



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 6, 2026 – 07:01 AM UTC

PDB ID : 3UGM / pdb_00003ugm
Title : Structure of TAL effector PthXo1 bound to its DNA target
Authors : Mak, A.N.S.; Bradley, P.; Cernadas, R.A.; Bogdanove, A.J.; Stoddard, B.L.
Deposited on : 2011-11-02
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
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

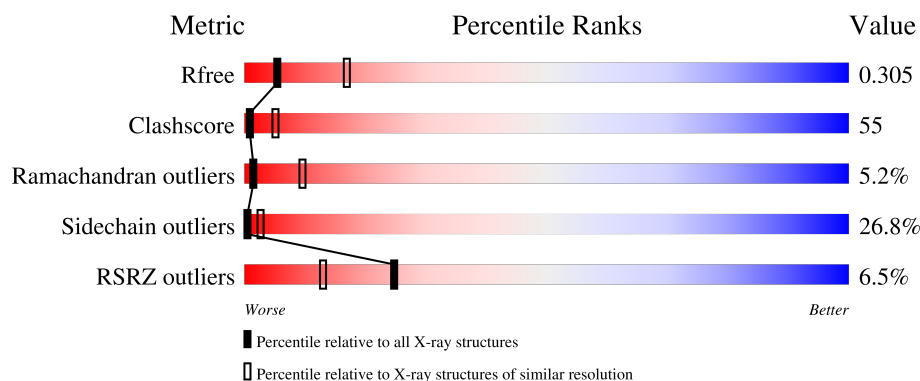
1 Overall quality at a glance

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	A	1047	5% 20% 37% 21% • 18%
2	B	38	8% 29% 61% 11%
3	C	38	3% 34% 63% •

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7853 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 protein called TAL effector AvrBs3/PthA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	854	Total	C	N	O	S	0	0	0
			6085	3792	1123	1147	23			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	103	MET	-	expression tag	UNP B2SU53
A	104	ALA	-	expression tag	UNP B2SU53
A	105	SER	-	expression tag	UNP B2SU53
A	106	SER	-	expression tag	UNP B2SU53
A	107	HIS	-	expression tag	UNP B2SU53
A	108	HIS	-	expression tag	UNP B2SU53
A	109	HIS	-	expression tag	UNP B2SU53
A	110	HIS	-	expression tag	UNP B2SU53
A	111	HIS	-	expression tag	UNP B2SU53
A	112	HIS	-	expression tag	UNP B2SU53
A	113	SER	-	expression tag	UNP B2SU53
A	114	SER	-	expression tag	UNP B2SU53
A	115	GLY	-	expression tag	UNP B2SU53
A	116	LEU	-	expression tag	UNP B2SU53
A	117	VAL	-	expression tag	UNP B2SU53
A	118	PRO	-	expression tag	UNP B2SU53
A	119	ARG	-	expression tag	UNP B2SU53
A	120	GLY	-	expression tag	UNP B2SU53
A	121	SER	-	expression tag	UNP B2SU53
A	122	SER	-	expression tag	UNP B2SU53
A	123	GLY	-	expression tag	UNP B2SU53
A	124	SER	-	expression tag	UNP B2SU53
A	125	SER	-	expression tag	UNP B2SU53
A	126	MET	-	expression tag	UNP B2SU53
A	173	ARG	GLY	engineered mutation	UNP B2SU53
A	198	GLN	ARG	engineered mutation	UNP B2SU53
A	209	THR	LYS	engineered mutation	UNP B2SU53

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	212	HIS	ASP	engineered mutation	UNP B2SU53
A	213	ILE	MET	engineered mutation	UNP B2SU53
A	215	THR	ALA	engineered mutation	UNP B2SU53
A	244	ASP	VAL	engineered mutation	UNP B2SU53
A	272	MET	VAL	engineered mutation	UNP B2SU53

- Molecule 2 is a DNA chain called DNA-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	38	Total	C	N	O	P	0	0	0
			764	367	137	223	37			

- Molecule 3 is a DNA chain called DNA-2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	38	Total	C	N	O	P	0	0	0
			788	376	146	229	37			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	159	Total	O	0	0
			159	159		
4	B	23	Total	O	0	0
			23	23		
4	C	34	Total	O	0	0
			34	34		



