



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 01:18 PM UTC

PDB ID : 1N5Y / pdb_00001n5y
Title : HIV-1 Reverse Transcriptase Crosslinked to Post-Translocation AZTMP-Terminated DNA (Complex P)
Authors : Sarafianos, S.G.; Clark Jr., A.D.; Das, K.; Tuske, S.; Birktoft, J.J.; Ilankumaran, P.; Ramesha, A.R.; Sayer, J.M.; Jerina, D.M.; Boyer, P.L.; Hughes, S.H.; Arnold, E.
Deposited on : 2002-11-07
Resolution : 3.10 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:

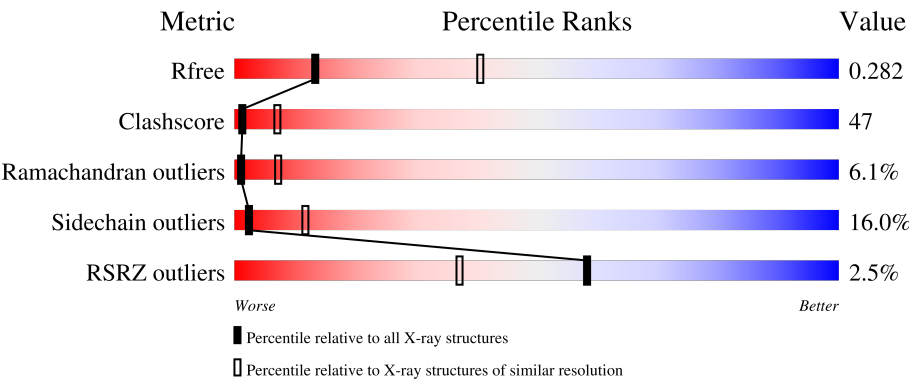
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1456 (3.10-3.10)
Clashscore	190562	1539 (3.10-3.10)
Ramachandran outliers	187476	1467 (3.10-3.10)
Sidechain outliers	187428	1467 (3.10-3.10)
RSRZ outliers	180081	1456 (3.10-3.10)

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	4% 48% 26% 22% .
2	P	21	5% 43% 48% 5%
3	A	558	3% 35% 47% 15% .
4	B	430	3% 28% 51% 19% .

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	MRG	P	817	-	-	X	-

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12185 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*TP*GP*C*TP*AP*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	21	Total	C	N	O	P	0	0	0
			432	204	87	121	20			

- 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*(ATM))-3'.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	P	20	Total	C	N	O	P	S	0	0	0
			408	195	72	121	19	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	18	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	1	Total	Mg	0	0
			1	1		

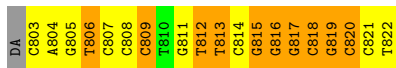
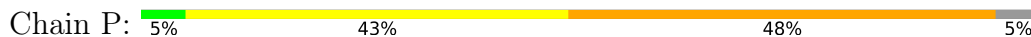
3 Residue-property plots

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.

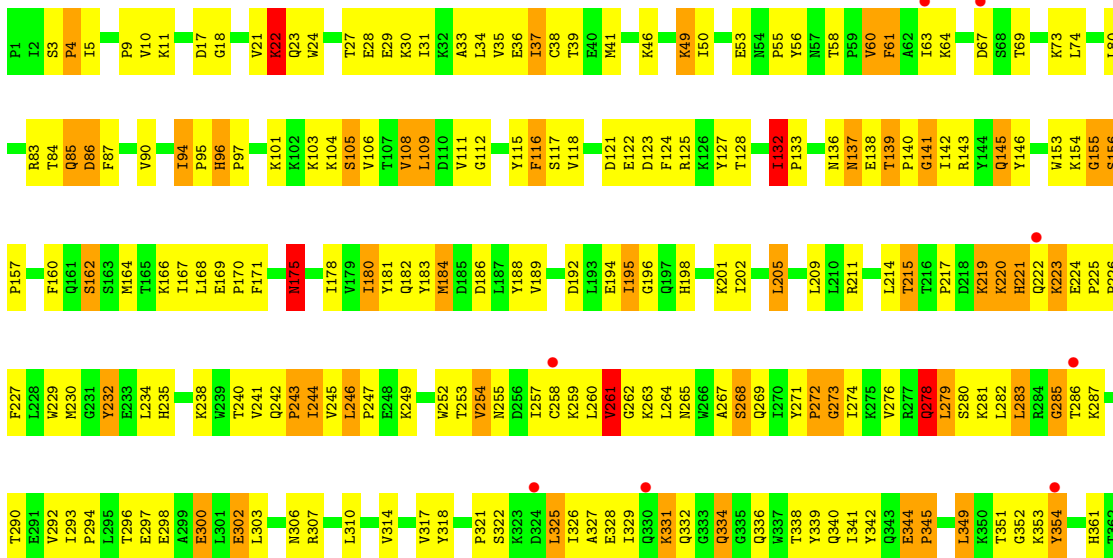
- Molecule 1: 5'-D(*AP*TP*GP*C*TP*AP*GP*GP*CP*GP*CP*CP*CP*GP*AP*AP*CP*AP*GP*GP*GP*AP*CP*TP*GP*TP*G)-3'

