wwPDB X-ray Structure Validation Summary Report (i)

Sep 4, 2023 – 11:51 PM EDT

PDB ID	:	3U8Z
Title	:	human merlin FERM domain
Authors	:	Yogesha, S.D.; Sharff, A.J.; Giovannini, M.; Bricogne, G.; Izard, T.
Deposited on	:	2011-10-17
Resolution	:	2.64 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 (i) 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 (1)) were used in the production of this report:

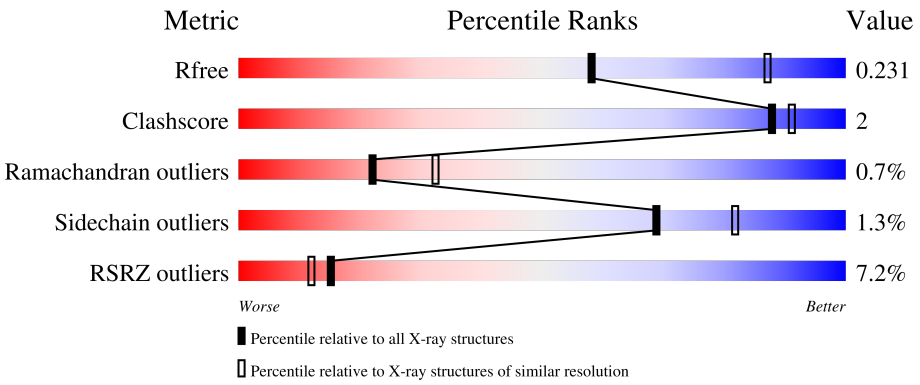
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.64 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	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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 <=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	300	3% 80% 7% 13%
1	B	300	8% 80% 6% 13%
1	C	300	3% 82% 6% • 11%
1	D	300	11% 74% 5% 21%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 9077 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 Merlin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	260	2173	1411	362	387	13	0	0	0
1	B	260	2172	1411	361	387	13	0	0	0
1	C	266	2222	1443	369	397	13	0	0	0
1	D	237	1973	1288	324	350	11	0	0	0

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	13	GLY	-	expression tag	UNP P35240
A	14	PRO	-	expression tag	UNP P35240
A	15	LEU	-	expression tag	UNP P35240
A	16	GLY	-	expression tag	UNP P35240
A	17	SER	-	expression tag	UNP P35240
B	13	GLY	-	expression tag	UNP P35240
B	14	PRO	-	expression tag	UNP P35240
B	15	LEU	-	expression tag	UNP P35240
B	16	GLY	-	expression tag	UNP P35240
B	17	SER	-	expression tag	UNP P35240
C	13	GLY	-	expression tag	UNP P35240
C	14	PRO	-	expression tag	UNP P35240
C	15	LEU	-	expression tag	UNP P35240
C	16	GLY	-	expression tag	UNP P35240
C	17	SER	-	expression tag	UNP P35240
D	13	GLY	-	expression tag	UNP P35240
D	14	PRO	-	expression tag	UNP P35240
D	15	LEU	-	expression tag	UNP P35240
D	16	GLY	-	expression tag	UNP P35240
D	17	SER	-	expression tag	UNP P35240

- Molecule 2 is water.

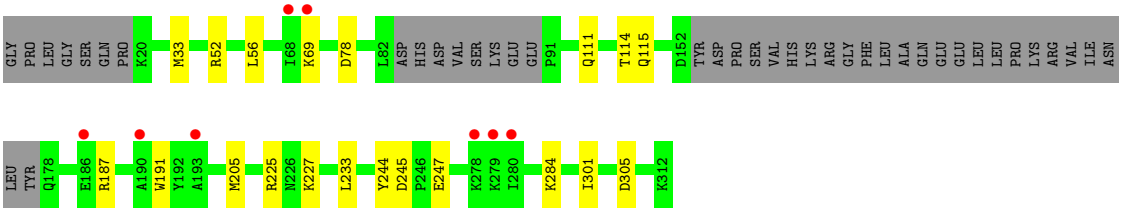
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	164	164	164	0	0
			Total	O		
2	B	141	141	141	0	0
			Total	O		
2	C	162	162	162	0	0
			Total	O		
2	D	70	70	70	0	0

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.

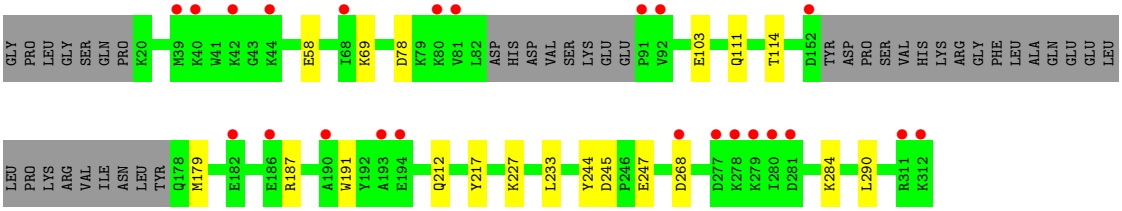
- Molecule 1: Merlin

Chain A:



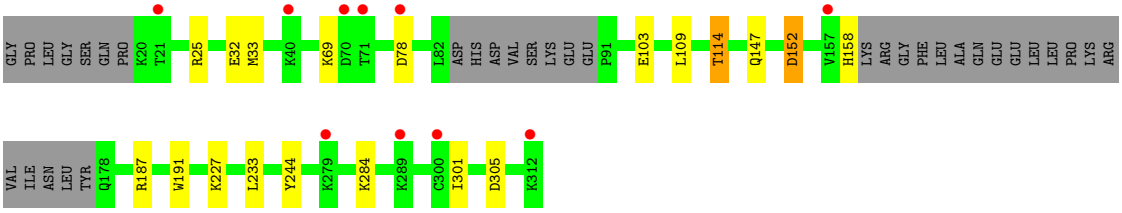
- Molecule 1: Merlin

Chain B:



