wwPDB X-ray Structure Validation Summary Report (i)

Mar 18, 2024 – 05:09 PM JST

PDB ID	:	6LOH
Title	:	Crystal structure of the catalytic domain of human ubiquitin ligase AREL1
Authors	:	Chen, Z.Z.; Li, Z.H.; Shang, G.H.
Deposited on	:	2020-01-05
Resolution	:	3.21 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 (i)) were used in the production of this report:

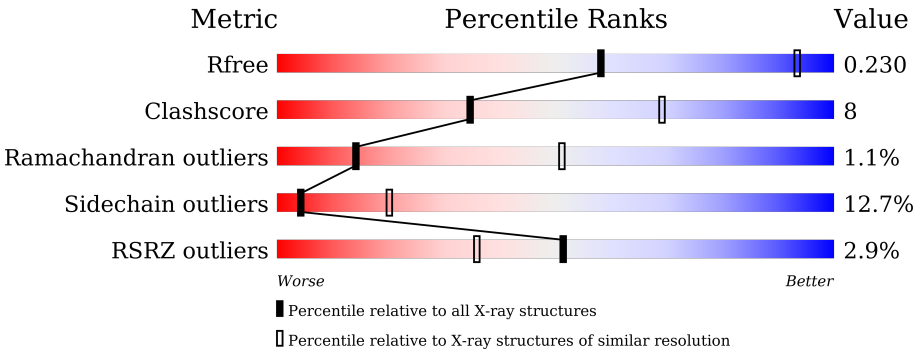
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.21 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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	389	2% 76% 18%
1	B	389	2% 70% 25%
1	C	389	4% 66% 25%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 8990 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 Apoptosis-resistant E3 ubiquitin protein ligase 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	373	Total	C	N	O	S	0	0	0
			3003	1921	517	551	14			
1	B	375	Total	C	N	O	S	0	0	0
			3002	1927	514	546	15			
1	C	372	Total	C	N	O	S	0	0	0
			2979	1907	509	548	15			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	423	HIS	-	expression tag	UNP O15033
A	424	HIS	-	expression tag	UNP O15033
A	425	HIS	-	expression tag	UNP O15033
A	426	HIS	-	expression tag	UNP O15033
A	427	HIS	-	expression tag	UNP O15033
A	428	HIS	-	expression tag	UNP O15033
A	429	GLU	-	expression tag	UNP O15033
A	430	ASN	-	expression tag	UNP O15033
A	431	LEU	-	expression tag	UNP O15033
A	432	TYR	-	expression tag	UNP O15033
A	433	PHE	-	expression tag	UNP O15033
A	434	GLN	-	expression tag	UNP O15033
B	423	HIS	-	expression tag	UNP O15033
B	424	HIS	-	expression tag	UNP O15033
B	425	HIS	-	expression tag	UNP O15033
B	426	HIS	-	expression tag	UNP O15033
B	427	HIS	-	expression tag	UNP O15033
B	428	HIS	-	expression tag	UNP O15033
B	429	GLU	-	expression tag	UNP O15033
B	430	ASN	-	expression tag	UNP O15033
B	431	LEU	-	expression tag	UNP O15033
B	432	TYR	-	expression tag	UNP O15033
B	433	PHE	-	expression tag	UNP O15033

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	434	GLN	-	expression tag	UNP O15033
C	423	HIS	-	expression tag	UNP O15033
C	424	HIS	-	expression tag	UNP O15033
C	425	HIS	-	expression tag	UNP O15033
C	426	HIS	-	expression tag	UNP O15033
C	427	HIS	-	expression tag	UNP O15033
C	428	HIS	-	expression tag	UNP O15033
C	429	GLU	-	expression tag	UNP O15033
C	430	ASN	-	expression tag	UNP O15033
C	431	LEU	-	expression tag	UNP O15033
C	432	TYR	-	expression tag	UNP O15033
C	433	PHE	-	expression tag	UNP O15033
C	434	GLN	-	expression tag	UNP O15033

- Molecule 2 is water.