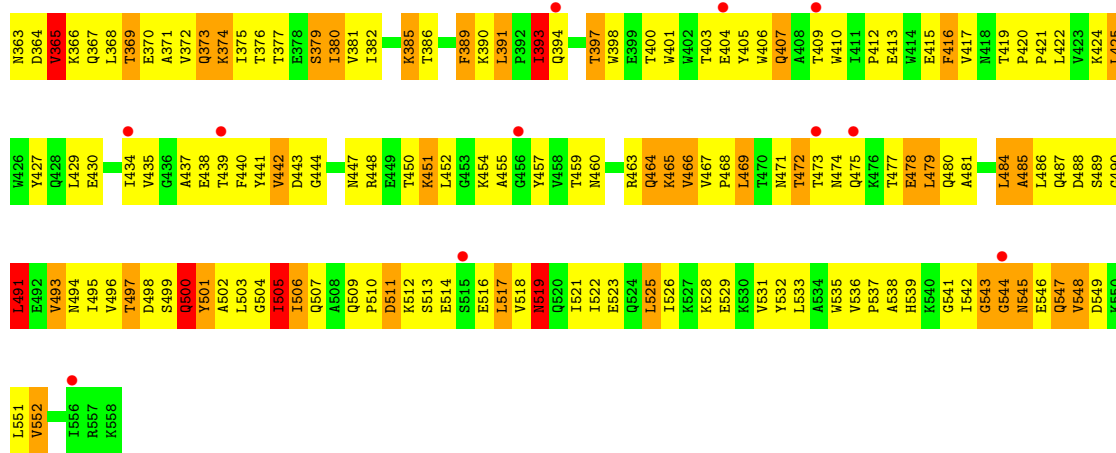


- Molecule 2: 5'-D(*A*CP*AP*GP*TP*CP*CP*CP*TP*GP*TP*TP*CP*GP*GP*(MRG)P*C P*GP*CP*CP*(ATM))-3'

