



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 08:47 AM UTC

PDB ID : 1T03 / pdb_00001t03
Title : HIV-1 reverse transcriptase crosslinked to tenofovir terminated template-primer (complex P)
Authors : Tuske, S.; Sarafianos, S.G.; Ding, J.; Arnold, E.
Deposited on : 2004-04-07
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

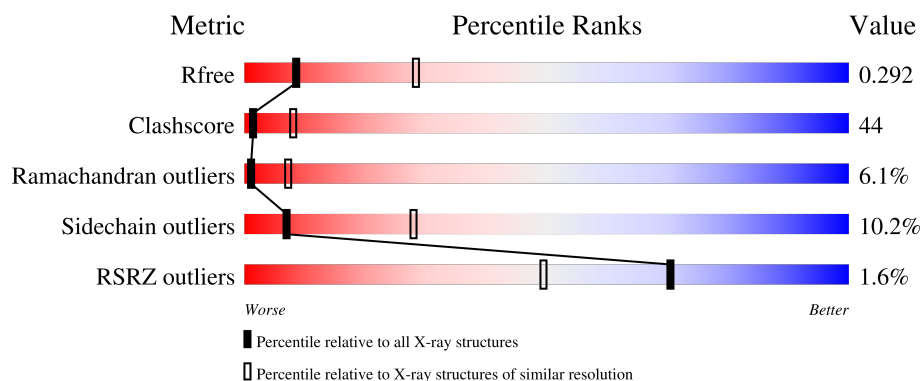
1 Overall quality at a glance

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	7% 33% 22% 30% 15%
2	P	21	5% 24% 48% 24% 5%
3	A	558	2% 38% 50% 11% .
4	B	437	% 33% 50% 14% ..
5	L	211	36% 55% 8%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
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	TFO	P	822[A]	-	X	-	X
2	TFO	P	822[B]	-	-	-	X

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 12252 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 Synthetic oligonucleotide template.

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 Synthetic oligonucleotide primer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	P	20	Total	C	N	O	P	0	2	0
			434	209	80	125	20			

- Molecule 3 is a protein called POL polyprotein.

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 POL polyprotein.

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

There are 9 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	280	SER	CYS	engineered mutation	UNP P03366
B	430	GLY	-	cloning artifact	UNP P03366

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	431	GLY	-	cloning artifact	UNP P03366
B	432	HIS	-	expression tag	UNP P03366
B	433	HIS	-	expression tag	UNP P03366
B	434	HIS	-	expression tag	UNP P03366
B	435	HIS	-	expression tag	UNP P03366
B	436	HIS	-	expression tag	UNP P03366
B	437	HIS	-	expression tag	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			

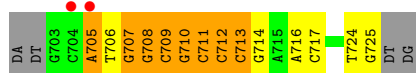
- Molecule 6 is a protein called monoclonal antibody heavy chain.

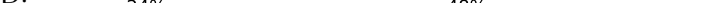
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	H	225	Total	C	N	O	S	0	0	0
			1685	1060	276	340	9			

