



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 4, 2026 – 08:51 PM UTC

PDB ID : 1N6Q / pdb_00001n6q
Title : HIV-1 Reverse Transcriptase Crosslinked to pre-translocation AZTMP-terminated DNA (complex N)
Authors : Sarafianos, S.G.; Clark Jr., A.D.; Das, K.; Tuske, S.; Birktoft, J.J.; Ilankumaran, I.; Ramesha, A.R.; Sayer, J.M.; Jerina, D.M.; Boyer, P.L.; Hughes, S.H.; Arnold, E.
Deposited on : 2002-11-11
Resolution : 3.00 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

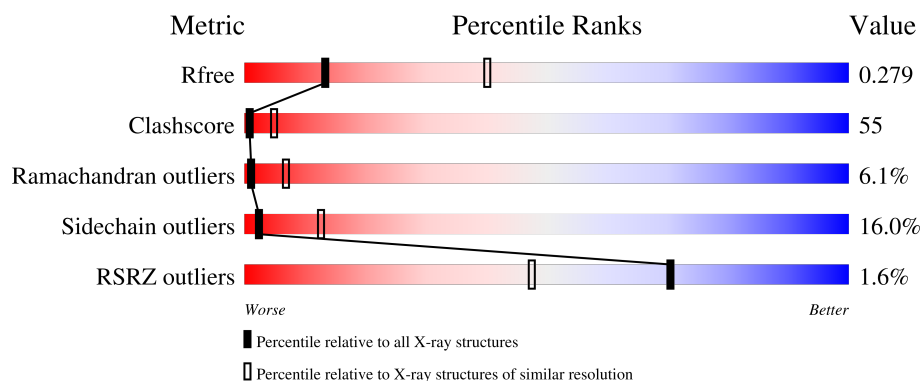
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.00 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	2672 (3.00-3.00)
Clashscore	190562	2977 (3.00-3.00)
Ramachandran outliers	187476	2877 (3.00-3.00)
Sidechain outliers	187428	2880 (3.00-3.00)
RSRZ outliers	180081	2671 (3.00-3.00)

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	T	27	
2	P	22	
3	A	558	
4	B	430	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
5	L	211	
6	H	225	

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
2	ATM	P	823	-	-	X	-

2 Entry composition

There are 8 unique types of molecules in this entry. The entry contains 12250 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 DNA chain called 5'-D(*AP*T*GP*CP*AP*TP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*GP*TP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	T	23	Total	C	N	O	P	0	0	0
			473	223	95	133	22			

- Molecule 2 is a DNA chain called 5'-D(*A*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*CP*GP*CP*CP*AP*(ATM))-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	P	21	Total	C	N	O	P	S	0	0	0
			429	205	77	126	20	1			

- Molecule 3 is a protein called Reverse Transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	558	Total	C	N	O	S	15	0	0
			4482	2901	741	832	8			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	258	CYS	GLN	engineered mutation	UNP P03366
A	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 4 is a protein called Reverse Transcriptase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	429	Total	C	N	O	S	12	0	0
			3534	2304	586	637	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366

- Molecule 5 is a protein called Monoclonal Antibody (Light Chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	L	211	Total	C	N	O	S	0	0	0
			1643	1025	270	342	6			

- Molecule 6 is a protein called Monoclonal Antibody (Heavy Chain).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	225	Total	C	N	O	S	0	0	0
			1685	1060	276	340	9			

- Molecule 7 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	Mg	0	0
			2	2		

- Molecule 8 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
8	A	2	Total	O	0	0
			2	2		

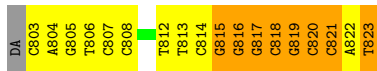
3 Residue-property plots

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.