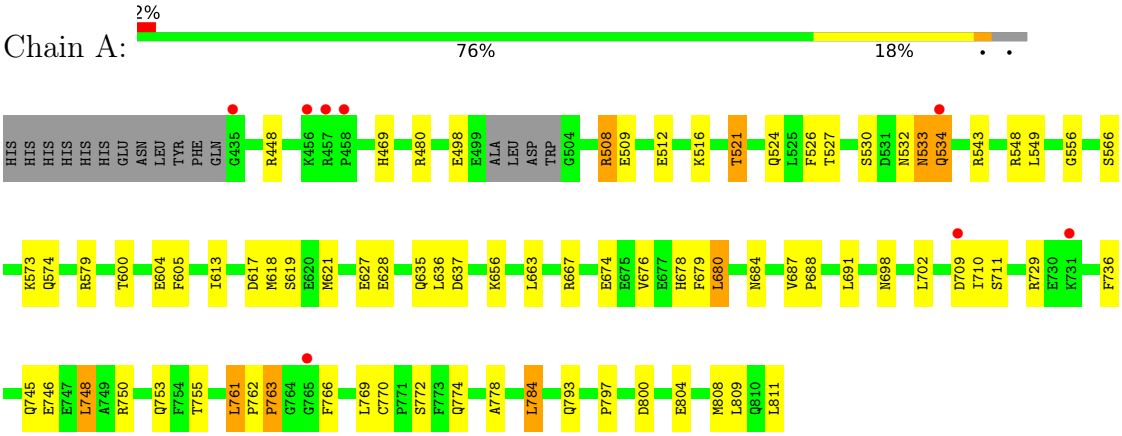
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	3	Total O 3 3	0	0
2	B	3	Total O 3 3	0	0

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.

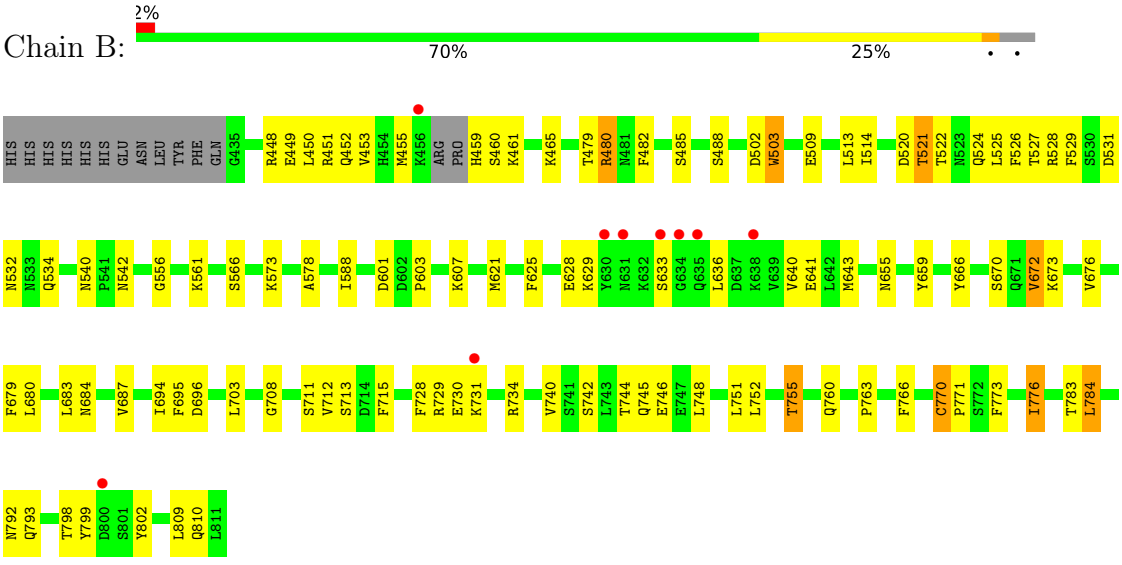
• Molecule 1: Apoptosis-resistant E3 ubiquitin protein ligase 1

Chain A:



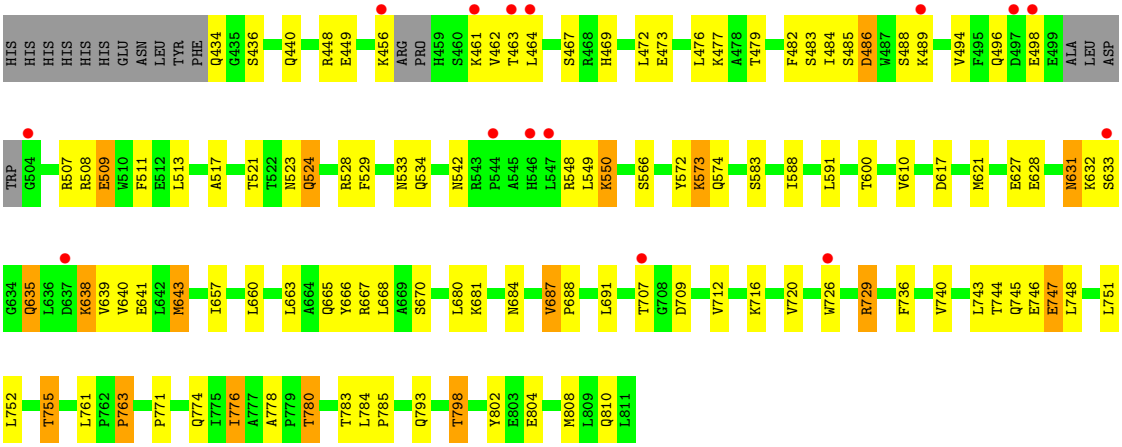
• Molecule 1: Apoptosis-resistant E3 ubiquitin protein ligase 1

