



Full wwPDB X-ray Structure Validation Report ⓘ

Apr 27, 2026 – 05:52 PM EDT

PDB ID : 8AYG / pdb_00008ayg
Title : Crystal structure of an intramolecular i-motif at the insulin-linked polymorphic region (ILPR)
Authors : Parkinson, G.N.; Alexandrou, E.; Waller, Z.A.E.; El-Omari, K.
Deposited on : 2022-09-02
Resolution : 2.25 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
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

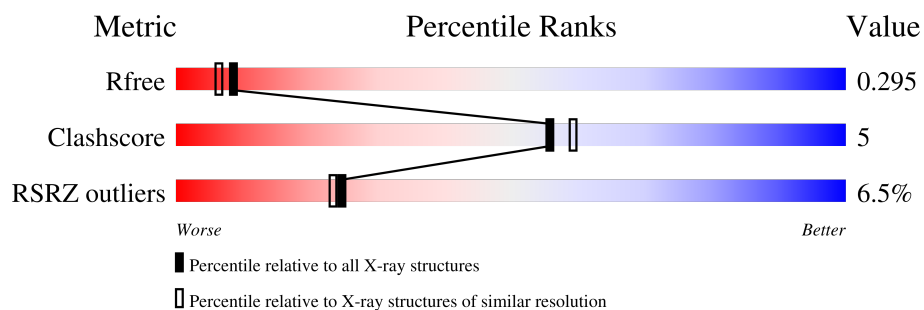
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.25 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	1898 (2.26-2.26)
Clashscore	190562	2005 (2.26-2.26)
RSRZ outliers	180081	1898 (2.26-2.26)

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 1234 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 Insulin-linked polymorphic region, ILPR DNA (31-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	31	Total	C	N	O	P	0	0	0
			606	292	101	183	30			
1	B	31	Total	C	N	O	P	0	0	0
			606	292	101	183	30			

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	10	Total	O	0	0
			10	10		
2	B	12	Total	O	0	0
			12	12		

SEQUENCE-PLOTS INFOmissingINFO

3 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	47.41Å 51.51Å 69.85Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	41.46 – 2.25 41.46 – 2.25	Depositor EDS
% Data completeness (in resolution range)	93.1 (41.46-2.25) 94.8 (41.46-2.25)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.47 (at 2.24Å)	Xtriage
Refinement program	PHENIX 1.20.1_4487	Depositor
R, R_{free}	0.249 , 0.295 0.248 , 0.295	Depositor DCC
R_{free} test set	430 reflections (4.28%)	wwPDB-VP
Wilson B-factor (Å ²)	79.4	Xtriage
Anisotropy	0.716	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 59.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.51$, $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	1234	wwPDB-VP
Average B, all atoms (Å ²)	80.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.41% 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.

4 Model quality [i](#)

4.1 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	0.52	0/674	0.56	0/1031
1	B	0.59	0/674	0.61	0/1031
All	All	0.55	0/1348	0.59	0/2062

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

4.2 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	606	0	349	3	0
1	B	606	0	349	6	0
2	A	10	0	0	0	0
2	B	12	0	0	0	0
All	All	1234	0	698	9	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 5.

All (9) 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:29:DT:H2''	1:B:30:DA:H5''	1.53	0.88

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:29:DT:H2'	1:B:29:DT:O2	1.99	0.63
1:B:16:DA:H4'	1:B:17:DT:OP2	2.13	0.48
1:B:16:DA:N3	1:B:16:DA:H5''	2.29	0.47
1:B:30:DA:H2''	1:B:31:DT:H72	1.98	0.45
1:A:6:DC:H4'	1:A:7:DC:OP1	2.18	0.44
1:B:6:DC:N3	1:B:20:DC:N3	2.66	0.44
1:A:12:DC:N3	1:A:26:DC:N3	2.68	0.41
1:A:6:DC:H2''	1:A:7:DC:O5'	2.21	0.41

There are no symmetry-related clashes.

4.3 Torsion angles [i](#)

4.3.1 Protein backbone [i](#)

There are no protein molecules in this entry.

4.3.2 Protein sidechains [i](#)

There are no protein molecules in this entry.

4.3.3 RNA [i](#)

There are no RNA molecules in this entry.

4.4 Non-standard residues in protein, DNA, RNA chains [i](#)

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

4.5 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

4.6 Ligand geometry [i](#)

There are no ligands in this entry.

4.7 Other polymers [i](#)

There are no such residues in this entry.

4.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

5 Fit of model and data [i](#)

5.1 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	31/31 (100%)	0.50	1 (3%) 50 50	61, 70, 102, 128	0
1	B	31/31 (100%)	0.85	3 (9%) 13 12	67, 76, 114, 139	0
All	All	62/62 (100%)	0.68	4 (6%) 25 23	61, 74, 114, 139	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	31	DT	3.2
1	B	31	DT	2.5
1	B	30	DA	2.4
1	B	8	DA	2.2

5.2 Non-standard residues in protein, DNA, RNA chains [i](#)

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

5.3 Carbohydrates [i](#)

There are no oligosaccharides in this entry.

5.4 Ligands [i](#)

There are no ligands in this entry.

5.5 Other polymers [i](#)

There are no such residues in this entry.