



Full wwPDB X-ray Structure Validation Report ⓘ

Mar 5, 2026 – 02:11 AM UTC

PDB ID : 5GHF / pdb_00005ghf
Title : Transaminase with L-ala
Authors : Kim, J.; Park, J.
Deposited on : 2016-06-19
Resolution : 2.00 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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

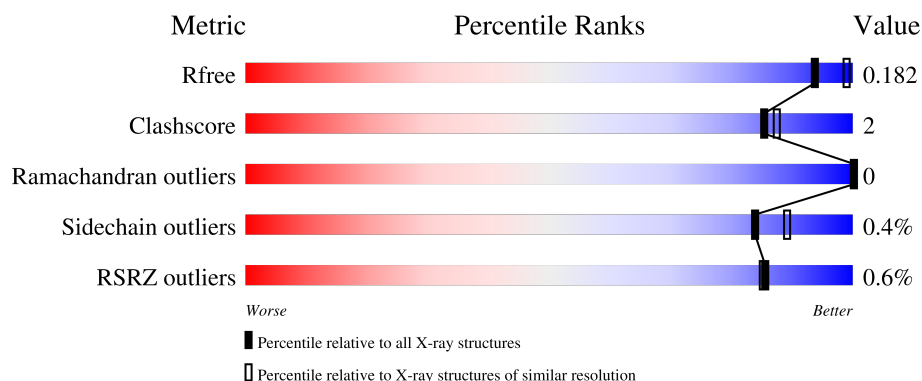
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	10052 (2.00-2.00)
Clashscore	190562	11152 (2.00-2.00)
Ramachandran outliers	187476	11031 (2.00-2.00)
Sidechain outliers	187428	11029 (2.00-2.00)
RSRZ outliers	180081	10067 (2.00-2.00)

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	456	 89% 5% 5%
1	B	456	 86% 7% 7%

2 Entry composition [i](#)

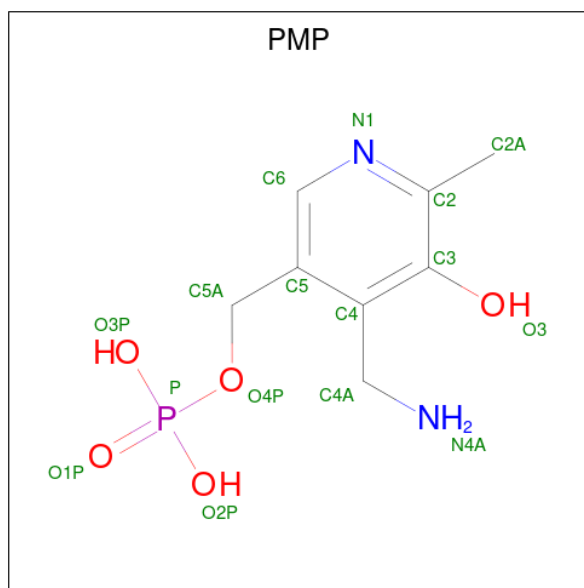
There are 3 unique types of molecules in this entry. The entry contains 7526 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Aminotransferase class-III.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	431	Total	C	N	O	S	0	0	0
			3311	2093	581	624	13			
1	B	424	Total	C	N	O	S	0	0	0
			3243	2054	564	612	13			

- Molecule 2 is 4'-DEOXY-4'-AMINOPYRIDOXAL-5'-PHOSPHATE (CCD ID: PMP) (formula: $C_8H_{13}N_2O_5P$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	N	O	P	0	0
			16	8	2	5	1		

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	516	Total	O	0	0
			516	516		

Continued on next page...

Continued from previous page...

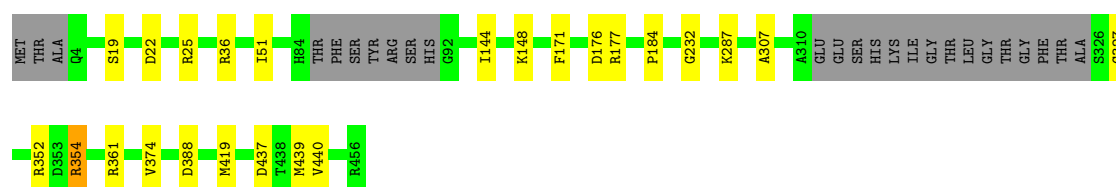
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	440	Total	O	0	0
			440	440		

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.

