



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:18 PM GMT

PDB ID : 1T05
Title : HIV-1 reverse transcriptase crosslinked to template-primer with tenofovir-diphosphate bound as the incoming nucleotide substrate
Authors : Tuske, S.; Sarafianos, S.G.; Ding, J.; Arnold, E.
Deposited on : 2004-04-07
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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : trunk26865

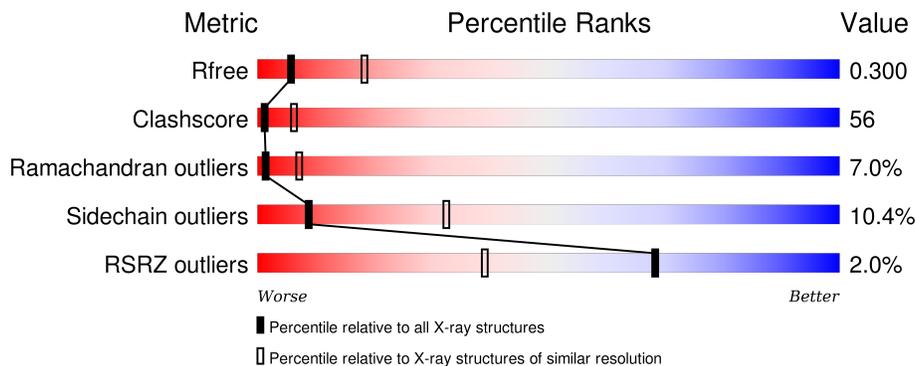
1 Overall quality at a glance

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}	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (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 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	T	27	
2	P	21	
3	A	558	
4	B	437	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
7	GOL	A	824	-	-	X	-
7	GOL	B	438	-	-	-	X

2 Entry composition [i](#)

There are 8 unique types of molecules in this entry. The entry contains 8898 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 oligonucleotide template.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	T	24	494	233	97	141	23	0	0	0

- Molecule 2 is a DNA chain called oligonucleotide primer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	P	20	403	192	72	120	19	0	0	0

- Molecule 3 is a protein called POL polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	554	4506	2917	749	832	8	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

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

- Molecule 4 is a protein called POL polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	B	414	3415	2221	567	621	6	0	0	0

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	ENGINEERED	UNP P04585
B	430	GLY	-	CLONING ARTIFACT	UNP P04585

Continued on next page...

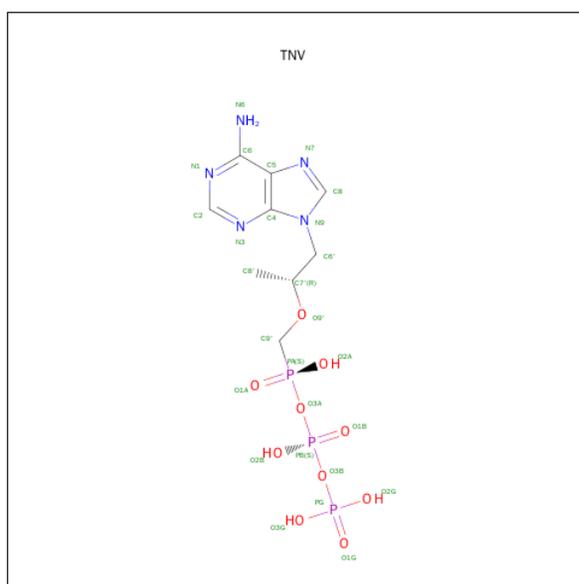
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	431	GLY	-	CLONING ARTIFACT	UNP P04585
B	432	HIS	-	EXPRESSION TAG	UNP P04585
B	433	HIS	-	EXPRESSION TAG	UNP P04585
B	434	HIS	-	EXPRESSION TAG	UNP P04585
B	435	HIS	-	EXPRESSION TAG	UNP P04585
B	436	HIS	-	EXPRESSION TAG	UNP P04585
B	437	HIS	-	EXPRESSION TAG	UNP P04585

