16732
A	414	TYR	-	EXPRESSION TAG	UNP P16732
A	415	PHE	-	EXPRESSION TAG	UNP P16732
A	416	GLN	-	EXPRESSION TAG	UNP P16732

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	417	GLY	-	EXPRESSION TAG	UNP P16732
B	396	MET	-	EXPRESSION TAG	UNP P16732
B	397	GLY	-	EXPRESSION TAG	UNP P16732
B	398	HIS	-	EXPRESSION TAG	UNP P16732
B	399	HIS	-	EXPRESSION TAG	UNP P16732
B	400	HIS	-	EXPRESSION TAG	UNP P16732
B	401	HIS	-	EXPRESSION TAG	UNP P16732
B	402	HIS	-	EXPRESSION TAG	UNP P16732
B	403	HIS	-	EXPRESSION TAG	UNP P16732
B	404	ASP	-	EXPRESSION TAG	UNP P16732
B	405	TYR	-	EXPRESSION TAG	UNP P16732
B	406	ASP	-	EXPRESSION TAG	UNP P16732
B	407	ILE	-	EXPRESSION TAG	UNP P16732
B	408	PRO	-	EXPRESSION TAG	UNP P16732
B	409	THR	-	EXPRESSION TAG	UNP P16732
B	410	THR	-	EXPRESSION TAG	UNP P16732
B	411	GLU	-	EXPRESSION TAG	UNP P16732
B	412	ASN	-	EXPRESSION TAG	UNP P16732
B	413	LEU	-	EXPRESSION TAG	UNP P16732
B	414	TYR	-	EXPRESSION TAG	UNP P16732
B	415	PHE	-	EXPRESSION TAG	UNP P16732
B	416	GLN	-	EXPRESSION TAG	UNP P16732
B	417	GLY	-	EXPRESSION TAG	UNP P16732
C	396	MET	-	EXPRESSION TAG	UNP P16732
C	397	GLY	-	EXPRESSION TAG	UNP P16732
C	398	HIS	-	EXPRESSION TAG	UNP P16732
C	399	HIS	-	EXPRESSION TAG	UNP P16732
C	400	HIS	-	EXPRESSION TAG	UNP P16732
C	401	HIS	-	EXPRESSION TAG	UNP P16732
C	402	HIS	-	EXPRESSION TAG	UNP P16732
C	403	HIS	-	EXPRESSION TAG	UNP P16732
C	404	ASP	-	EXPRESSION TAG	UNP P16732
C	405	TYR	-	EXPRESSION TAG	UNP P16732
C	406	ASP	-	EXPRESSION TAG	UNP P16732
C	407	ILE	-	EXPRESSION TAG	UNP P16732
C	408	PRO	-	EXPRESSION TAG	UNP P16732
C	409	THR	-	EXPRESSION TAG	UNP P16732
C	410	THR	-	EXPRESSION TAG	UNP P16732
C	411	GLU	-	EXPRESSION TAG	UNP P16732
C	412	ASN	-	EXPRESSION TAG	UNP P16732
C	413	LEU	-	EXPRESSION TAG	UNP P16732
C	414	TYR	-	EXPRESSION TAG	UNP P16732

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	415	PHE	-	EXPRESSION TAG	UNP P16732
C	416	GLN	-	EXPRESSION TAG	UNP P16732
C	417	GLY	-	EXPRESSION TAG	UNP P16732
D	396	MET	-	EXPRESSION TAG	UNP P16732
D	397	GLY	-	EXPRESSION TAG	UNP P16732
D	398	HIS	-	EXPRESSION TAG	UNP P16732
D	399	HIS	-	EXPRESSION TAG	UNP P16732
D	400	HIS	-	EXPRESSION TAG	UNP P16732
D	401	HIS	-	EXPRESSION TAG	UNP P16732
D	402	HIS	-	EXPRESSION TAG	UNP P16732
D	403	HIS	-	EXPRESSION TAG	UNP P16732
D	404	ASP	-	EXPRESSION TAG	UNP P16732
D	405	TYR	-	EXPRESSION TAG	UNP P16732
D	406	ASP	-	EXPRESSION TAG	UNP P16732
D	407	ILE	-	EXPRESSION TAG	UNP P16732
D	408	PRO	-	EXPRESSION TAG	UNP P16732
D	409	THR	-	EXPRESSION TAG	UNP P16732
D	410	THR	-	EXPRESSION TAG	UNP P16732
D	411	GLU	-	EXPRESSION TAG	UNP P16732
D	412	ASN	-	EXPRESSION TAG	UNP P16732
D	413	LEU	-	EXPRESSION TAG	UNP P16732
D	414	TYR	-	EXPRESSION TAG	UNP P16732
D	415	PHE	-	EXPRESSION TAG	UNP P16732
D	416	GLN	-	EXPRESSION TAG	UNP P16732
D	417	GLY	-	EXPRESSION TAG	UNP P16732