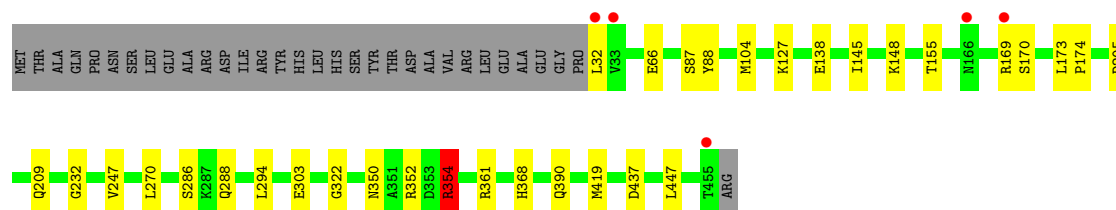
• Molecule 1: Aminotransferase class-III

Chain A: 



• Molecule 1: Aminotransferase class-III

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	78.80Å 102.00Å 116.73Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	27.53 – 2.00 27.53 – 2.00	Depositor EDS
% Data completeness (in resolution range)	96.5 (27.53-2.00) 96.4 (27.53-2.00)	Depositor EDS
R_{merge}	0.15	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	7.23 (at 1.99Å)	Xtriage
Refinement program	REFMAC 5.8.0135	Depositor
R, R_{free}	0.135 , 0.174 0.148 , 0.182	Depositor DCC
R_{free} test set	3136 reflections (4.88%)	wwPDB-VP
Wilson B-factor (Å ²)	9.6	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 47.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	7526	wwPDB-VP
Average B, all atoms (Å ²)	11.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.42% 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: PMP

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	1.33	6/3373 (0.2%)	1.08	2/4571 (0.0%)
1	B	1.34	6/3306 (0.2%)	1.11	6/4481 (0.1%)
All	All	1.34	12/6679 (0.2%)	1.10	8/9052 (0.1%)

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	286	SER	C-O	7.29	1.29	1.24
1	B	145	ILE	C-O	-6.41	1.17	1.24
1	B	247	VAL	N-CA	-6.20	1.39	1.46
1	B	138	GLU	C-O	-6.13	1.17	1.24
1	A	144	ILE	C-O	-5.84	1.17	1.24
1	A	307	ALA	N-CA	5.83	1.52	1.46
1	A	374	VAL	C-O	-5.76	1.17	1.24
1	B	148	LYS	C-O	-5.47	1.17	1.23
1	B	288	GLN	C-O	-5.42	1.17	1.24
1	A	388	ASP	C-O	5.17	1.29	1.24
1	A	184	PRO	N-CA	-5.16	1.43	1.47
1	A	327	GLY	N-CA	5.02	1.51	1.45

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	354	ARG	CB-CG-CD	8.32	130.45	111.30
1	B	354	ARG	CB-CG-CD	8.00	129.69	111.30
1	B	155	THR	N-CA-C	-5.72	103.16	110.53
1	B	294	LEU	CA-C-N	-5.72	113.20	120.23
1	B	294	LEU	C-N-CA	-5.72	113.20	120.23
1	B	66	GLU	CA-C-N	-5.24	113.64	119.19
1	B	66	GLU	C-N-CA	-5.24	113.64	119.19

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	354	ARG	CG-CD-NE	5.16	123.36	112.00

There are no chirality outliers.

There are no planarity outliers.

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	A	3311	0	3328	15	0
1	B	3243	0	3262	16	0
2	A	16	0	10	2	0
3	A	516	0	0	4	1
3	B	440	0	0	4	0
All	All	7526	0	6600	29	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 2.

