



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2026 – 07:03 PM UTC

PDB ID : 1MUH / pdb_00001muh
Title : CRYSTAL STRUCTURE OF TN5 TRANSPOSASE COMPLEXED WITH
TRANSPOSON END DNA
Authors : Thoden, J.B.; Holden, H.M.; Davies, D.R.; Goryshin, I.Y.; Reznikoff, W.S.;
Rayment, I.
Deposited on : 2002-09-23
Resolution : 2.30 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
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

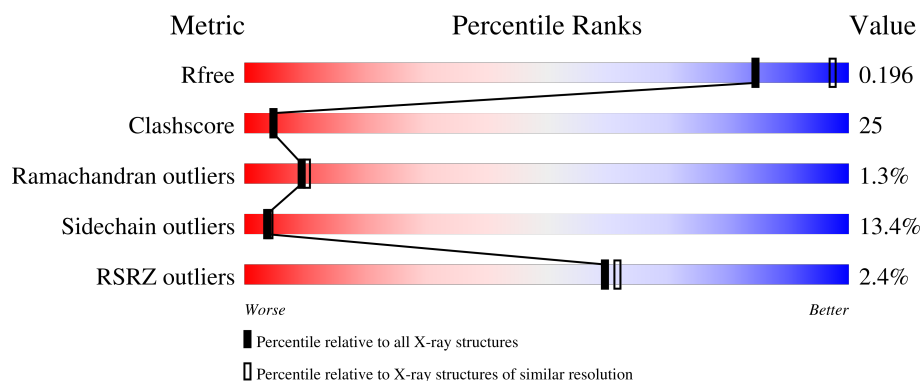
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.30 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	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	B	20	
2	C	20	
3	A	481	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 4691 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 DNA TRANSFERRED STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	20	Total	C	N	O	P	0	0	0
			413	198	78	118	19			

- Molecule 2 is a DNA chain called DNA NON-TRANSFERRED STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	20	Total	C	N	O	P	0	0	0
			401	194	70	118	19			

- Molecule 3 is a protein called Tn5 transposase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	455	Total	C	N	O	S	0	1	0
			3569	2250	653	654	12			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	54	LYS	GLU	engineered mutation	UNP Q46731
A	56	ALA	MET	engineered mutation	UNP Q46731
A	372	PRO	LEU	engineered mutation	UNP Q46731
A	477	GLY	-	cloning artifact	UNP Q46731

- Molecule 4 is MANGANESE (II) ION (CCD ID: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Mn	0	0
			1	1		

- Molecule 5 is MAGNESIUM ION (CCD ID: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	1	Total 1	Mg 1	0	0

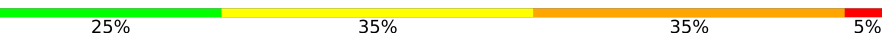
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	27	Total 27	O 27	0	0
6	C	22	Total 22	O 22	0	0
6	A	257	Total 257	O 257	0	0

3 Residue-property plots

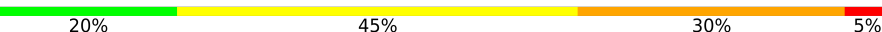
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: DNA TRANSFERRED STRAND

Chain B: 



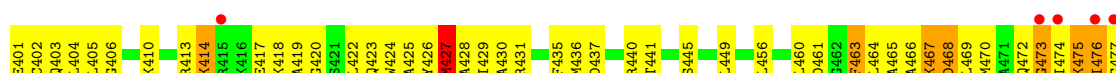
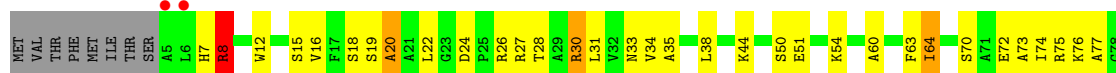
• Molecule 2: DNA NON-TRANSFERRED STRAND

Chain C: 



