



wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 8, 2026 – 01:10 PM UTC

PDB ID : 5BP5 / pdb_00005bp5
Title : Crystal structure of HA17-HA33-IPT
Authors : Lee, K.; Lam, K.; Jin, R.
Deposited on : 2015-05-27
Resolution : 2.18 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

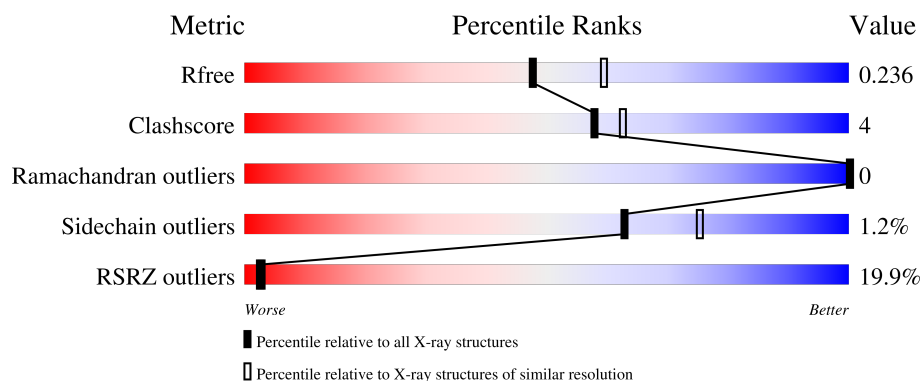
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.18 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	8975 (2.20-2.16)
Clashscore	190562	9786 (2.20-2.16)
Ramachandran outliers	187476	9664 (2.20-2.16)
Sidechain outliers	187428	9664 (2.20-2.16)
RSRZ outliers	180081	8979 (2.20-2.16)

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	296	8% 90% 6% ..
1	B	296	38% 86% 10% ..
2	C	147	5% 83% 14% ..

2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 6043 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 HA-33.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	286	Total	C	N	O	S	4	0	0
			2333	1476	399	455	3			
1	B	285	Total	C	N	O	S	70	0	0
			2325	1472	397	453	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	294	PRO	-	expression tag	UNP Q45871
A	295	GLY	-	expression tag	UNP Q45871
A	296	SER	-	expression tag	UNP Q45871
A	297	ALA	-	expression tag	UNP Q45871
B	294	PRO	-	expression tag	UNP Q45871
B	295	GLY	-	expression tag	UNP Q45871
B	296	SER	-	expression tag	UNP Q45871
B	297	ALA	-	expression tag	UNP Q45871

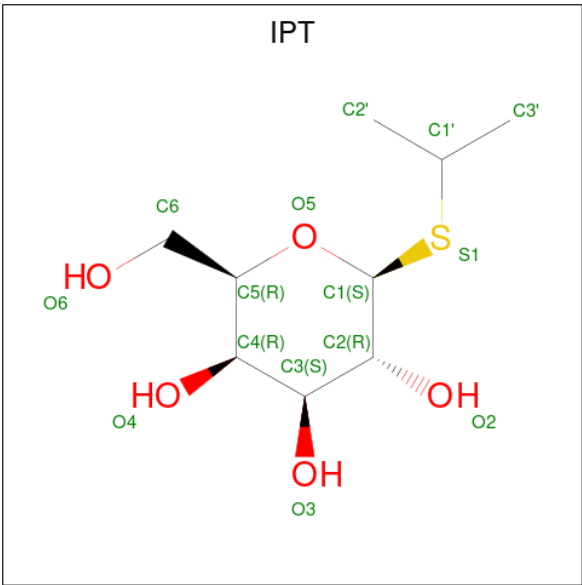
- Molecule 2 is a protein called HA-17.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	144	Total	C	N	O	S	0	0	0
			1184	764	189	226	5			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-1	GLY	-	expression tag	UNP Q45878
C	0	PRO	-	expression tag	UNP Q45878

- Molecule 3 is 1-methylethyl 1-thio-beta-D-galactopyranoside (CCD ID: IPT) (formula: C₉H₁₈O₅S).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	A	1	Total	C	O	S	0	0
			15	9	5	1		
3	B	1	Total	C	O	S	0	0
			15	9	5	1		