Chain B:



• Molecule 1: Apoptosis-resistant E3 ubiquitin protein ligase 1

Chain C:


HIS HIS HIS HIS HIS HIS GLU ASN LEU TYR PHE Q434 G435 S436 Q440 R448 E449 K456 ARG PRO H459 S460 K461 V462 T463 L464 S467 R468 H469 L472 E473 L476 K477 A478 T479 F482 S483 I484 S485 D486 W487 S488 K489 V494 F495 Q496 D497 E498 E499 ALA LEU ASP
TRP G504 R507 R508 E509 W510 F511 E512 L513 A517 T521 T522 N523 Q524 R528 F529 N533 Q534 N542 R543 P544 A545 H546 L547 R548 L549 K550 S566 Y572 K573 Q574 S583 I588 L591 T600 V610 D617 M621 E627 E628 N631 K632 S633
G634 Q635 L636 D637 K638 V639 V640 E641 L642 M643 I657 L660 L663 A664 Q665 Y666 R667 L668 A669 S670 L680 K681 N684 V687 P688 L691 T707 G708 D709 V712 K716 V720 W726 R729 F736 V740 L743 T744 Q745 E746 E747 L748 L751
L752 T755 L761 P762 P763 P771 Q774 I775 I776 A777 A778 P779 T780 T783 L784 P785 Q793 T798 Y802 E803 E804 M808 L809 Q810 L811

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	150.23Å 150.23Å 342.64Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	49.22 – 3.21 49.17 – 3.21	Depositor EDS
% Data completeness (in resolution range)	99.8 (49.22-3.21) 99.9 (49.17-3.21)	Depositor EDS
R_{merge}	0.12	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.81 (at 3.19Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.211 , 0.225 0.213 , 0.230	Depositor DCC
R_{free} test set	1878 reflections (4.91%)	wwPDB-VP
Wilson B-factor (Å^2)	71.9	Xtriage
Anisotropy	0.030	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 71.5	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	0.92	EDS
Total number of atoms	8990	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.56% 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, 0.333 respectively for untwinned datasets, and 0.375, 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	0.67	0/3075	0.87	0/4158
1	B	0.68	0/3075	0.85	0/4160
1	C	0.71	0/3048	0.88	0/4119
All	All	0.69	0/9198	0.86	0/12437

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	3003	0	2935	36	0
1	B	3002	0	2924	57	0
1	C	2979	0	2900	54	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
All	All	8990	0	8759	147	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 8.

The worst 5 of 147 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:770:CYS:CB	1:B:771:PRO:HD3	1.71	1.16
1:B:776:ILE:HD11	1:B:793:GLN:OE1	1.56	1.04
1:B:776:ILE:CD1	1:B:793:GLN:OE1	2.13	0.97
1:C:804:GLU:O	1:C:808:MET:HG2	1.72	0.89
1:B:770:CYS:CB	1:B:771:PRO:CD	2.52	0.84

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	369/389 (95%)	336 (91%)	29 (8%)	4 (1%)	14	51
1	B	371/389 (95%)	338 (91%)	28 (8%)	5 (1%)	12	47
1	C	366/389 (94%)	337 (92%)	26 (7%)	3 (1%)	19	58
All	All	1106/1167 (95%)	1011 (91%)	83 (8%)	12 (1%)	14	51

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	498	GLU
1	B	521	THR
1	B	633	SER
1	B	636	LEU
1	A	636	LEU

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	326/346 (94%)	292 (90%)	34 (10%)	7	28
1	B	321/346 (93%)	288 (90%)	33 (10%)	7	29
1	C	321/346 (93%)	265 (83%)	56 (17%)	2	10
All	All	968/1038 (93%)	845 (87%)	123 (13%)	4	20

5 of 123 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	742	SER
1	C	687	VAL
1	C	462	VAL
1	C	681	LYS
1	C	761	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	806	HIS
1	C	792	ASN
1	B	810	GLN
1	C	635	GLN
1	B	806	HIS

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, 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	373/389 (95%)	0.04	8 (2%)	63	49	41, 64, 112, 141	0
1	B	375/389 (96%)	0.10	9 (2%)	59	44	44, 69, 113, 146	0
1	C	372/389 (95%)	0.21	15 (4%)	38	25	40, 74, 128, 174	0
All	All	1120/1167 (95%)	0.11	32 (2%)	51	36	40, 69, 121, 174	0

The worst 5 of 32 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	504	GLY	4.9
1	C	497	ASP	4.4
1	B	731	LYS	3.9
1	C	633	SER	3.4
1	B	456	LYS	3.3

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 monosaccharides in this entry.

6.4 Ligands

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.