All (29) 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:439:MET:SD	3:A:1046:HOH:O	2.21	0.99
1:A:232:GLY:HA2	1:A:419:MET:HE2	1.45	0.97
1:A:25:ARG:NH2	3:A:601:HOH:O	2.06	0.87
1:A:361:ARG:NH1	1:A:437:ASP:OD1	2.17	0.76
1:B:232:GLY:HA2	1:B:419:MET:HE2	1.75	0.68
1:B:205:ASP:OD2	1:B:209:GLN:NE2	2.30	0.64
1:A:148:LYS:NZ	3:A:604:HOH:O	2.32	0.63
1:B:303:GLU:OE2	3:B:501:HOH:O	2.18	0.56
1:B:350:ASN:O	1:B:354:ARG:HG2	2.06	0.56
1:B:352:ARG:HD3	3:B:586:HOH:O	2.08	0.54
1:A:287:LYS:HE2	2:A:501:PMP:H4A2	1.91	0.53
1:B:361:ARG:NH1	1:B:437:ASP:OD1	2.27	0.52
1:A:352:ARG:NH2	3:A:602:HOH:O	2.24	0.51
1:A:232:GLY:HA2	1:A:419:MET:CE	2.31	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:176:ASP:O	1:A:177:ARG:HB2	2.15	0.47
1:B:104:MET:HE2	1:B:270:LEU:HD21	1.96	0.46
1:A:19:SER:HB3	1:B:322:GLY:HA3	1.98	0.46
1:A:171:PHE:CD1	1:B:127:LYS:HG2	2.51	0.46
1:A:287:LYS:CE	2:A:501:PMP:H4A2	2.46	0.46
1:A:51:ILE:HG21	1:A:439:MET:HE2	1.97	0.45
1:B:390:GLN:NE2	3:B:518:HOH:O	2.48	0.45
1:A:361:ARG:HB3	1:A:440:VAL:HG11	2.00	0.44
1:B:32:LEU:N	3:B:520:HOH:O	2.51	0.44
1:A:22:ASP:C	1:A:22:ASP:OD1	2.61	0.43
1:B:232:GLY:HA2	1:B:419:MET:CE	2.48	0.42
1:B:173:LEU:HB3	1:B:174:PRO:HA	2.02	0.41
1:B:368:HIS:CE1	1:B:447:LEU:HB3	2.56	0.41
1:B:169:ARG:O	1:B:170:SER:HB2	2.22	0.40
1:B:87:SER:C	1:B:88:TYR:CD2	2.99	0.40

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)
3:A:621:HOH:O	3:A:671:HOH:O[4_545]	2.18	0.02

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	425/456 (93%)	408 (96%)	17 (4%)	0	100	100
1	B	422/456 (92%)	409 (97%)	13 (3%)	0	100	100
All	All	847/912 (93%)	817 (96%)	30 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	
1	A	350/370 (95%)	348 (99%)	2 (1%)	78	85
1	B	343/370 (93%)	342 (100%)	1 (0%)	86	91
All	All	693/740 (94%)	690 (100%)	3 (0%)	84	89

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	36	ARG
1	A	354	ARG
1	B	354	ARG

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	338	ASN
1	B	338	ASN
1	B	390	GLN
1	B	410	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

1 ligand is 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	PMP	A	501	-	16,16,16	2.64	3 (18%)	22,23,23	1.64	5 (22%)

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	PMP	A	501	-	-	2/8/8/8	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	501	PMP	C3-C4	5.79	1.48	1.40
2	A	501	PMP	C3-C2	5.76	1.47	1.41
2	A	501	PMP	C5-C4	5.70	1.48	1.40

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	501	PMP	O4P-C5A-C5	3.44	115.81	109.36
2	A	501	PMP	O2P-P-O4P	-2.93	99.04	106.67
2	A	501	PMP	O3-C3-C2	2.69	123.15	117.58
2	A	501	PMP	C4-C3-C2	-2.33	116.40	119.91
2	A	501	PMP	O3P-P-O2P	2.25	116.24	107.80

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	501	PMP	C3-C4-C4A-N4A
2	A	501	PMP	C5-C4-C4A-N4A

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	501	PMP	2	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	A	431/456 (94%)	-0.75	0 100 100	3, 7, 20, 43	0
1	B	424/456 (92%)	-0.64	5 (1%) 76 76	2, 8, 24, 59	0
All	All	855/912 (93%)	-0.70	5 (0%) 85 85	2, 8, 21, 59	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	32	LEU	3.1
1	B	455	THR	3.0
1	B	169	ARG	2.7
1	B	166	ASN	2.4
1	B	33	VAL	2.2

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
2	PMP	A	501	16/16	0.98	0.06	9,12,21,28	0

6.5 Other polymers [i](#)

There are no such residues in this entry.