wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 07:05 AM GMT

PDB ID	:	2ZIX
Title	:	Crystal structure of the Mus81-Eme1 complex
Authors	:	Chang, J.H.; Kim, J.J.; Choi, J.M.; Lee, J.H.; Cho, Y.
Deposited on	:	2008-02-25
Resolution	:	3.50 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

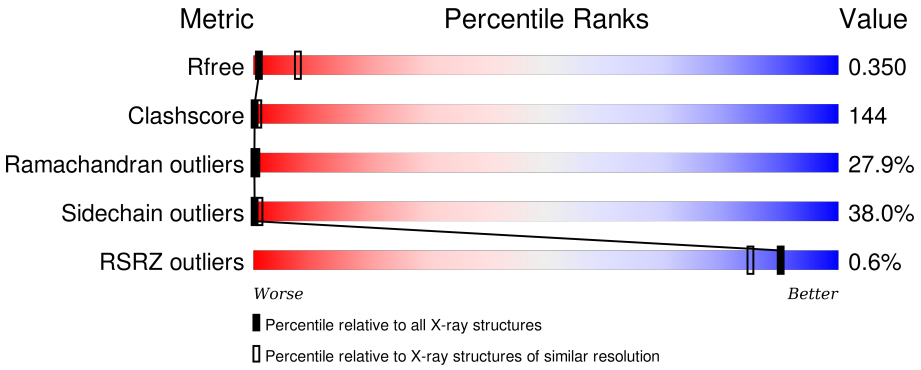
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

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 3.50 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}	91344	1051 (3.60-3.40)
Clashscore	102246	1157 (3.60-3.40)
Ramachandran outliers	100387	1120 (3.60-3.40)
Sidechain outliers	100360	1121 (3.60-3.40)
RSRZ outliers	91569	1058 (3.60-3.40)

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 on the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. 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	307	5% 41% 36% 7% 12%
2	B	341	% 6% 33% 30% 7% 24%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 4154 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 Crossover junction endonuclease MUS81.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	270	Total	C	N	O	S	0	0	0
			2136	1340	397	391	8			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	245	MET	-	INITIATING METHIONINE	UNP Q96NY9

- Molecule 2 is a protein called Crossover junction endonuclease EME1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	259	Total	C	N	O	S	0	0	0
			2018	1269	354	383	12			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	230	MET	-	EXPRESSION TAG	UNP Q96AY2
B	231	GLY	-	EXPRESSION TAG	UNP Q96AY2
B	232	SER	-	EXPRESSION TAG	UNP Q96AY2
B	233	SER	-	EXPRESSION TAG	UNP Q96AY2
B	234	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	235	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	236	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	237	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	238	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	239	HIS	-	EXPRESSION TAG	UNP Q96AY2
B	240	SER	-	EXPRESSION TAG	UNP Q96AY2
B	241	GLN	-	EXPRESSION TAG	UNP Q96AY2
B	242	ASP	-	EXPRESSION TAG	UNP Q96AY2
B	243	PRO	-	EXPRESSION TAG	UNP Q96AY2
B	244	ASN	-	EXPRESSION TAG	UNP Q96AY2

Continued on next page...

Continued from previous page...