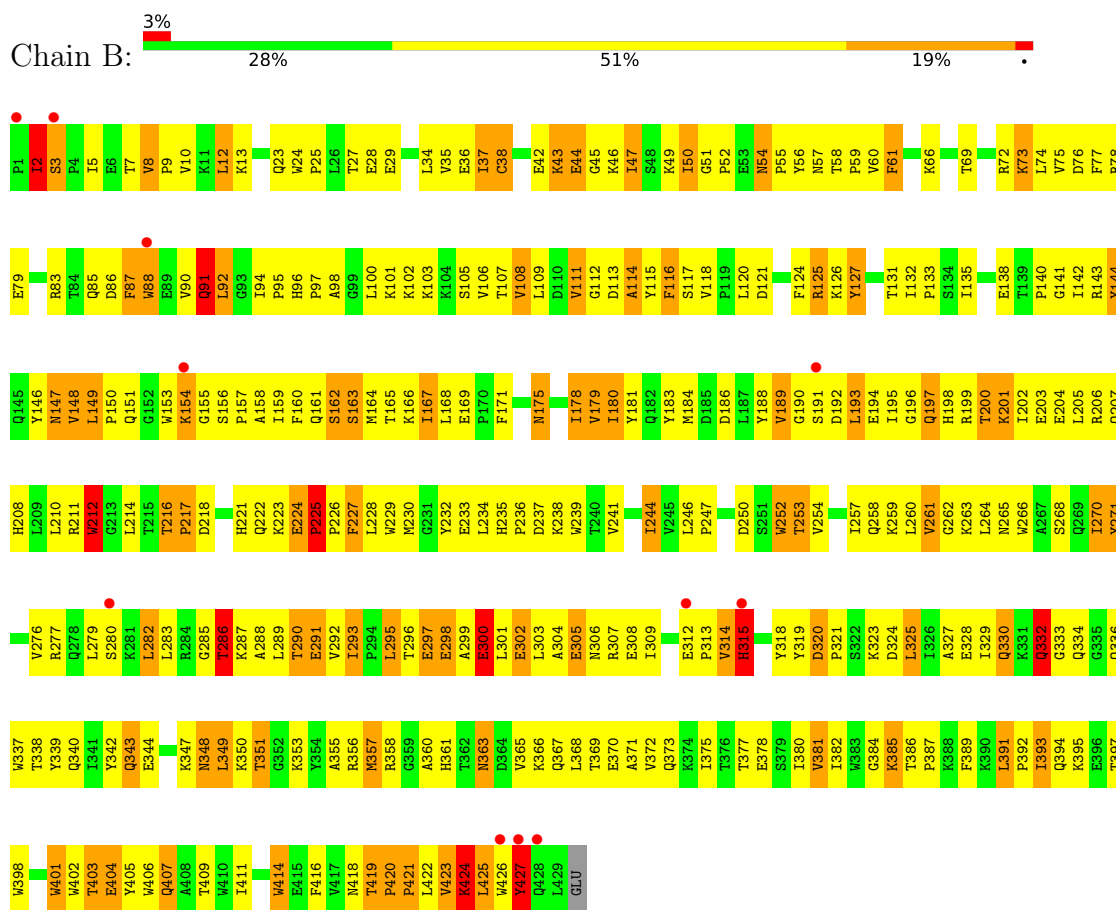


- Molecule 3: REVERSE TRANSCRIPTASE

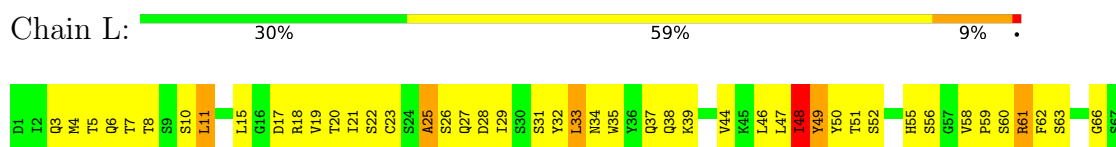


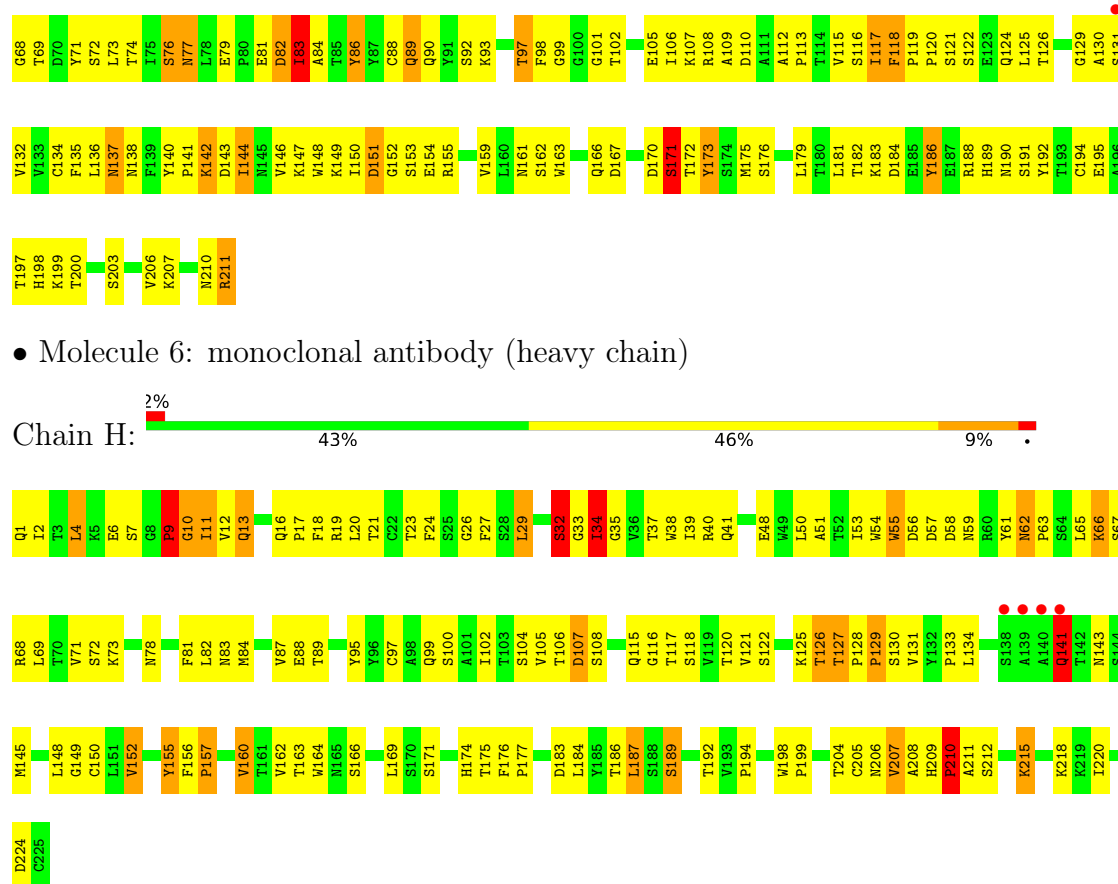


• Molecule 4: REVERSE TRANSCRIPTASE



• 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	166.70Å 166.70Å 221.12Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.10 20.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	91.8 (20.00-3.10) 91.4 (20.00-3.10)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.11	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.54 (at 3.12Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.255 , 0.285 0.253 , 0.282	Depositor DCC
R_{free} test set	2502 reflections (4.05%)	wwPDB-VP
Wilson B-factor (Å ²)	85.4	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 43.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.054 for -h,-k,l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	12185	wwPDB-VP
Average B, all atoms (Å ²)	83.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.91% 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: MG, MRG, ATM

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	1.04	2/486 (0.4%)	1.46	10/749 (1.3%)
2	P	1.28	5/400 (1.2%)	1.86	15/612 (2.5%)