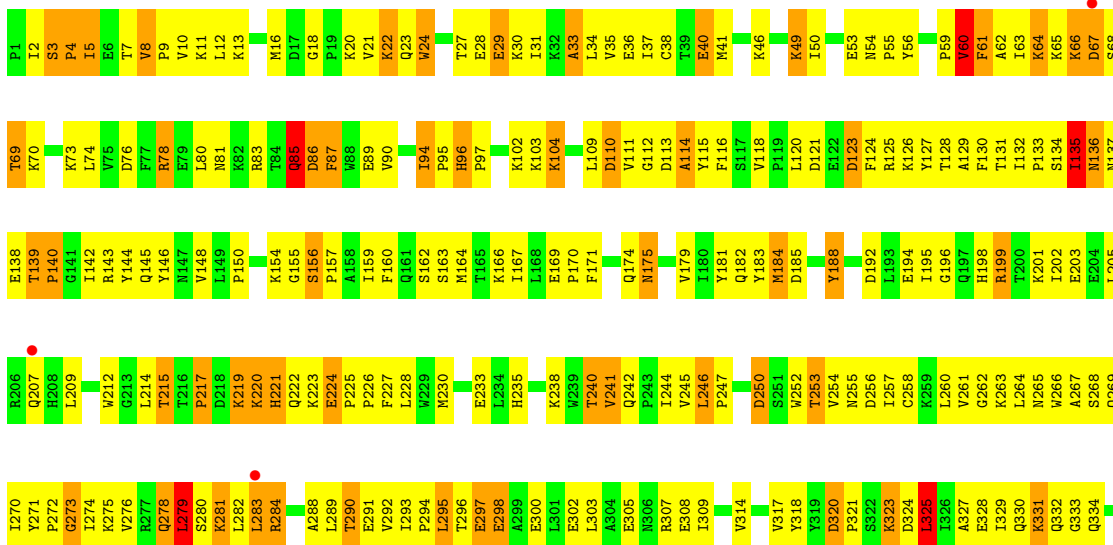
- Molecule 1: 5'-D(*AP*T*GP*CP*AP*TP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*GP*TP*G)-3'

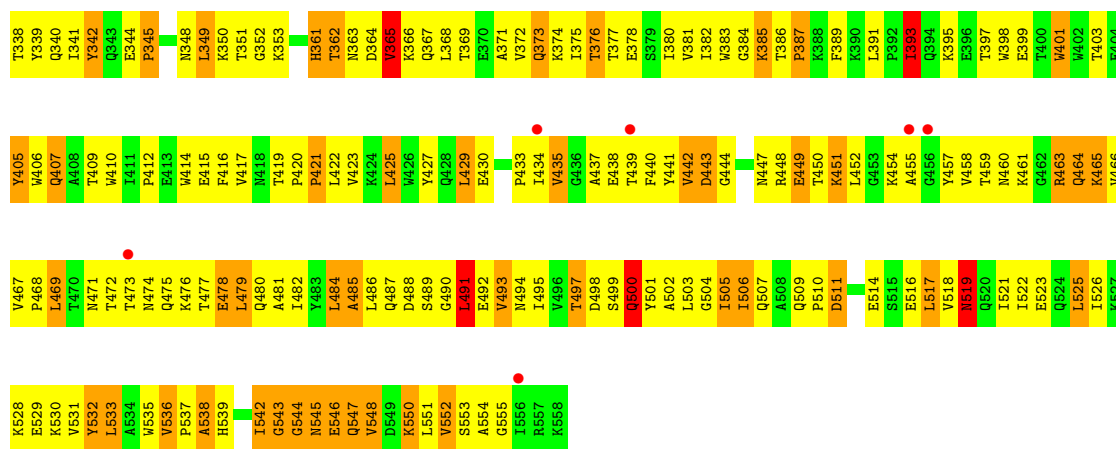


- Molecule 2: 5'-D(*A*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*C P*GP*CP*CP*AP*(ATM))-3'