- 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		

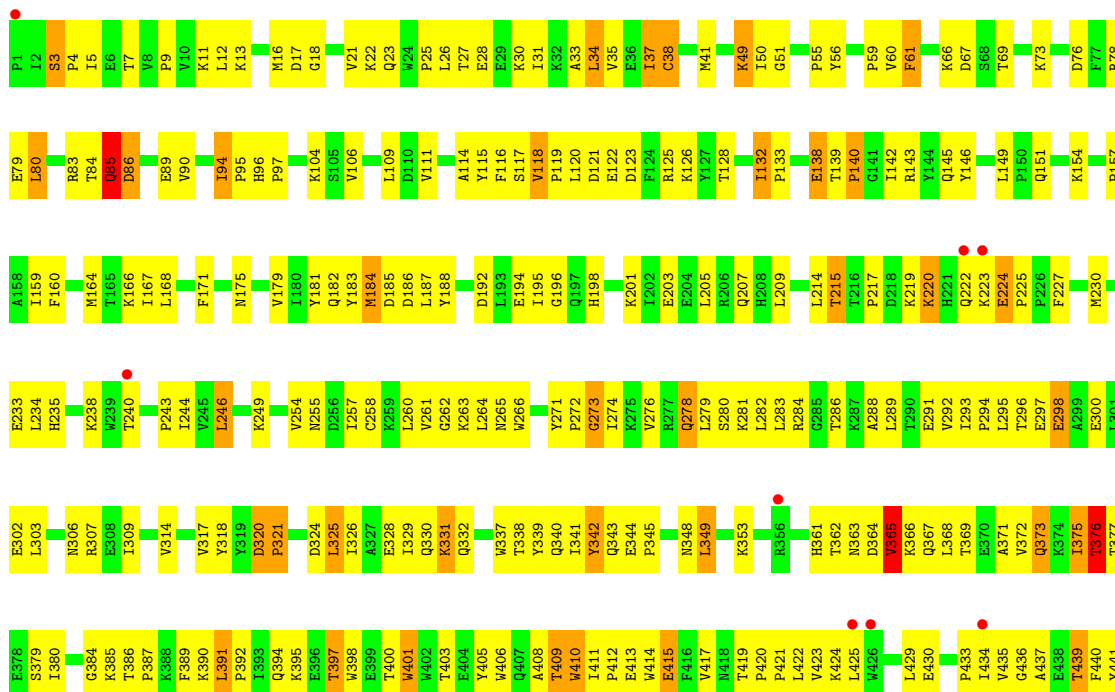
- Molecule 1: Synthetic oligonucleotide template

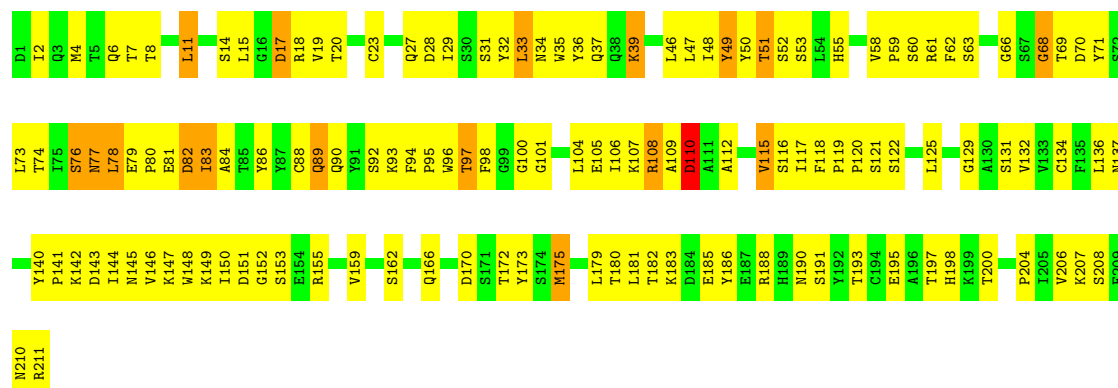


- Chain P: 



- Chain A: 2% 38% 50% 11%





Chain H:

40% 50% 8%

Q1 Q2 Q3 Q4 Q5 Q6 Q7 Q8 Q9 Q10 Q11 Q12 Q13 Q16 Q17 Q18 Q21 Q22 Q23 Q24 Q27 Q28 Q29 Q30 Q31 Q32 Q33 Q34 Q35 Q38 Q39 Q40 Q41 Q42 Q43 Q44 Q47 Q48 Q49 Q50 Q53 Q54 Q55 Q56 Q57 Q58 Q59 Q60 Q61 Q62 Q63 Q64 Q69 Q70 Q71 Q72 Q73 Q74 Q77 Q78 Q79 Q80 Q81 Q82 Q83 Q84 Q85 Q86 Q87 Q88 Q89 Q90 Q91 Q92 Q93 Q94 Q95 Q96 Q97 Q98 Q99 S104 S105 S106 S107 S108 S109 S110 S111 S112 S113 S114 S115 S116 S117 S118 S119 S120 S121 S122 S123 S124 S125 S126 S127 S128 S129 S130 S131 S132 S133 S134 S135 S136 S137 S138 S139 S140 S141 S142 S143 S144 S145 S146 S147 S148 S149 S150 S151 S152 S153 S154 S155 S156 S157 S158 S159 S160 S161 S162 S163 S164 S165 S166 S167 S168 S169 S170 S171 S172 S173 S174 S175 S176 S177 S178 S179 S180 S181 S182 S183 S184 S185 S186 S187 S188 S189 S190 S191 S192 S193 S194 S195 S196 S197 S198 S199 S200 S201 S202 S203 S204 S205 S206 S207 S208 S209 S210 S211 S212 S213 S214 S215 S216 S217 S218 S219 S220 S221 S222 S223 S224 S225