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	95.58Å 248.48Å 54.65Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 3.00 50.00 – 3.00	Depositor EDS
% Data completeness (in resolution range)	95.3 (50.00-3.00) 95.8 (50.00-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.01 (at 3.01Å)	Xtriage
Refinement program	REFMAC, CNS	Depositor
R, R_{free}	0.266 , 0.294 0.267 , 0.305	Depositor DCC
R_{free} test set	1326 reflections (4.84%)	wwPDB-VP
Wilson B-factor (Å ²)	76.4	Xtriage
Anisotropy	0.259	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 114.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.89	EDS
Total number of atoms	7853	wwPDB-VP
Average B, all atoms (Å ²)	85.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.50% 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

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	A	1.48	69/6153 (1.1%)	1.70	205/8397 (2.4%)
2	B	0.78	0/855	1.16	5/1314 (0.4%)
3	C	0.67	0/885	1.05	2/1368 (0.1%)
All	All	1.35	69/7893 (0.9%)	1.57	212/11079 (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	A	0	9

The worst 5 of 69 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	214	ILE	CA-C	-18.32	1.31	1.52
1	A	201	ALA	CA-C	-17.63	1.30	1.52
1	A	169	LYS	CA-CB	-15.88	1.21	1.53
1	A	214	ILE	C-O	-15.04	1.08	1.24
1	A	212	HIS	CA-CB	-14.63	1.30	1.53

The worst 5 of 212 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	740	SER	N-CA-C	16.14	130.07	112.97
1	A	214	ILE	N-CA-C	14.02	128.66	108.48
1	A	207	ALA	N-CA-C	-12.29	99.07	113.21
1	A	853	VAL	N-CA-C	-12.23	98.96	110.82
1	A	593	GLY	CA-C-N	12.04	139.18	121.72

There are no chirality outliers.

5 of 9 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	169	LYS	Peptide
1	A	200	PRO	Mainchain
1	A	209	THR	Peptide
1	A	211	GLN	Peptide
1	A	215	THR	Peptide

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	A	6085	0	6244	718	0
2	B	764	0	427	51	0
3	C	788	0	432	49	0
4	A	159	0	0	88	0
4	B	23	0	0	12	0
4	C	34	0	0	12	0
All	All	7853	0	7103	803	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 803 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)
1:A:454:CYS:HA	1:A:457:HIS:O	1.26	1.27
1:A:729:GLY:O	1:A:730:LEU:HD23	1.14	1.26
1:A:1008:VAL:HG11	4:A:101:HOH:O	1.37	1.21
1:A:684:GLN:HB2	4:A:1175:HOH:O	1.43	1.16
1:A:389:HIS:CE1	1:A:416:LEU:HD23	1.79	1.16

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
1	A	839/1047 (80%)	549 (65%)	246 (29%)	44 (5%)	1 9

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	200	PRO
1	A	303	GLY
1	A	371	GLY
1	A	1034	GLY
1	A	221	THR

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
1	A	635/816 (78%)	465 (73%)	170 (27%)	0 2

5 of 170 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	779	LYS
1	A	921	VAL
1	A	789	LEU
1	A	857	LEU
1	A	939	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 36

such sidechains are listed below:

Mol	Chain	Res	Type
1	A	882	GLN
1	A	1039	GLN
1	A	888	GLN
1	A	950	GLN
1	A	514	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](#)

There are no ligands in this entry.

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	A	854/1047 (81%)	0.45	56 (6%) 24 12	27, 84, 153, 200	0
2	B	38/38 (100%)	0.37	3 (7%) 18 10	31, 46, 132, 137	0
3	C	38/38 (100%)	0.60	1 (2%) 57 34	43, 70, 129, 135	0
All	All	930/1123 (82%)	0.45	60 (6%) 25 13	27, 83, 153, 200	0

The worst 5 of 60 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1040	VAL	5.2
1	A	593	GLY	4.2
1	A	834	PRO	4.2
1	A	1042	ALA	3.9
1	A	830	HIS	3.8

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.