



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 1, 2026 – 10:21 AM UTC

PDB ID : 4OSK / pdb_00004osk
Title : Crystal structure of TAL effector reveals the recognition between asparagine and guanine
Authors : Deng, D.; Wu, J.P.; Yan, C.Y.; Pan, X.J.; Yan, N.
Deposited on : 2014-02-13
Resolution : 2.40 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

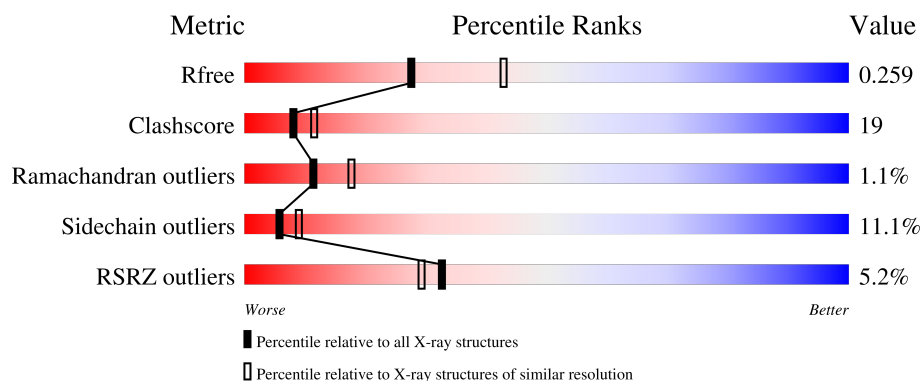
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.40 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	4912 (2.40-2.40)
Clashscore	190562	5391 (2.40-2.40)
Ramachandran outliers	187476	5320 (2.40-2.40)
Sidechain outliers	187428	5321 (2.40-2.40)
RSRZ outliers	180081	4916 (2.40-2.40)

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	499	4% 62% 32% 5%
1	B	499	5% 61% 31% 6%
2	G	17	76% 18% 6%
2	I	17	18% 53% 47%
3	H	17	18% 71% 29%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	J	17	18% 41% 59%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 8734 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 Hax3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	490	Total	C	N	O	S	5	1	0
			3551	2215	659	665	12			
1	B	491	Total	C	N	O	S	5	0	0
			3561	2224	659	665	13			

There are 46 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	230	MET	-	expression tag	UNP Q3ZD72
A	300	HIS	ASN	engineered mutation	UNP Q3ZD72
A	301	ASP	ILE	engineered mutation	UNP Q3ZD72
A	369	SER	ILE	engineered mutation	UNP Q3ZD72
A	402	ASN	HIS	engineered mutation	UNP Q3ZD72
A	403	SER	ASP	engineered mutation	UNP Q3ZD72
A	470	ASN	HIS	engineered mutation	UNP Q3ZD72
A	471	GLY	ASP	engineered mutation	UNP Q3ZD72
A	538	HIS	ASN	engineered mutation	UNP Q3ZD72
A	539	ASP	SER	engineered mutation	UNP Q3ZD72
A	573	GLY	SER	engineered mutation	UNP Q3ZD72
A	606	ASN	HIS	engineered mutation	UNP Q3ZD72
A	607	SER	ASP	engineered mutation	UNP Q3ZD72
A	641	ASN	ILE	engineered mutation	UNP Q3ZD72
A	675	SER	GLY	engineered mutation	UNP Q3ZD72
A	721	LEU	-	expression tag	UNP Q3ZD72
A	722	GLU	-	expression tag	UNP Q3ZD72
A	723	HIS	-	expression tag	UNP Q3ZD72
A	724	HIS	-	expression tag	UNP Q3ZD72
A	725	HIS	-	expression tag	UNP Q3ZD72
A	726	HIS	-	expression tag	UNP Q3ZD72
A	727	HIS	-	expression tag	UNP Q3ZD72
A	728	HIS	-	expression tag	UNP Q3ZD72
B	230	MET	-	expression tag	UNP Q3ZD72
B	300	HIS	ASN	engineered mutation	UNP Q3ZD72