4 Data and refinement statistics

Property	Value	Source
Space group	P 32 1 2	Depositor
Cell constants a, b, c, α , β , γ	166.78Å 166.78Å 221.13Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.10 20.00 – 3.10	Depositor EDS
% Data completeness (in resolution range)	93.9 (20.00-3.10) 94.6 (20.00-3.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.01 (at 3.12Å)	Xtriage
Refinement program	CNS	Depositor
R, R_{free}	0.256 , 0.295 0.251 , 0.292	Depositor DCC
R_{free} test set	2454 reflections (4.04%)	wwPDB-VP
Wilson B-factor (Å ²)	70.1	Xtriage
Anisotropy	0.010	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 32.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.43$, $\langle L^2 \rangle = 0.26$	Xtriage
Estimated twinning fraction	0.057 for -h,-k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	12252	wwPDB-VP
Average B, all atoms (Å ²)	64.0	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: TFO, MG, 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	0.43	0/532	1.06	12/820 (1.5%)
2	P	0.39	0/420	0.88	4/640 (0.6%)
3	A	0.53	0/4600	1.01	20/6259 (0.3%)
4	B	0.63	1/3639 (0.0%)	1.15	37/4949 (0.7%)
5	L	0.51	0/1681	1.05	7/2283 (0.3%)
6	H	0.61	0/1729	1.08	6/2372 (0.3%)
All	All	0.56	1/12601 (0.0%)	1.06	86/17323 (0.5%)

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

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	132	ILE	CA-CB	-6.07	1.49	1.54

The worst 5 of 86 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	86	ASP	N-CA-C	-12.51	97.31	111.82
4	B	427	TYR	N-CA-C	12.11	128.10	110.42
4	B	225	PRO	N-CA-C	11.32	124.51	110.70
3	A	365	VAL	N-CA-C	-10.60	100.54	110.82
1	T	705	DA	N9-C1'-C2'	-9.89	98.66	113.50

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	B	188	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	32	0
2	P	434	0	248	23	0
3	A	4482	0	4485	430	0
4	B	3534	0	3568	329	1
5	L	1643	0	1565	165	0
6	H	1685	0	1640	127	0
7	A	1	0	0	0	0
All	All	12252	0	11763	1062	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 44.

The worst 5 of 1062 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:501:TYR:CE1	3:A:505:ILE:HD11	1.86	1.10
3:A:50:ILE:HD11	3:A:145:GLN:HB2	1.10	1.09
3:A:344:GLU:HG3	3:A:345:PRO:HD2	1.36	1.07
3:A:138:GLU:HG2	3:A:139:THR:H	1.10	1.06
5:L:61:ARG:HB2	5:L:76:SER:HB3	1.36	1.06

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:O	4:B:2:ILE:O[6_565]	1.67	0.53

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%)	431 (78%)	92 (16%)	33 (6%)	1	8
4	B	427/437 (98%)	324 (76%)	75 (18%)	28 (7%)	1	6
5	L	209/211 (99%)	165 (79%)	34 (16%)	10 (5%)	2	11
6	H	223/225 (99%)	185 (83%)	23 (10%)	15 (7%)	1	6
All	All	1415/1431 (99%)	1105 (78%)	224 (16%)	86 (6%)	1	7