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	68	Total	O	0	0
			68	68		
4	C	38	Total	O	0	0
			38	38		
4	B	65	Total	O	0	0
			65	65		

- Molecule 1: HA-33



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	105.87Å 118.93Å 162.20Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.95 – 2.18 47.95 – 2.18	Depositor EDS
% Data completeness (in resolution range)	99.4 (47.95-2.18) 99.4 (47.95-2.18)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.39 (at 2.18Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.227 , 0.238 (Not available) , 0.236	Depositor DCC
R_{free} test set	2711 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	44.9	Xtriage
Anisotropy	0.143	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.35 , 35.3	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	6043	wwPDB-VP
Average B, all atoms (Å ²)	54.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.77% 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: IPT

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	0/2384	0.71	0/3255
1	B	0.54	0/2376	0.71	0/3244
2	C	0.58	0/1214	0.76	0/1649
All	All	0.57	0/5974	0.72	0/8148

There are no bond length outliers.

There are no bond angle outliers.

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	2333	0	2250	12	0
1	B	2325	0	2244	24	0
2	C	1184	0	1141	15	0
3	A	15	0	18	0	0
3	B	15	0	18	0	0
4	A	68	0	0	4	0
4	B	65	0	0	8	0
4	C	38	0	0	1	0
All	All	6043	0	5671	51	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 4.

The worst 5 of 51 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:B:175:LEU:HD12	1:B:225:VAL:HG22	1.48	0.93
1:A:246:THR:HG21	1:A:249:THR:OG1	1.80	0.81
2:C:7:LEU:HD11	2:C:46:VAL:HG11	1.75	0.69
1:B:244:SER:O	1:B:245:ASP:HB2	1.92	0.69
1:B:232:ASN:N	4:B:401:HOH:O	2.09	0.68

There are no symmetry-related clashes.

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	284/296 (96%)	275 (97%)	9 (3%)	0	100	100
1	B	283/296 (96%)	275 (97%)	8 (3%)	0	100	100
2	C	142/147 (97%)	141 (99%)	1 (1%)	0	100	100
All	All	709/739 (96%)	691 (98%)	18 (2%)	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	262/270 (97%)	259 (99%)	3 (1%)	65	77
1	B	261/270 (97%)	258 (99%)	3 (1%)	65	77
2	C	135/138 (98%)	133 (98%)	2 (2%)	57	70
All	All	658/678 (97%)	650 (99%)	8 (1%)	63	75

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	225	VAL
1	B	213	LEU
2	C	60	VAL
2	C	25	ASN
1	B	89	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
2	C	138	GLN
1	B	202	GLN
1	B	292	ASN
1	A	202	GLN
1	A	169	GLN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates ⓘ

There are no oligosaccharides in this entry.

5.6 Ligand geometry ⓘ

2 ligands 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	$\# Z > 2$
3	IPT	A	301	-	15,15,15	0.91	1 (6%)	18,21,21	0.63	0
3	IPT	B	301	-	15,15,15	0.70	0	18,21,21	0.48	0

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
3	IPT	A	301	-	-	1/6/26/26	0/1/1/1
3	IPT	B	301	-	-	1/6/26/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	301	IPT	O5-C1	2.09	1.45	1.42

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	301	IPT	O5-C5-C6-O6
3	A	301	IPT	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues ⓘ

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	286/296 (96%)	0.69	23 (8%) 18 17	28, 48, 72, 95	2 (0%)
1	B	284/296 (95%)	1.85	112 (39%) 1 0	11, 54, 101, 118	38 (13%)
2	C	144/147 (97%)	0.62	7 (4%) 35 34	29, 47, 64, 75	0
All	All	714/739 (96%)	1.14	142 (19%) 3 2	11, 49, 92, 118	40 (5%)

The worst 5 of 142 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	276	GLN	15.1
1	B	261	ALA	12.6
1	B	265	TYR	8.6
1	B	277	VAL	8.1
1	B	264	LEU	8.1

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
3	IPT	B	301	15/15	0.73	0.22	30,33,34,34	15
3	IPT	A	301	15/15	0.86	0.14	28,31,32,34	0

6.5 Other polymers [i](#)

There are no such residues in this entry.