- Molecule 2 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	2	Total Mg 2 2	0	0
2	A	1	Total Mg 1 1	0	0
2	D	2	Total Mg 2 2	0	0
2	C	2	Total Mg 2 2	0	0

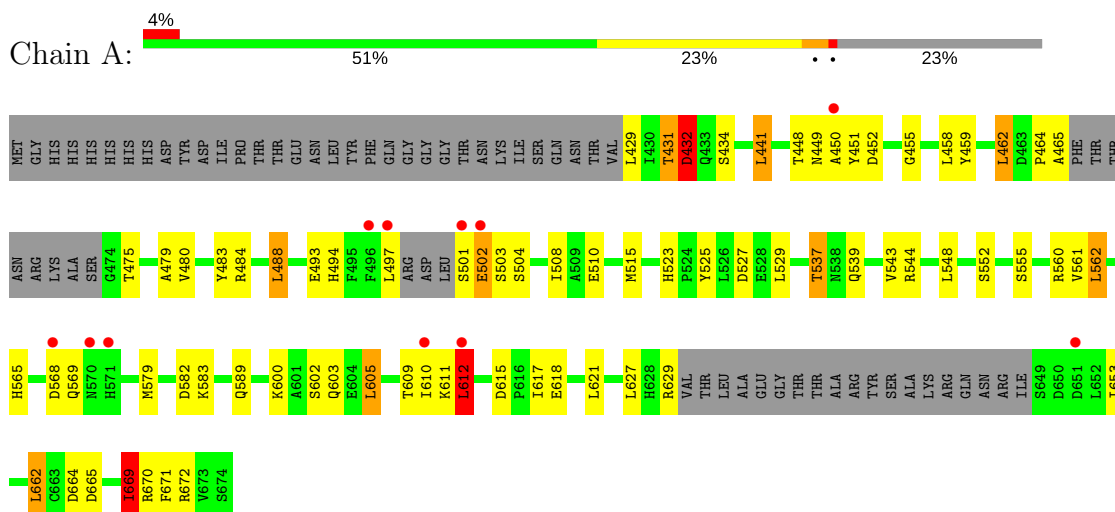
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	35	Total 35	O 35	0	0
3	B	25	Total 25	O 25	0	0
3	C	27	Total 27	O 27	0	0
3	D	11	Total 11	O 11	0	0

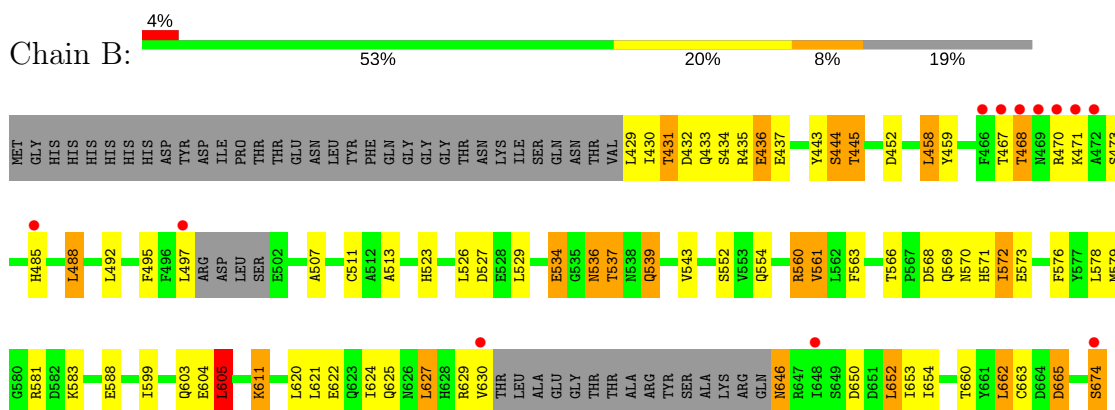
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA 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.