5 of 86 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	67	ASP
3	A	138	GLU
3	A	273	GLY
3	A	466	VAL
3	A	474	ASN

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	
3	A	485/498 (97%)	437 (90%)	48 (10%)	7	29
4	B	388/397 (98%)	344 (89%)	44 (11%)	5	23
5	L	190/190 (100%)	177 (93%)	13 (7%)	14	42
6	H	196/196 (100%)	173 (88%)	23 (12%)	5	22
All	All	1259/1281 (98%)	1131 (90%)	128 (10%)	7	28

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

Mol	Chain	Res	Type
6	H	123	SER
6	H	159	PRO
4	B	12	LEU
3	A	552	VAL
6	H	163	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
4	B	258	GLN
5	L	27	GLN
4	B	315	HIS
4	B	367	GLN
5	L	89	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

3 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	TFO	P	822[A]	2	15,19,20	3.08	9 (60%)	20,26,29	6.21	10 (50%)
2	TFO	P	822[B]	2	15,19,20	0.51	0	20,26,29	1.77	3 (15%)
2	MRG	P	817	2,1	21,24,29	0.42	0	30,35,42	0.61	1 (3%)

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	TFO	P	822[A]	2	-	4/7/9/10	0/2/2/2
2	TFO	P	822[B]	2	-	2/7/9/10	0/2/2/2
2	MRG	P	817	2,1	-	0/7/21/27	0/3/3/3

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	P	822[A]	TFO	C8'-C7'	5.77	1.73	1.51
2	P	822[A]	TFO	C6'-N9	-5.15	1.39	1.46
2	P	822[A]	TFO	O9'-C7'	-4.65	1.36	1.44
2	P	822[A]	TFO	C4-N9	-4.32	1.32	1.37
2	P	822[A]	TFO	C8-N9	3.62	1.42	1.36

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	P	822[A]	TFO	C7'-C6'-N9	22.06	147.53	112.68
2	P	822[A]	TFO	C6'-N9-C4	-9.87	114.95	126.41
2	P	822[A]	TFO	C6'-N9-C8	7.24	138.20	126.63
2	P	822[A]	TFO	N3-C4-N9	-7.05	118.15	126.87
2	P	822[B]	TFO	C7'-C6'-N9	5.44	121.28	112.68

There are no chirality outliers.

5 of 6 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	P	822[A]	TFO	N9-C6'-C7'-O9'
2	P	822[A]	TFO	N9-C6'-C7'-C8'
2	P	822[A]	TFO	C7'-C6'-N9-C4
2	P	822[B]	TFO	C7'-C6'-N9-C4
2	P	822[A]	TFO	C7'-C6'-N9-C8

There are no ring outliers.

3 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	P	822[A]	TFO	2	0
2	P	822[B]	TFO	5	0
2	P	817	MRG	3	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	23/27 (85%)	0.59	2 (8%) 16 9	53, 99, 132, 135	0
2	P	18/21 (85%)	0.33	1 (5%) 30 16	33, 87, 109, 121	1 (5%)
3	A	556/558 (99%)	-0.00	12 (2%) 62 41	16, 69, 117, 131	1 (0%)
4	B	429/437 (98%)	-0.31	5 (1%) 76 58	15, 48, 112, 128	1 (0%)
5	L	211/211 (100%)	-0.25	0 100 100	26, 60, 105, 116	0
6	H	225/225 (100%)	-0.41	3 (1%) 75 56	16, 48, 92, 129	0
All	All	1462/1479 (98%)	-0.18	23 (1%) 70 49	15, 58, 113, 135	3 (0%)

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	556	ILE	4.9
3	A	1	PRO	4.3
4	B	429	LEU	3.5
3	A	222	GLN	3.5
1	T	705	DA	3.2

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	TFO	P	822[A]	18/19	0.72	0.44	85,98,98,98	18
2	TFO	P	822[B]	18/19	0.72	0.44	86,98,98,98	18
2	MRG	P	817	22/27	0.87	0.12	97,109,112,115	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	559	1/1	0.99	0.14	18,18,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.