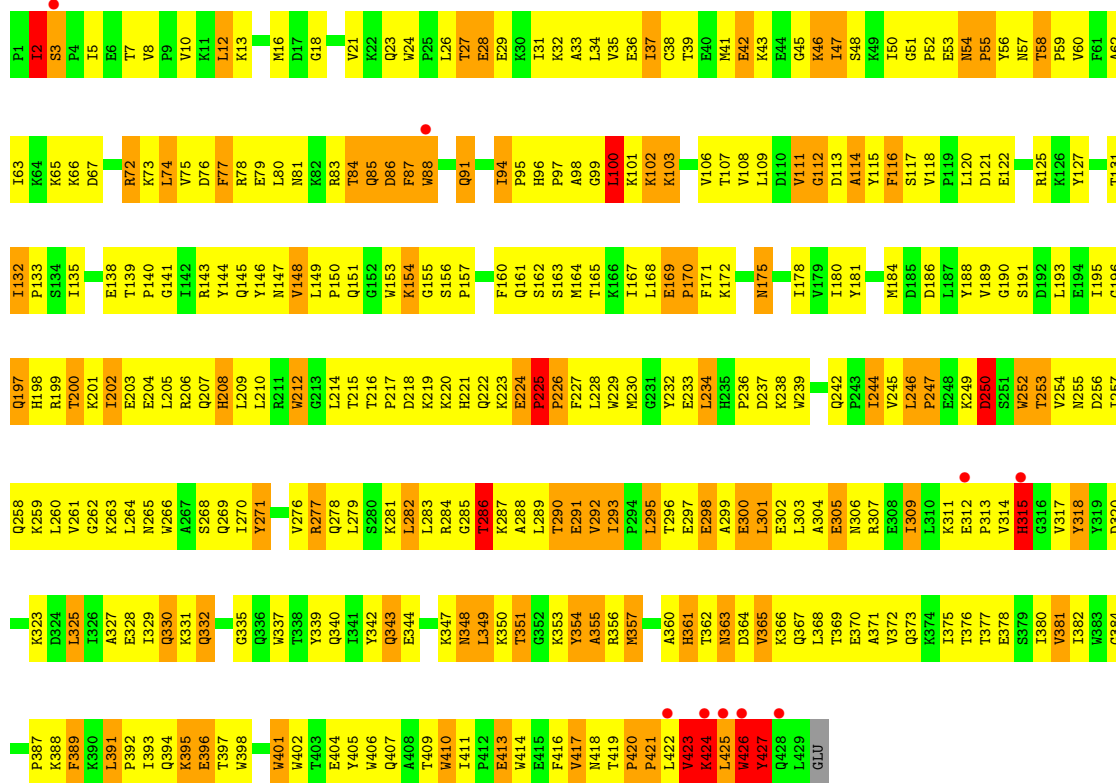


- Molecule 3: Reverse Transcriptase

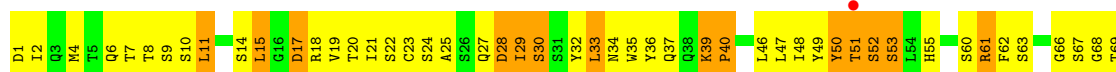


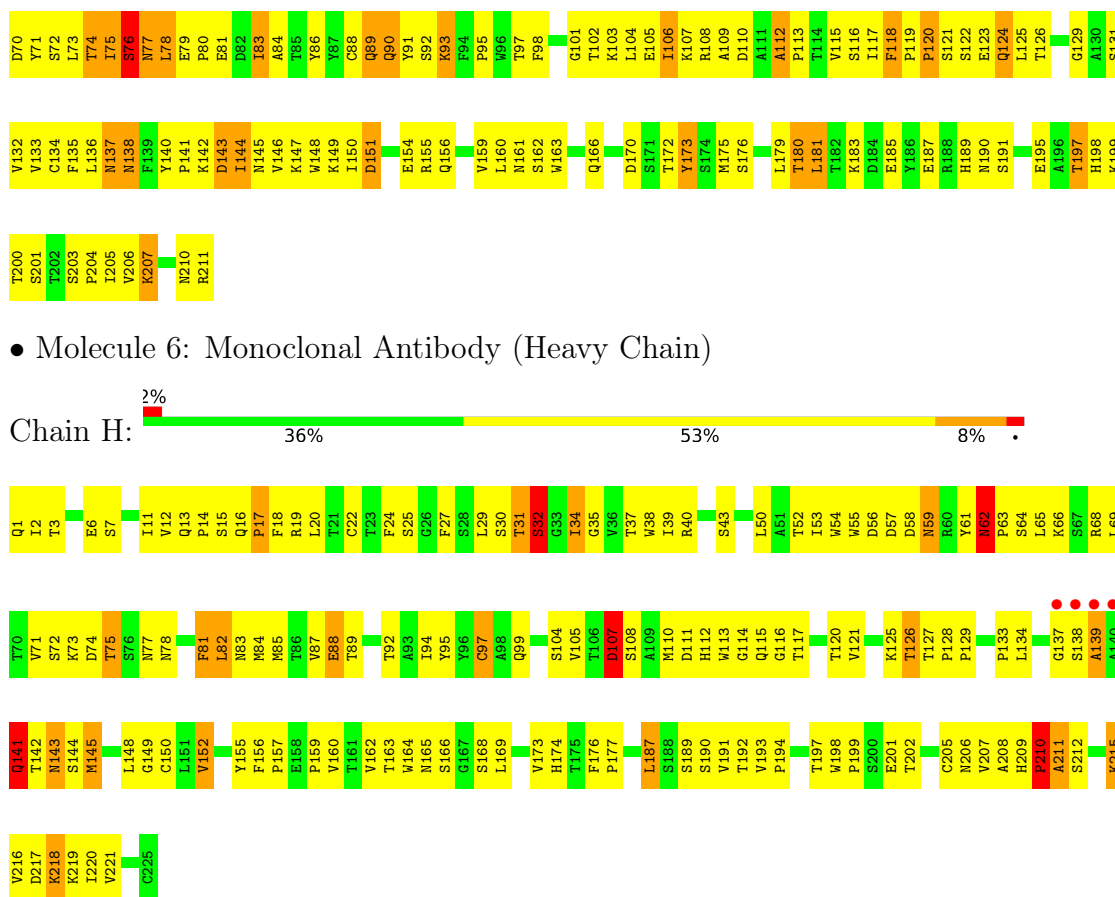


• Molecule 4: Reverse Transcriptase



• Molecule 5: Monoclonal Antibody (Light Chain)





● Molecule 6: Monoclonal Antibody (Heavy Chain)



4 Data and refinement statistics

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α , β , γ	166.35Å 166.35Å 220.96Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 20.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	85.0 (20.00-3.00) 84.6 (20.00-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.10	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.32 (at 3.01Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.247 , 0.284 0.246 , 0.279	Depositor DCC
R_{free} test set	2631 reflections (4.06%)	wwPDB-VP
Wilson B-factor (Å ²)	67.1	Xtriage
Anisotropy	0.053	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 49.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.45$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.052 for -h,-k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12250	wwPDB-VP
Average B, all atoms (Å ²)	73.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.92% 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: MRG, ATM, 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	T	0.45	0/532	1.14	11/820 (1.3%)