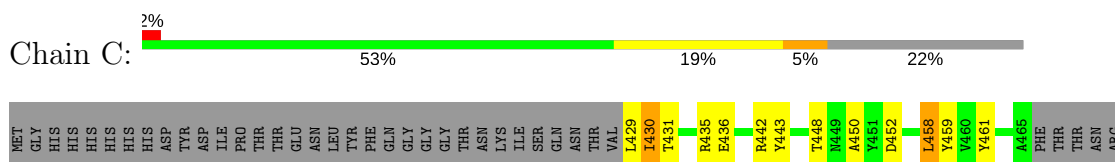
• Molecule 1: Terminase subunit UL89 protein

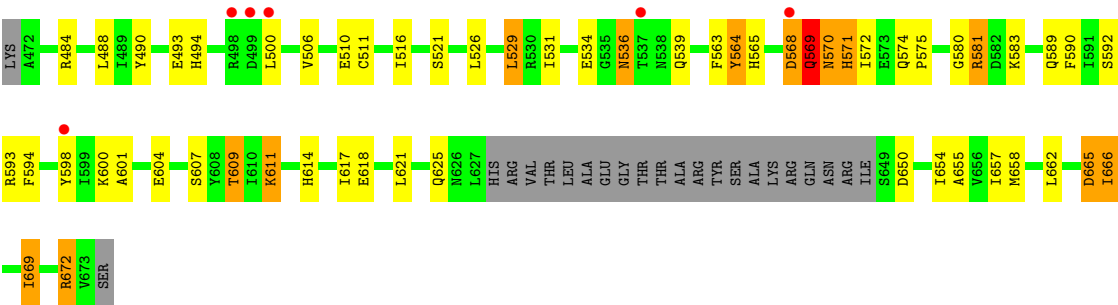


• Molecule 1: Terminase subunit UL89 protein

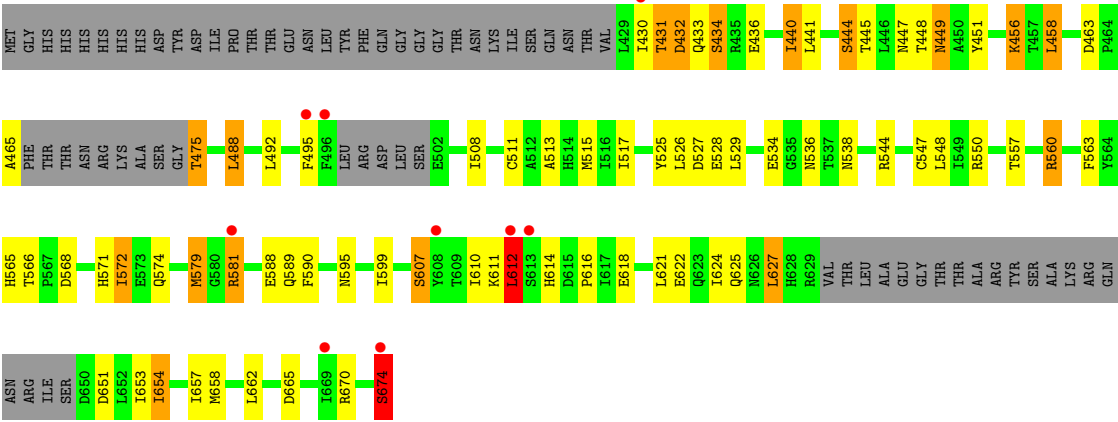


• Molecule 1: Terminase subunit UL89 protein





• Molecule 1: Terminase subunit UL89 protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	82.86Å 87.96Å 188.38Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.99 – 2.15 19.99 – 2.15	Depositor EDS
% Data completeness (in resolution range)	98.7 (19.99-2.15) 98.7 (19.99-2.15)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.53 (at 2.15Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.219 , 0.272 0.230 , 0.275	Depositor DCC
R_{free} test set	3760 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	43.7	Xtriage
Anisotropy	0.232	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 29.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7146	wwPDB-VP
Average B, all atoms (Å ²)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: 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	A	1.20	5/1770 (0.3%)	1.16	12/2399 (0.5%)
1	B	1.03	5/1875 (0.3%)	1.00	4/2542 (0.2%)
1	C	1.14	5/1781 (0.3%)	1.05	2/2416 (0.1%)
1	D	0.97	1/1754 (0.1%)	1.00	7/2378 (0.3%)
All	All	1.09	16/7180 (0.2%)	1.05	25/9735 (0.3%)