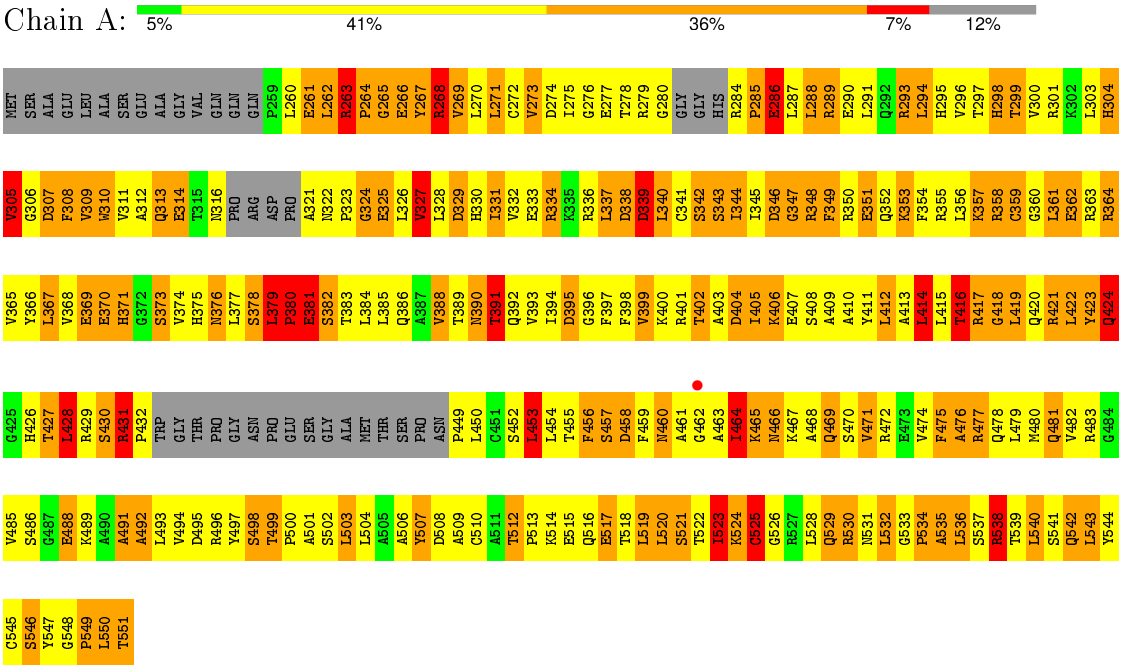
Chain	Residue	Modelled	Actual	Comment	Reference
B	245	SER	-	EXPRESSION TAG	UNP Q96AY2

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

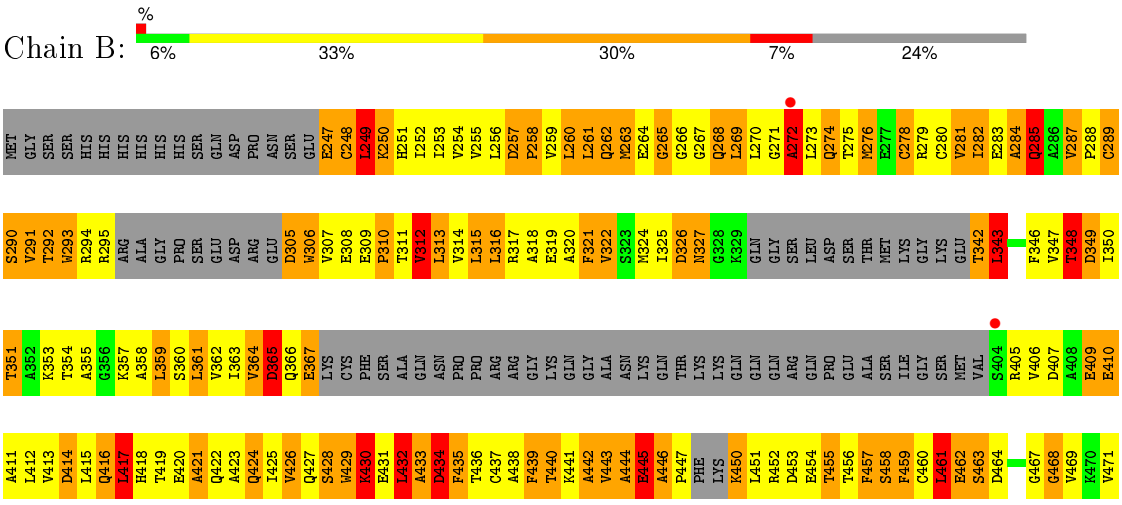
- Molecule 1: Crossover junction endonuclease MUS81

Chain A:



- Molecule 2: Crossover junction endonuclease EME1

Chain B:




D472
L473
A474
G475
R476
G477
L478
A479
L480
V481
W482
R483
R484
Q485
T486
Q487
Q488
L489
N490
R491
V492
S493
L494
E495
M496
A497
S498
A499
V500
V501
N502
A503
Y504
P505
S506
P507
Q508
L509
L510
V511
Q512
A513
Y514
Q515
Q516
C517
P518
S519
D520
K521
E522
R523
Q524
N525
L526
L527
A528
D529
I530
Q531