2	P	0.41	0/424	1.23	8/649 (1.2%)
3	A	0.56	1/4600 (0.0%)	1.09	30/6259 (0.5%)
4	B	0.73	5/3639 (0.1%)	1.27	56/4949 (1.1%)
5	L	0.62	2/1681 (0.1%)	1.12	15/2283 (0.7%)
6	H	0.64	0/1729	1.18	14/2372 (0.6%)
All	All	0.63	8/12605 (0.1%)	1.17	134/17332 (0.8%)

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
4	B	0	1

The worst 5 of 8 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	425	LEU	C-N	-12.43	1.16	1.33
4	B	426	TRP	N-CA	-9.91	1.33	1.46
5	L	50	TYR	C-N	-8.06	1.22	1.33
3	A	421	PRO	C-N	-8.02	1.22	1.33
4	B	3	SER	C-N	6.11	1.41	1.34

The worst 5 of 134 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	365	VAL	N-CA-C	-13.73	98.24	110.74
4	B	225	PRO	N-CA-C	13.07	126.65	110.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	224	GLU	CA-C-N	12.19	132.93	120.38
4	B	224	GLU	C-N-CA	12.19	132.93	120.38
2	P	815	DG	N9-C1'-C2'	12.03	131.55	113.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	B	427	TYR	Sidechain

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	T	473	0	257	34	0
2	P	429	0	242	28	0
3	A	4482	0	4484	534	0
4	B	3534	0	3567	423	0
5	L	1643	0	1565	215	0
6	H	1685	0	1640	148	0
7	A	2	0	0	0	0
8	A	2	0	0	0	0
All	All	12250	0	11755	1315	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 55.