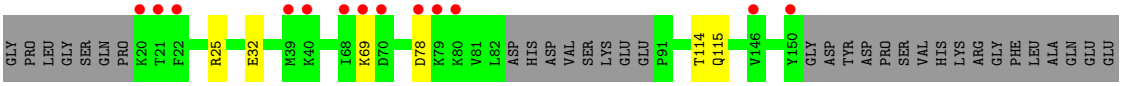
- Molecule 1: Merlin

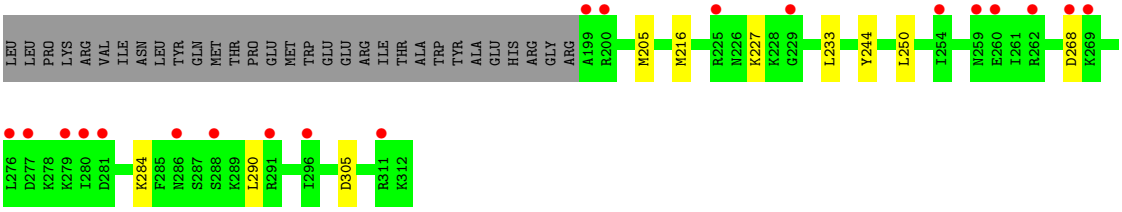
Chain C:



- Molecule 1: Merlin

Chain D:




LEU LEU PRO LYS ARG VAL ILE ASN LEU TYR GLN MET THR PRO GLU MET TRP GLU GLU ARG ILE THR ALA TRP TYR ALA GLU HIS ARG GLY ARG
A199 R200 M205 M216 R225 N226 K227 K228 G229 L233 Y244 L250 I254 N259 E260 I261 R262 D268 K269
L276 D277 K278 K279 I280 D281 K284 F285 N286 S287 S288 K289 L290 R291 I296 D305 R311 K312

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	105.45Å 105.45Å 330.00Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.36 – 2.64 26.28 – 2.64	Depositor EDS
% Data completeness (in resolution range)	94.1 (26.36-2.64) 94.2 (26.28-2.64)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 2.64Å)	Xtriage
Refinement program	BUSTER 2.13.0	Depositor
R, R_{free}	0.200 , 0.229 0.204 , 0.231	Depositor DCC
R_{free} test set	2666 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å^2)	47.9	Xtriage
Anisotropy	0.114	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 75.9	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	9077	wwPDB-VP
Average B, all atoms (Å^2)	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.44 % 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 3.1713e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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	0.43	0/2223	0.54	0/2992
1	B	0.41	0/2222	0.53	0/2990
1	C	0.43	0/2275	0.56	0/3064
1	D	0.40	0/2016	0.54	0/2711
All	All	0.42	0/8736	0.54	0/11757

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	2173	0	2185	16	0
1	B	2172	0	2184	9	0
1	C	2222	0	2225	11	0
1	D	1973	0	2007	7	0
2	A	164	0	0	2	0
2	B	141	0	0	1	0
2	C	162	0	0	0	0
2	D	70	0	0	0	0
All	All	9077	0	8601	34	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.

The worst 5 of 34 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:56:LEU:CD1	1:A:114:THR:HG22	2.02	0.90
1:A:52:ARG:HD3	1:C:152:ASP:OD2	1.81	0.80
1:B:58:GLU:OE1	1:B:114:THR:HG21	1.89	0.72
1:C:25:ARG:HE	1:C:32:GLU:HG3	1.57	0.69
1:A:56:LEU:HD12	1:A:114:THR:HG22	1.74	0.67

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	254/300 (85%)	249 (98%)	4 (2%)	1 (0%)	34	48
1	B	254/300 (85%)	249 (98%)	3 (1%)	2 (1%)	19	28
1	C	260/300 (87%)	250 (96%)	8 (3%)	2 (1%)	19	28
1	D	231/300 (77%)	228 (99%)	1 (0%)	2 (1%)	17	26
All	All	999/1200 (83%)	976 (98%)	16 (2%)	7 (1%)	22	32

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	152	ASP
1	D	69	LYS
1	A	69	LYS
1	B	69	LYS
1	C	69	LYS

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	235/271 (87%)	234 (100%)	1 (0%)	91	95
1	B	234/271 (86%)	231 (99%)	3 (1%)	69	82
1	C	240/271 (89%)	235 (98%)	5 (2%)	53	71
1	D	215/271 (79%)	212 (99%)	3 (1%)	67	80
All	All	924/1084 (85%)	912 (99%)	12 (1%)	69	82

5 of 12 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	147	GLN
1	C	158	HIS
1	D	290	LEU
1	D	78	ASP
1	B	290	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	147	GLN
1	C	195	HIS
1	D	147	GLN
1	D	115	GLN
1	B	147	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 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 (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	260/300 (86%)	-0.14	8 (3%) 49 45	28, 43, 85, 110	0
1	B	260/300 (86%)	0.22	23 (8%) 10 7	28, 61, 125, 156	0
1	C	266/300 (88%)	-0.07	10 (3%) 40 36	29, 51, 97, 122	0
1	D	237/300 (79%)	0.58	33 (13%) 2 1	48, 80, 138, 169	0
All	All	1023/1200 (85%)	0.14	74 (7%) 15 12	28, 58, 122, 169	0

The worst 5 of 74 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	280	ILE	7.7
1	C	279	LYS	5.3
1	D	70	ASP	5.1
1	B	279	LYS	4.8
1	D	150	TYR	4.8

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.