The worst 5 of 16 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	490	TYR	CD2-CE2	6.67	1.49	1.39
1	A	671	PHE	CE2-CZ	6.45	1.49	1.37
1	A	459	TYR	CD1-CE1	6.36	1.48	1.39
1	B	443	TYR	CD1-CE1	6.03	1.48	1.39
1	C	604	GLU	CB-CG	-5.78	1.41	1.52

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	612	LEU	CA-CB-CG	7.74	133.10	115.30
1	A	605	LEU	CA-CB-CG	7.35	132.20	115.30
1	B	662	LEU	CA-CB-CG	7.12	131.66	115.30
1	A	560	ARG	NE-CZ-NH2	-6.58	117.01	120.30
1	A	662	LEU	CB-CG-CD1	6.53	122.11	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	1736	0	1707	34	0
1	B	1838	0	1813	51	0
1	C	1747	0	1721	51	0
1	D	1720	0	1688	45	0
2	A	1	0	0	0	0
2	B	2	0	0	0	0
2	C	2	0	0	0	0
2	D	2	0	0	0	0
3	A	35	0	0	2	0
3	B	25	0	0	1	0
3	C	27	0	0	1	0
3	D	11	0	0	0	0
All	All	7146	0	6929	171	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 12.

The worst 5 of 171 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:D:550:ARG:HD3	1:D:674:SER:HB2	1.14	1.14
1:C:430:ILE:HD11	1:C:435:ARG:HG3	1.25	1.09
1:C:669:ILE:H	1:C:669:ILE:HD12	1.18	1.07
1:A:494:HIS:H	1:A:609:THR:HG21	1.24	1.03
1:C:430:ILE:HD11	1:C:435:ARG:CG	1.91	1.01

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone

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	208/279 (75%)	196 (94%)	9 (4%)	3 (1%)	12	6
1	B	222/279 (80%)	206 (93%)	13 (6%)	3 (1%)	12	6
1	C	212/279 (76%)	204 (96%)	3 (1%)	5 (2%)	6	1
1	D	205/279 (74%)	190 (93%)	13 (6%)	2 (1%)	17	10
All	All	847/1116 (76%)	796 (94%)	38 (4%)	13 (2%)	11	5

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	569	GLN
1	A	612	LEU
1	B	467	THR
1	C	571	HIS
1	C	569	GLN

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	188/242 (78%)	156 (83%)	32 (17%)	2	0
1	B	199/242 (82%)	165 (83%)	34 (17%)	2	0
1	C	189/242 (78%)	166 (88%)	23 (12%)	5	2
1	D	186/242 (77%)	147 (79%)	39 (21%)	1	0
All	All	762/968 (79%)	634 (83%)	128 (17%)	2	1

5 of 128 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	611	LYS
1	C	488	LEU
1	D	621	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	625	GLN
1	B	674	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	565	HIS
1	B	574	GLN
1	D	449	ASN
1	B	539	GLN
1	B	554	GLN

5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 7 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	A	216/279 (77%)	0.54	11 (5%) 28 35	28, 53, 94, 132	2 (0%)
1	B	227/279 (81%)	0.64	12 (5%) 26 34	36, 62, 112, 137	2 (0%)
1	C	218/279 (78%)	0.52	6 (2%) 53 61	31, 63, 101, 138	1 (0%)
1	D	212/279 (75%)	0.64	9 (4%) 36 45	39, 73, 108, 139	6 (2%)
All	All	873/1116 (78%)	0.59	38 (4%) 34 43	28, 62, 104, 139	11 (1%)

The worst 5 of 38 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	497	LEU	12.9
1	C	500	LEU	8.9
1	D	612	LEU	8.2
1	B	469	ASN	6.7
1	B	471	LYS	6.0

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 carbohydrates 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(\AA^2)	Q<0.9
2	MG	B	1	1/1	0.41	0.31	151,151,151,151	0
2	MG	C	2	1/1	0.45	0.35	130,130,130,130	0
2	MG	A	4	1/1	0.54	0.39	100,100,100,100	0
2	MG	C	7	1/1	0.55	0.23	102,102,102,102	0
2	MG	B	5	1/1	0.77	0.21	96,96,96,96	0
2	MG	D	6	1/1	0.81	0.40	130,130,130,130	0
2	MG	D	3	1/1	0.82	0.24	122,122,122,122	0

6.5 Other polymers [i](#)

There are no such residues in this entry.