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

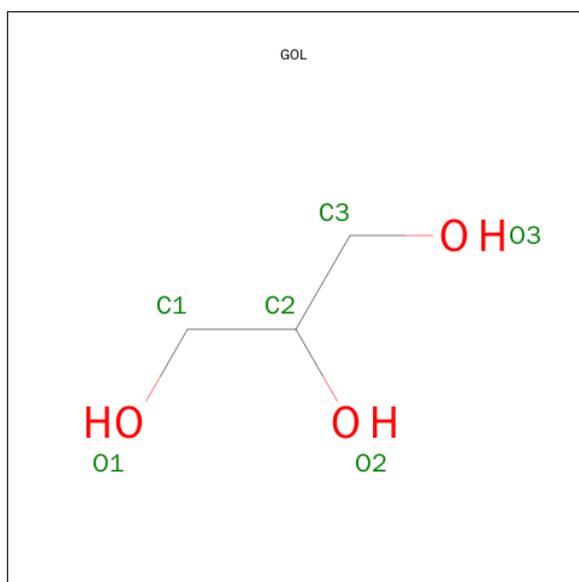
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	2	Total Mg 2 2	0	0

- Molecule 6 is [2-(6-AMINO-9H-PURIN-9-YL)-1-METHYLETHOXY]METHYL-TRIPHOSPHATE (three-letter code: TNV) (formula: C₉H₁₆N₅O₁₀P₃).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	1	Total C N O P 27 9 5 10 3	0	0

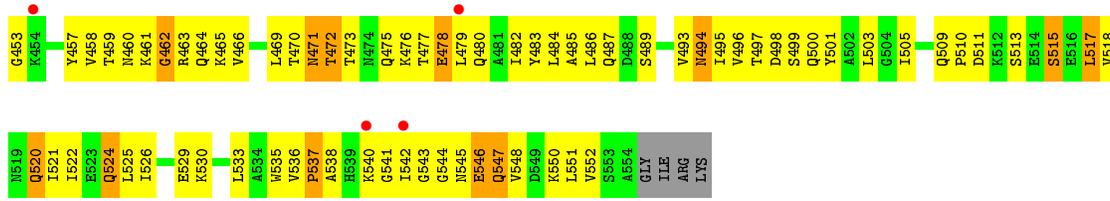
- Molecule 7 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



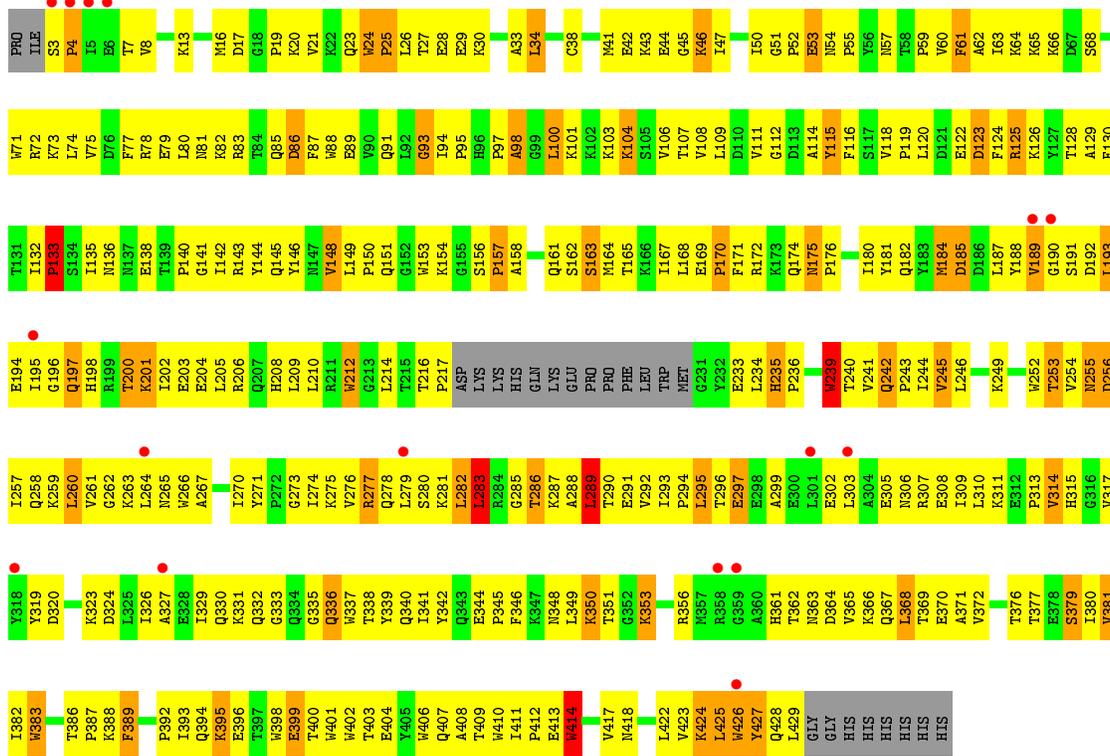
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	1	Total C O 6 3 3	0	0
7	A	1	Total C O 6 3 3	0	0
7	A	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0
7	B	1	Total C O 6 3 3	0	0