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	301	ASP	ILE	engineered mutation	UNP Q3ZD72
B	369	SER	ILE	engineered mutation	UNP Q3ZD72
B	402	ASN	HIS	engineered mutation	UNP Q3ZD72
B	403	SER	ASP	engineered mutation	UNP Q3ZD72
B	470	ASN	HIS	engineered mutation	UNP Q3ZD72
B	471	GLY	ASP	engineered mutation	UNP Q3ZD72
B	538	HIS	ASN	engineered mutation	UNP Q3ZD72
B	539	ASP	SER	engineered mutation	UNP Q3ZD72
B	573	GLY	SER	engineered mutation	UNP Q3ZD72
B	606	ASN	HIS	engineered mutation	UNP Q3ZD72
B	607	SER	ASP	engineered mutation	UNP Q3ZD72
B	641	ASN	ILE	engineered mutation	UNP Q3ZD72
B	675	SER	GLY	engineered mutation	UNP Q3ZD72
B	721	LEU	-	expression tag	UNP Q3ZD72
B	722	GLU	-	expression tag	UNP Q3ZD72
B	723	HIS	-	expression tag	UNP Q3ZD72
B	724	HIS	-	expression tag	UNP Q3ZD72
B	725	HIS	-	expression tag	UNP Q3ZD72
B	726	HIS	-	expression tag	UNP Q3ZD72
B	727	HIS	-	expression tag	UNP Q3ZD72
B	728	HIS	-	expression tag	UNP Q3ZD72

- Molecule 2 is a DNA chain called DNA (5'-D(*TP*GP*TP*CP*CP*AP*AP*CP*TP*AP*CP*TP*AP*GP*AP*CP*T)-3').

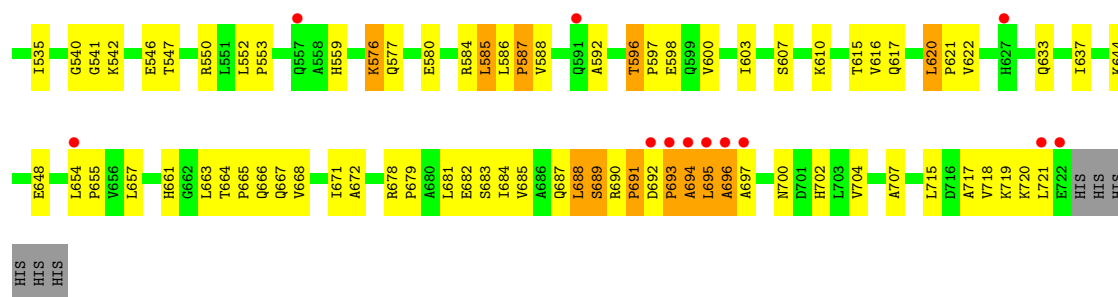
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	I	17	Total	C	N	O	P	0	0	0
			341	165	60	100	16			
2	G	16	Total	C	N	O	P	0	0	0
			321	155	58	93	15			

- Molecule 3 is a DNA chain called DNA (5'-D(*AP*GP*TP*CP*TP*AP*GP*TP*AP*GP*TP*TP*GP*GP*AP*CP*A)-3').

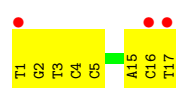
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	J	17	Total	C	N	O	P	0	0	0
			350	168	66	100	16			
3	H	17	Total	C	N	O	P	0	0	0
			350	168	66	100	16			

- Molecule 4 is water.

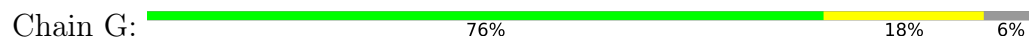
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	99	Total 99	O 99	0	0
4	I	27	Total 27	O 27	0	0
4	J	7	Total 7	O 7	0	0
4	B	90	Total 90	O 90	0	0
4	G	24	Total 24	O 24	0	0
4	H	13	Total 13	O 13	0	0



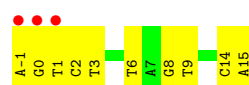
- Molecule 2: DNA (5'-D(*TP*GP*TP*CP*CP*AP*AP*CP*TP*AP*CP*TP*AP*GP*AP*CP*T)-3')



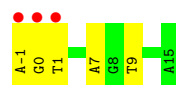
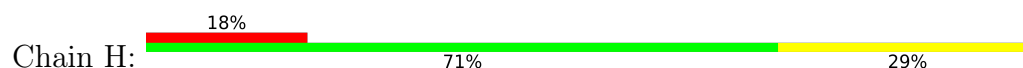
- Molecule 2: DNA (5'-D(*TP*GP*TP*CP*CP*AP*AP*CP*TP*AP*CP*TP*AP*GP*AP*CP*T)-3')



