



Full wwPDB X-ray Structure Validation Report ⓘ

Dec 18, 2023 – 01:17 PM EST

PDB ID : 8UGC
Title : FD15: Flat repeat helix-turn-helix-turn protein
Authors : Davila-Hernandez, F.; Bera, A.K.; Kang, A.; Baker, D.
Deposited on : 2023-10-05
Resolution : 4.00 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 : 4.02b-467
Xtrriage (Phenix) : 1.13
EDS : 2.36
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.36

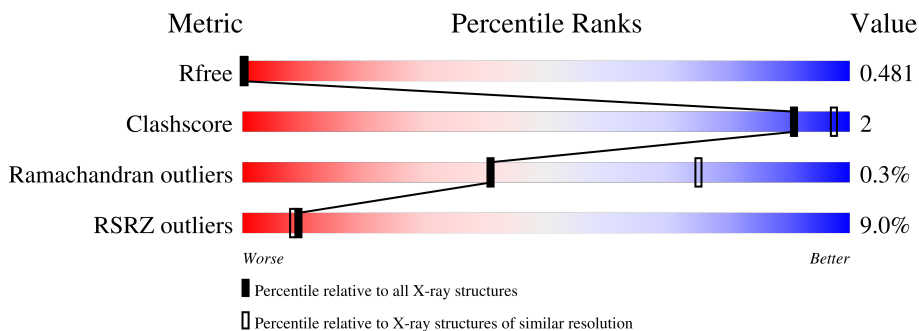
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.00 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}	130704	1087 (4.30-3.70)
Clashscore	141614	1148 (4.30-3.70)
Ramachandran outliers	138981	1108 (4.30-3.70)
RSRZ outliers	127900	1028 (4.34-3.66)

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	393	 7% 95%
1	B	393	 11% 95%

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 3816 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.

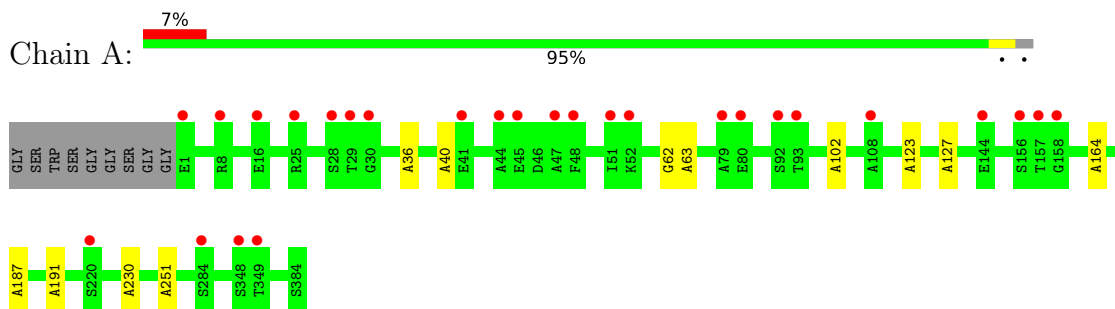
- Molecule 1 is a protein called FD15.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
1	A	384	1908	1140	384	384	0	0	0
1	B	384	1908	1140	384	384	0	0	0

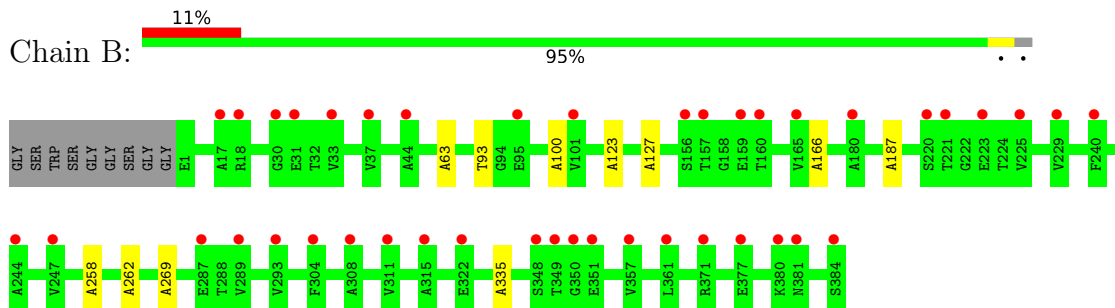
3 Residue-property plots [i](#)

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: FD15



- Molecule 1: FD15



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	65.67Å 122.78Å 90.00Å 90.00° 97.61° 90.00°	Depositor
Resolution (Å)	89.20 – 4.00 89.20 – 4.00	Depositor EDS
% Data completeness (in resolution range)	93.2 (89.20-4.00) 97.6 (89.20-4.00)	Depositor EDS
R_{merge}	0.45	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.54 (at 4.01Å)	Xtrriage
Refinement program	PHENIX dev_4761, PHENIX dev_4761	Depositor
R, R_{free}	0.465 , 0.492 0.466 , 0.481	Depositor DCC
R_{free} test set	560 reflections (4.76%)	wwPDB-VP
Wilson B-factor (Å ²)	108.0	Xtrriage
Anisotropy	0.801	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.30 , 306.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.40$, $\langle L^2 \rangle = 0.23$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.70	EDS
Total number of atoms	3816	wwPDB-VP
Average B, all atoms (Å ²)	103.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 28.01 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.9872e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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	0.24	0/1907	0.34	0/2661
1	B	0.24	0/1907	0.35	0/2661
All	All	0.24	0/3814	0.34	0/5322

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 [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	1908	0	1256	6	0
1	B	1908	0	1256	5	0
All	All	3816	0	2512	11	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 2.