3	A	0.73	1/4600 (0.0%)	1.13	31/6259 (0.5%)
4	B	0.83	2/3639 (0.1%)	1.28	54/4949 (1.1%)
5	L	0.66	0/1681	1.10	12/2283 (0.5%)
6	H	0.80	0/1729	1.26	18/2372 (0.8%)
All	All	0.80	10/12535 (0.1%)	1.23	140/17224 (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
2	P	1	0
3	A	0	1
5	L	0	1
All	All	1	2

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	812	DT	C4-C5	8.87	1.62	1.44
2	P	806	DT	C4-C5	8.56	1.61	1.44
1	T	720	DG	C5-C6	8.19	1.58	1.42
4	B	393	ILE	CA-CB	-7.11	1.50	1.55
1	T	721	DG	C5-C6	6.19	1.54	1.42

The worst 5 of 140 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	816	DG	N9-C1'-C2'	14.69	135.53	113.50
2	P	815	DG	N9-C1'-C2'	14.48	135.22	113.50
4	B	225	PRO	N-CA-C	13.41	127.06	110.70
4	B	86	ASP	N-CA-C	-12.43	97.73	111.28
6	H	127	THR	CA-C-N	12.16	128.54	119.66

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	P	815	DG	C1'

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	354	TYR	Sidechain
5	L	49	TYR	Sidechain

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	T	432	0	235	38	0
2	P	408	0	230	35	0
3	A	4482	0	4484	449	0
4	B	3534	0	3568	380	1
5	L	1643	0	1565	149	0
6	H	1685	0	1640	109	0
7	A	1	0	0	0	0
All	All	12185	0	11722	1106	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 47.