- Molecule 3: DNA (5'-D(*AP*GP*TP*CP*TP*AP*GP*TP*AP*GP*TP*TP*GP*GP*AP*CP*A)-3')



- Molecule 3: DNA (5'-D(*AP*GP*TP*CP*TP*AP*GP*TP*AP*GP*TP*TP*GP*GP*AP*CP*A)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	84.63Å 88.13Å 90.48Å 90.00° 104.96° 90.00°	Depositor
Resolution (Å)	34.97 – 2.40 34.97 – 2.40	Depositor EDS
% Data completeness (in resolution range)	99.6 (34.97-2.40) 99.5 (34.97-2.40)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.20 (at 2.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.1_743)	Depositor
R, R_{free}	0.214 , 0.255 0.220 , 0.259	Depositor DCC
R_{free} test set	2552 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	43.0	Xtriage
Anisotropy	0.220	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 33.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	8734	wwPDB-VP
Average B, all atoms (Å ²)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% 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](#)

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	0.59	0/3599	0.81	1/4915 (0.0%)
1	B	0.56	0/3609	0.83	0/4927
2	G	0.47	0/359	0.98	0/551
2	I	0.36	0/381	0.83	0/585
3	H	0.36	0/393	0.77	0/606
3	J	0.33	0/393	0.71	0/606
All	All	0.54	0/8734	0.82	1/12190 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	585	LEU	N-CA-C	5.16	119.73	113.38

There are no chirality outliers.

There are no planarity outliers.

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	3551	0	3708	145	1
1	B	3561	0	3718	140	2
2	G	321	0	182	3	0
2	I	341	0	194	10	1
3	H	350	0	194	5	0
3	J	350	0	194	25	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	A	99	0	0	10	0
4	B	90	0	0	6	0
4	G	24	0	0	2	0
4	H	13	0	0	0	0
4	I	27	0	0	4	0
4	J	7	0	0	5	0
All	All	8734	0	8190	321	2

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 19.

The worst 5 of 321 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:J:2:DC:OP2	4:J:105:HOH:O	1.60	1.17
1:A:550:ARG:HH12	1:A:551:LEU:CD2	1.62	1.11
1:A:550:ARG:HH12	1:A:551:LEU:CG	1.64	1.09
1:A:550:ARG:NH1	1:A:551:LEU:CD2	2.21	1.03
1:A:550:ARG:NH1	1:A:550:ARG:HG2	1.57	1.02

All (2) 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)
1:A:233:SER:N	1:B:232:TRP:O[1_656]	2.10	0.10
2:I:1:DT:O5'	1:B:233:SER:OG[1_656]	2.15	0.05

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	
1	A	489/499 (98%)	454 (93%)	33 (7%)	2 (0%)	30	43
1	B	487/499 (98%)	442 (91%)	36 (7%)	9 (2%)	6	9
All	All	976/998 (98%)	896 (92%)	69 (7%)	11 (1%)	11	18

5 of 11 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	691	PRO
1	B	694	ALA
1	A	553	PRO
1	B	486	VAL
1	B	693	PRO

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	379/387 (98%)	332 (88%)	47 (12%)	4	6
1	B	380/387 (98%)	343 (90%)	37 (10%)	8	12
All	All	759/774 (98%)	675 (89%)	84 (11%)	6	9

5 of 84 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	347	LEU
1	B	576	LYS
1	B	351	LEU
1	B	455	GLN
1	B	607	SER

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	334	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	436	HIS
1	B	687	GLN
1	B	421	GLN
1	B	457	HIS

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	490/499 (98%)	0.48	21 (4%) 40 36	16, 51, 87, 128	10 (2%)
1	B	491/499 (98%)	0.43	24 (4%) 35 31	24, 48, 86, 114	7 (1%)
2	G	16/17 (94%)	-0.75	0 100 100	27, 33, 51, 70	0
2	I	17/17 (100%)	-0.09	3 (17%) 4 3	32, 40, 101, 146	0
3	H	17/17 (100%)	0.41	3 (17%) 4 3	20, 50, 86, 130	0
3	J	17/17 (100%)	1.04	3 (17%) 4 3	20, 51, 117, 134	0
All	All	1048/1066 (98%)	0.44	54 (5%) 33 29	16, 49, 87, 146	17 (1%)

The worst 5 of 54 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	233	SER	5.2
3	J	1	DT	4.7
1	B	694	ALA	4.0
1	B	722	GLU	3.8
1	A	524	ALA	3.7

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.