The worst 5 of 1315 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)
6:H:92:THR:HG23	6:H:120:THR:HA	1.34	1.07
3:A:22:LYS:H	3:A:22:LYS:HD3	1.13	1.06
3:A:441:TYR:CE2	3:A:544:GLY:HA3	1.91	1.05
4:B:60:VAL:HG12	4:B:75:VAL:HG22	1.37	1.04
3:A:501:TYR:CE1	3:A:505:ILE:HD11	1.93	1.04

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	
3	A	556/558 (100%)	408 (73%)	110 (20%)	38 (7%)	1	5
4	B	427/430 (99%)	333 (78%)	62 (14%)	32 (8%)	1	4
5	L	209/211 (99%)	167 (80%)	31 (15%)	11 (5%)	1	9
6	H	223/225 (99%)	187 (84%)	30 (14%)	6 (3%)	4	22
All	All	1415/1424 (99%)	1095 (77%)	233 (16%)	87 (6%)	1	7

5 of 87 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	66	LYS
3	A	195	ILE
3	A	278	GLN
3	A	345	PRO
3	A	393	ILE

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	
3	A	485/498 (97%)	403 (83%)	82 (17%)	2	11
4	B	388/392 (99%)	321 (83%)	67 (17%)	2	10
5	L	190/190 (100%)	166 (87%)	24 (13%)	4	20

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
6	H	196/196 (100%)	168 (86%)	28 (14%)	3	16
All	All	1259/1276 (99%)	1058 (84%)	201 (16%)	2	12

5 of 201 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	250	ASP
4	B	410	TRP
6	H	216	VAL
4	B	277	ARG
4	B	330	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 51 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	242	GLN
4	B	418	ASN
6	H	206	ASN
4	B	255	ASN
4	B	363	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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	$\# Z > 2$
2	MRG	P	817	1,3,2	25,28,29	2.00	4 (16%)	32,39,42	2.49	3 (9%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
2	ATM	P	823	1,2	20,23,24	2.32	8 (40%)	26,32,35	2.75	9 (34%)

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
2	MRG	P	817	1,3,2	-	2/12/26/27	0/3/3/3
2	ATM	P	823	1,2	-	1/10/24/25	0/2/2/2

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	823	ATM	C6-C5	7.10	1.46	1.34
2	P	817	MRG	C2-N2	5.27	1.45	1.34
2	P	817	MRG	C21-N2	-5.22	1.34	1.46
2	P	817	MRG	C2-N1	-3.93	1.30	1.36
2	P	823	ATM	N4'-N3'	3.30	1.31	1.23

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	817	MRG	C21-N2-C2	-12.75	99.41	123.36
2	P	823	ATM	C5A-C5-C4	8.36	127.70	118.78
2	P	823	ATM	C5-C4-N3	5.76	120.33	115.32
2	P	823	ATM	C5A-C5-C6	-5.05	116.02	122.85
2	P	823	ATM	O4-C4-C5	-3.37	121.06	124.92

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	817	MRG	N2-C21-C22-C23
2	P	823	ATM	C3'-N3'-N4'-N5'
2	P	817	MRG	C21-C22-C23-S24

There are no ring outliers.

2 monomers are involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	817	MRG	1	0
2	P	823	ATM	8	0

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](#)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
4	B	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	B	425:LEU	C	426:TRP	N	1.16

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	T	23/27 (85%)	0.36	1 (4%) 40 21	57, 99, 110, 121	0
2	P	19/22 (86%)	0.21	0 100 100	74, 88, 106, 110	0
3	A	556/558 (99%)	-0.07	9 (1%) 70 47	35, 80, 110, 110	1 (0%)
4	B	428/430 (99%)	-0.20	9 (2%) 63 40	27, 62, 108, 110	1 (0%)
5	L	211/211 (100%)	-0.09	1 (0%) 87 72	39, 71, 106, 110	0
6	H	225/225 (100%)	-0.25	4 (1%) 67 44	35, 62, 99, 110	0
All	All	1462/1473 (99%)	-0.13	24 (1%) 70 47	27, 71, 109, 121	2 (0%)

The worst 5 of 24 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	556	ILE	5.2
4	B	428	GLN	4.9
4	B	3	SER	4.0
4	B	425	LEU	3.8
4	B	315	HIS	3.8

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(Å ²)	Q<0.9
2	MRG	P	817	26/27	0.87	0.12	79,79,79,80	0
2	ATM	P	823	22/23	0.93	0.13	62,68,81,86	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
7	MG	A	1002	1/1	0.71	0.17	58,58,58,58	0
7	MG	A	1001	1/1	0.95	0.20	54,54,54,54	0

6.5 Other polymers [i](#)

There are no such residues in this entry.