V532
R533
R534
G535
E536
GLY
VAL
THR
SER
T541
S542
R543
R544
I545
G546
P547
E548
L549
S550
R551
R552
I553
Y554
L555
Q556
M557
T558
T559
L560
Q561
P562
H563
L564
S565
L566
D567
S568
ALA
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α, β, γ	85.81Å 85.81Å 176.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	20.00 – 3.50 19.61 – 3.50	Depositor EDS
% Data completeness (in resolution range)	92.7 (20.00-3.50) 92.7 (19.61-3.50)	Depositor EDS
R_{merge}	0.13	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	5.69 (at 3.52Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.285 , 0.346 0.283 , 0.350	Depositor DCC
R_{free} test set	441 reflections (5.04%)	DCC
Wilson B-factor (Å^2)	73.1	Xtriage
Anisotropy	0.317	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.10 , -28.4	EDS
Estimated twinning fraction	0.188 for -h,-k,l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.32$, $< L^2 > = 0.15$	Xtriage
Outliers	1 of 9187 reflections (0.011%)	Xtriage
F_o,F_c correlation	0.83	EDS
Total number of atoms	4154	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

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	0.62	1/2168 (0.0%)	1.03	6/2923 (0.2%)
2	B	0.63	1/2043 (0.0%)	0.98	3/2765 (0.1%)
All	All	0.63	2/4211 (0.0%)	1.01	9/5688 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	1	6
2	B	0	9
All	All	1	15

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	484	ARG	CG-CD	9.51	1.75	1.51
1	A	380	PRO	CG-CD	5.26	1.68	1.50

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	339	ASP	N-CA-C	-8.72	87.44	111.00
1	A	380	PRO	N-CA-C	6.48	128.94	112.10
2	B	365	ASP	CB-CG-OD1	6.05	123.74	118.30
1	A	538	ARG	NE-CZ-NH2	-5.44	117.58	120.30
2	B	249	LEU	CA-CB-CG	5.42	127.77	115.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	A	380	PRO	CA

5 of 15 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	268	ARG	Peptide
1	A	338	ASP	Peptide
1	A	378	SER	Peptide
1	A	379	LEU	Peptide
1	A	431	ARG	Peptide

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	2136	0	2188	673	1
2	B	2018	0	2033	628	1
All	All	4154	0	4221	1208	2

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 144.

The worst 5 of 1208 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)
2:B:484:ARG:CD	2:B:484:ARG:CG	1.75	1.64
1:A:491:ALA:CB	1:A:492:ALA:HB2	1.30	1.59
1:A:414:LEU:HD21	2:B:413:VAL:CG2	1.41	1.47
1:A:414:LEU:CD2	2:B:413:VAL:HG21	1.43	1.45
1:A:491:ALA:HB3	1:A:492:ALA:CB	1.44	1.44

All (2) 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)
2:B:266:GLY:O	2:B:266:GLY:O[6_555]	1.90	0.30

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:380:PRO:CD	1:A:538:ARG:NH1[4_555]	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	262/307 (85%)	127 (48%)	59 (22%)	76 (29%)	0 0
2	B	247/341 (72%)	100 (40%)	81 (33%)	66 (27%)	0 0
All	All	509/648 (78%)	227 (45%)	140 (28%)	142 (28%)	0 0

5 of 142 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	262	LEU
1	A	263	ARG
1	A	279	ARG
1	A	285	PRO
1	A	305	VAL

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	233/260 (90%)	150 (64%)	83 (36%)	0 1
2	B	220/290 (76%)	131 (60%)	89 (40%)	0 1

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	453/550 (82%)	281 (62%)	172 (38%)	0 1

5 of 172 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	530	ARG
2	B	292	THR
2	B	536	GLU
1	A	543	LEU
2	B	263	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	542	GLN
2	B	327	ASN
2	B	490	ASN
1	A	481	GLN
1	A	531	ASN

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 carbohydrates in this entry.

5.6 Ligand geometry

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, 95^{th} 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	270/307 (87%)	-0.70	1 (0%)	93	90	8, 34, 44, 53	0
2	B	259/341 (75%)	-0.64	2 (0%)	87	80	15, 35, 45, 52	0
All	All	529/648 (81%)	-0.67	3 (0%)	90	85	8, 34, 44, 53	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	404	SER	3.5
1	A	462	GLY	2.8
2	B	272	ALA	2.4

6.2 Non-standard residues in protein, DNA, RNA chains

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.