• Molecule 3: Tn5 transposase

Chain A: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α , β , γ	113.70Å 113.70Å 228.10Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 2.30 30.00 – 2.30	Depositor EDS
% Data completeness (in resolution range)	100.0 (30.00-2.30) 99.7 (30.00-2.30)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.90 (at 2.31Å)	Xtriage
Refinement program	TNT, CNS	Depositor
R, R_{free}	0.204 , 0.267 0.199 , 0.196	Depositor DCC
R_{free} test set	4023 reflections (10.21%)	wwPDB-VP
Wilson B-factor (Å ²)	44.4	Xtriage
Anisotropy	0.244	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 147.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	4691	wwPDB-VP
Average B, all atoms (Å ²)	52.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.03% 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: MN, 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	B	0.75	0/464	2.18	24/716 (3.4%)
2	C	0.75	0/448	1.96	21/688 (3.1%)
3	A	1.27	2/3642 (0.1%)	1.92	77/4914 (1.6%)
All	All	1.19	2/4554 (0.0%)	1.96	122/6318 (1.9%)

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
1	B	0	1
2	C	0	1
All	All	0	2

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	250	ARG	NE-CZ	5.38	1.39	1.33
3	A	64	ILE	CA-CB	-5.05	1.48	1.54

The worst 5 of 122 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	19	DA	C4-N9-C1'	-19.41	97.94	127.05
1	B	19	DA	C8-N9-C1'	14.72	149.14	127.05
1	B	1	DG	C4-N9-C1'	-13.71	106.44	127.00
1	B	1	DG	C8-N9-C1'	12.61	145.92	127.00
2	C	4	DA	C8-N9-C1'	11.80	144.75	127.05

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	19	DA	Sidechain
2	C	11	DT	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	B	413	0	227	12	0
2	C	401	0	228	22	0
3	A	3569	0	3601	185	0
4	A	1	0	0	0	0
5	A	1	0	0	0	0
6	A	257	0	0	7	0
6	B	27	0	0	0	0
6	C	22	0	0	3	0
All	All	4691	0	4056	213	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 25.

The worst 5 of 213 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:189:ARG:HB3	3:A:212:LYS:HB2	1.22	1.19
1:B:4:DT:H5"	1:B:4:DT:H6	1.31	0.94
3:A:194:HIS:CD2	3:A:274:GLN:HG3	2.02	0.94
3:A:194:HIS:CE1	3:A:274:GLN:HB2	2.08	0.88
3:A:472:GLN:HG3	3:A:474:ILE:HD11	1.55	0.88

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	
3	A	452/481 (94%)	423 (94%)	23 (5%)	6 (1%)	9	10

5 of 6 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	274	GLN
3	A	192	ASP
3	A	468	ASP
3	A	467	LYS
3	A	473	GLY

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	367/393 (93%)	318 (87%)	49 (13%)	4	4

5 of 49 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	270	ILE
3	A	297	LYS
3	A	273	LYS
3	A	279	LEU
3	A	343	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such

sidechains are listed below:

Mol	Chain	Res	Type
3	A	81	GLN
3	A	231	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 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](#)

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	B	20/20 (100%)	-0.65	0 100 100	33, 46, 84, 97	0
2	C	20/20 (100%)	-0.41	0 100 100	34, 56, 78, 90	0
3	A	455/481 (94%)	-0.08	12 (2%) 57 59	26, 47, 90, 100	1 (0%)
All	All	495/521 (95%)	-0.12	12 (2%) 59 62	26, 47, 90, 100	1 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	476	ILE	3.7
3	A	474	ILE	3.6
3	A	372	PRO	2.9
3	A	477	GLY	2.8
3	A	5	ALA	2.6

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 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
5	MG	A	479	1/1	0.98	0.04	50,50,50,50	0
4	MN	A	478	1/1	1.00	0.02	42,42,42,42	0

6.5 Other polymers [i](#)

There are no such residues in this entry.