All (11) 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:63:ALA:HB2	1:B:123:ALA:HB1	1.76	0.68
1:A:63:ALA:HB2	1:A:123:ALA:HB1	1.78	0.66

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:127:ALA:HB2	1:A:187:ALA:HB1	1.77	0.66
1:B:127:ALA:HB2	1:B:187:ALA:HB1	1.87	0.55
1:A:40:ALA:HB2	1:A:102:ALA:HA	1.96	0.47
1:A:191:ALA:HB2	1:A:251:ALA:HB1	1.97	0.46
1:B:269:ALA:HB1	1:B:335:ALA:HB2	2.00	0.43
1:B:258:ALA:O	1:B:262:ALA:N	2.50	0.42
1:B:100:ALA:HB1	1:B:166:ALA:HB2	2.01	0.42
1:A:164:ALA:HB1	1:A:230:ALA:HB2	2.01	0.42
1:A:36:ALA:HB1	1:A:102:ALA:HB2	2.01	0.41

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	382/393 (97%)	366 (96%)	15 (4%)	1 (0%)	41	75
1	B	382/393 (97%)	371 (97%)	10 (3%)	1 (0%)	41	75
All	All	764/786 (97%)	737 (96%)	25 (3%)	2 (0%)	41	75

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	62	GLY
1	B	93	THR

5.3.2 Protein sidechains [i](#)

There are no protein residues with a non-rotameric sidechain to report in this entry.

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 monosaccharides 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

6.1 Protein, DNA and RNA chains

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	384/393 (97%)	0.23	27 (7%) 16 13	56, 104, 136, 153	0
1	B	384/393 (97%)	0.37	42 (10%) 5 5	55, 103, 135, 177	0
All	All	768/786 (97%)	0.30	69 (8%) 9 8	55, 103, 137, 177	0

All (69) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	29	THR	7.4
1	A	44	ALA	7.0
1	A	93	THR	4.8
1	A	348	SER	4.8
1	B	18	ARG	3.9
1	A	28	SER	3.8
1	B	31	GLU	3.6
1	B	311	VAL	3.6
1	A	1	GLU	3.4
1	A	284	SER	3.4
1	B	351	GLU	3.3
1	A	45	GLU	3.2
1	B	348	SER	3.1
1	A	108	ALA	3.1
1	A	79	ALA	3.1
1	B	240	PHE	3.1
1	B	244	ALA	3.1
1	A	30	GLY	3.0
1	B	17	ALA	3.0
1	A	52	LYS	3.0
1	A	16	GLU	2.9
1	B	304	PHE	2.9
1	B	159	GLU	2.9
1	B	289	VAL	2.8

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	220	SER	2.7
1	B	357	VAL	2.7
1	B	380	LYS	2.6
1	A	220	SER	2.6
1	B	287	GLU	2.6
1	B	95	GLU	2.5
1	A	41	GLU	2.5
1	A	80	GLU	2.5
1	B	221	THR	2.5
1	B	350	GLY	2.5
1	B	384	SER	2.5
1	B	37	VAL	2.4
1	A	156	SER	2.4
1	B	293	VAL	2.4
1	B	223	GLU	2.4
1	B	101	VAL	2.3
1	B	349	THR	2.3
1	B	377	GLU	2.3
1	B	156	SER	2.3
1	A	48	PHE	2.3
1	B	229	VAL	2.3
1	B	308	ALA	2.3
1	B	30	GLY	2.3
1	A	92	SER	2.3
1	A	51	ILE	2.2
1	B	247	VAL	2.2
1	B	44	ALA	2.2
1	A	8	ARG	2.2
1	A	47	ALA	2.2
1	A	25	ARG	2.2
1	B	322	GLU	2.2
1	B	33	VAL	2.1
1	B	381	ASN	2.1
1	A	144	GLU	2.1
1	B	315	ALA	2.1
1	B	157	THR	2.1
1	B	371	ARG	2.1
1	A	157	THR	2.1
1	B	180	ALA	2.1
1	B	361	LEU	2.0
1	A	349	THR	2.0
1	A	158	GLY	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	165	VAL	2.0
1	B	225	VAL	2.0
1	B	160	THR	2.0

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 monosaccharides 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.