- Molecule 8 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	A	10	Total O 10 10	0	0
8	B	6	Total O 6 6	0	0
8	P	3	Total O 3 3	0	0
8	T	2	Total O 2 2	0	0



• Molecule 4: POL polyprotein



4 Data and refinement statistics

Property	Value	Source
Space group	P 31 1 2	Depositor
Cell constants a, b, c, α , β , γ	171.57Å 171.57Å 156.27Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.00 39.85 – 3.00	Depositor EDS
% Data completeness (in resolution range)	81.8 (20.00-3.00) 91.3 (39.85-3.00)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.35 (at 3.01Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.253 , 0.292 0.268 , 0.300	Depositor DCC
R_{free} test set	1922 reflections (4.02%)	DCC
Wilson B-factor (Å ²)	70.0	Xtriage
Anisotropy	0.278	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.27 , 66.3	EDS
Estimated twinning fraction	0.063 for -h,-k,l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Outliers	0 of 48012 reflections	Xtriage
F_o, F_c correlation	0.90	EDS
Total number of atoms	8898	wwPDB-VP
Average B, all atoms (Å ²)	84.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.97% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, MG, TNV, DDG, MRG

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	0.59	0/555	1.13	2/856 (0.2%)
2	P	0.59	0/400	1.24	5/612 (0.8%)
3	A	0.55	1/4624 (0.0%)	0.80	3/6282 (0.0%)
4	B	0.48	0/3512	0.78	4/4774 (0.1%)
All	All	0.53	1/9091 (0.0%)	0.84	14/12524 (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
3	A	0	1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	478	GLU	CD-OE2	7.31	1.33	1.25

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	3	SER	C-N-CD	-20.07	76.44	120.60
4	B	295	LEU	N-CA-C	6.77	129.27	111.00
1	T	718	DA	N9-C1'-C2'	6.35	124.66	112.60
3	A	100	LEU	CA-CB-CG	6.22	129.60	115.30
2	P	815	DG	N9-C1'-C2'	6.15	124.28	112.60

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
3	A	160	PHE	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	494	0	269	24	0
2	P	403	0	224	30	0
3	A	4506	0	4567	517	0
4	B	3415	0	3448	443	0
5	A	2	0	0	0	0
6	A	27	0	12	7	0
7	A	18	0	21	5	0
7	B	12	0	14	1	0
8	A	10	0	0	1	0
8	B	6	0	0	1	0
8	P	3	0	0	0	0
8	T	2	0	0	0	0
All	All	8898	0	8555	981	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 56.

The worst 5 of 981 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)
3:A:498:ASP:HB2	3:A:538:ALA:HB2	1.31	1.08
3:A:103:LYS:H	7:A:824:GOL:H32	1.13	1.08
3:A:205:LEU:HD22	3:A:209:LEU:HD11	1.35	1.07
3:A:408:ALA:HB1	4:B:364:ASP:HB3	1.37	1.07
3:A:72:ARG:HH12	6:A:823:TNV:H8'3	1.12	1.07

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	552/558 (99%)	410 (74%)	112 (20%)	30 (5%)	2	14
4	B	410/437 (94%)	278 (68%)	95 (23%)	37 (9%)	1	4
All	All	962/995 (97%)	688 (72%)	207 (22%)	67 (7%)	1	7

5 of 67 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	77	PHE
3	A	154	LYS
3	A	268	SER
3	A	345	PRO
3	A	438	GLU

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	495/498 (99%)	437 (88%)	58 (12%)	7	27
4	B	376/397 (95%)	343 (91%)	33 (9%)	12	42
All	All	871/895 (97%)	780 (90%)	91 (10%)	9	32

5 of 91 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	314	VAL

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	425	LEU
4	B	349	LEU
3	A	340	GLN
3	A	362	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
3	A	487	GLN
3	A	509	GLN
4	B	340	GLN
3	A	500	GLN
3	A	519	ASN

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	MRG	P	817	1,2	16,24,29	1.41	2 (12%)	20,35,42	2.69	4 (20%)
2	DDG	P	822	1,2	15,23,24	1.15	1 (6%)	16,33,36	2.96	4 (25%)