The worst 5 of 1106 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)
5:L:61:ARG:HB2	5:L:76:SER:HB3	1.25	1.17

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
6:H:166:SER:H	6:H:206:ASN:ND2	1.51	1.07
6:H:166:SER:N	6:H:206:ASN:HD21	1.52	1.05
3:A:441:TYR:CE2	3:A:544:GLY:HA3	1.91	1.05
4:B:244:ILE:HD13	4:B:244:ILE:H	1.18	1.04

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)
4:B:2:ILE:CD1	4:B:2:ILE:CD1[6_565]	1.62	0.58

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	
3	A	556/558 (100%)	424 (76%)	102 (18%)	30 (5%)	1	9
4	B	427/430 (99%)	317 (74%)	77 (18%)	33 (8%)	1	4
5	L	209/211 (99%)	165 (79%)	32 (15%)	12 (6%)	1	8
6	H	223/225 (99%)	191 (86%)	21 (9%)	11 (5%)	1	11
All	All	1415/1424 (99%)	1097 (78%)	232 (16%)	86 (6%)	1	7

5 of 86 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	137	ASN
3	A	139	THR
3	A	249	LYS
3	A	254	VAL
3	A	273	GLY

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%)	401 (83%)	84 (17%)	2	9
4	B	388/392 (99%)	322 (83%)	66 (17%)	2	10
5	L	190/190 (100%)	168 (88%)	22 (12%)	5	22
6	H	196/196 (100%)	166 (85%)	30 (15%)	3	12
All	All	1259/1276 (99%)	1057 (84%)	202 (16%)	2	12

5 of 202 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
4	B	241	VAL
4	B	386	THR
6	H	210	PRO
4	B	271	TYR
4	B	318	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 43 such sidechains are listed below:

Mol	Chain	Res	Type
5	L	77	ASN
6	H	59	ASN
5	L	89	GLN
5	L	166	GLN
6	H	79	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](#)

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	ATM	P	822	2,1	20,23,24	2.51	7 (35%)	26,32,35	2.91	12 (46%)
2	MRG	P	817	2,1,3	25,28,29	1.79	6 (24%)	32,39,42	2.01	9 (28%)

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	ATM	P	822	2,1	-	3/10/24/25	0/2/2/2
2	MRG	P	817	2,1,3	-	2/12/26/27	0/3/3/3

The worst 5 of 13 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	822	ATM	C6-C5	6.87	1.45	1.34
2	P	822	ATM	N4'-N3'	5.26	1.36	1.23
2	P	817	MRG	C2-N2	5.15	1.45	1.34
2	P	822	ATM	C1'-N1	-3.87	1.38	1.48
2	P	817	MRG	C21-N2	-3.07	1.39	1.46

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	822	ATM	C5A-C5-C4	9.02	128.41	118.78
2	P	817	MRG	C22-C23-S24	7.50	136.04	113.10
2	P	822	ATM	C5A-C5-C6	-5.35	115.61	122.85
2	P	822	ATM	C5-C4-N3	4.96	119.64	115.32
2	P	822	ATM	C3'-C2'-C1'	4.68	108.22	103.21

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	822	ATM	C4'-C3'-N3'-N4'
2	P	822	ATM	C3'-N3'-N4'-N5'
2	P	817	MRG	C21-C22-C23-S24
2	P	817	MRG	C3'-C4'-C5'-O5'
2	P	822	ATM	C2'-C3'-N3'-N4'

There are no ring outliers.

1 monomer is involved in 9 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	817	MRG	9	0

5.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 1 ligands modelled in this entry, 1 is 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	T	21/27 (77%)	0.58	1 (4%) 35 19	64, 84, 98, 107	0
2	P	18/21 (85%)	0.61	0 100 100	75, 83, 100, 103	0
3	A	556/558 (99%)	0.03	19 (3%) 48 28	44, 91, 110, 110	1 (0%)
4	B	428/430 (99%)	-0.02	11 (2%) 57 36	30, 79, 110, 110	2 (0%)
5	L	211/211 (100%)	0.02	1 (0%) 87 73	56, 88, 110, 110	0
6	H	225/225 (100%)	-0.14	4 (1%) 67 47	51, 78, 105, 110	0
All	All	1459/1472 (99%)	-0.00	36 (2%) 58 37	30, 84, 110, 110	3 (0%)

The worst 5 of 36 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	556	ILE	5.2
3	A	354	TYR	4.5
4	B	315	HIS	4.1
6	H	140	ALA	4.0
4	B	428	GLN	3.9

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	0.78	0.19	92,92,92,92	0
2	ATM	P	822	22/23	0.88	0.15	62,68,87,87	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	1001	1/1	0.81	0.36	62,62,62,62	0

6.5 Other polymers [i](#)

There are no such residues in this entry.