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	1,2	-	0/3/21/27	0/3/3/3
2	DDG	P	822	1,2	-	0/3/18/19	0/3/3/3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	817	MRG	C2-N1	2.39	1.39	1.35
2	P	822	DDG	C6-N1	3.42	1.39	1.33
2	P	817	MRG	C6-N1	4.24	1.41	1.33

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	822	DDG	C5-C6-N1	-8.75	111.62	123.59
2	P	817	MRG	C5-C6-N1	-8.72	111.66	123.59
2	P	817	MRG	O3'-C3'-C2'	-3.16	100.28	110.74
2	P	817	MRG	N3-C2-N1	-2.73	123.29	127.44
2	P	822	DDG	N3-C2-N1	-2.42	123.76	127.44

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	817	MRG	3	0
2	P	822	DDG	2	0

5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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$
6	TNV	A	823	5	21,28,28	1.42	3 (14%)	21,43,43	1.53	4 (19%)
7	GOL	A	824	-	5,5,5	1.32	1 (20%)	5,5,5	2.03	2 (40%)
7	GOL	A	825	-	5,5,5	1.25	1 (20%)	5,5,5	2.00	2 (40%)
7	GOL	A	826	-	5,5,5	1.09	0	5,5,5	2.00	2 (40%)
7	GOL	B	438	-	5,5,5	0.85	0	5,5,5	2.03	2 (40%)
7	GOL	B	439	-	5,5,5	1.04	0	5,5,5	1.99	2 (40%)

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
6	TNV	A	823	5	-	0/17/22/22	0/2/2/2
7	GOL	A	824	-	-	0/4/4/4	0/0/0/0
7	GOL	A	825	-	-	0/4/4/4	0/0/0/0
7	GOL	A	826	-	-	0/4/4/4	0/0/0/0
7	GOL	B	438	-	-	0/4/4/4	0/0/0/0
7	GOL	B	439	-	-	0/4/4/4	0/0/0/0

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	823	TNV	C6'-N9	-3.06	1.45	1.48
6	A	823	TNV	PA-O3A	-3.02	1.55	1.58
7	A	825	GOL	C3-C2	-2.12	1.44	1.52
7	A	824	GOL	C3-C2	-2.08	1.44	1.52
6	A	823	TNV	C2-N3	2.24	1.36	1.32

The worst 5 of 14 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	823	TNV	PB-O3A-PA	-4.47	120.17	132.73
6	A	823	TNV	PB-O3B-PG	-3.08	122.35	132.67
6	A	823	TNV	O3A-PA-C9'	2.42	112.59	104.16
6	A	823	TNV	C9'-O9'-C7'	2.47	116.33	114.35
7	B	438	GOL	O2-C2-C1	2.71	121.07	108.65

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 13 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
6	A	823	TNV	7	0
7	A	824	GOL	4	0
7	A	825	GOL	1	0
7	B	438	GOL	1	0
7	B	439	GOL	1	0

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	24/27 (88%)	-0.23	0 100 100	42, 100, 120, 120	0
2	P	18/21 (85%)	-0.26	0 100 100	42, 104, 120, 120	0
3	A	554/558 (99%)	-0.25	4 (0%) 89 70	13, 71, 119, 120	0
4	B	414/437 (94%)	0.01	16 (3%) 43 18	30, 100, 120, 120	0
All	All	1010/1043 (96%)	-0.14	20 (1%) 68 39	13, 86, 120, 120	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	B	3	SER	9.5
4	B	4	PRO	6.0
4	B	5	ILE	4.6
4	B	358	ARG	3.6
4	B	279	LEU	3.6

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
2	DDG	P	822	21/22	0.96	0.19	-	39,53,60,74	0
2	MRG	P	817	22/27	0.94	0.12	-	37,60,89,98	0

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
7	GOL	B	438	6/6	0.90	0.47	9.43	56,70,72,74	6
5	MG	A	601	1/1	0.94	0.24	1.31	73,73,73,73	0
5	MG	A	600	1/1	0.90	0.21	1.11	16,16,16,16	0
7	GOL	A	824	6/6	0.83	0.25	0.98	54,58,60,63	6
6	TNV	A	823	27/27	0.96	0.21	0.43	34,54,64,71	0
7	GOL	A	825	6/6	0.88	0.18	0.19	45,59,68,81	6
7	GOL	A	826	6/6	0.85	0.29	-	36,42,49,50	6
7	GOL	B	439	6/6	0.89	0.47	-	63,70,74,75	6

6.5 Other polymers [i](#)

There are no such residues in this entry.