wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 03:08 PM EDT

PDB ID	:	5B0U
Title	:	Crystal structure of the mutated 19 kDa protein of Oplophorus luciferase (nanoKAZ)
Authors	:	Tomabechi, Y.; Ehara, H.; Sekine, S.I.; Shirouzu, M.
Deposited on	:	2015-11-04
Resolution	:	1.71 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:

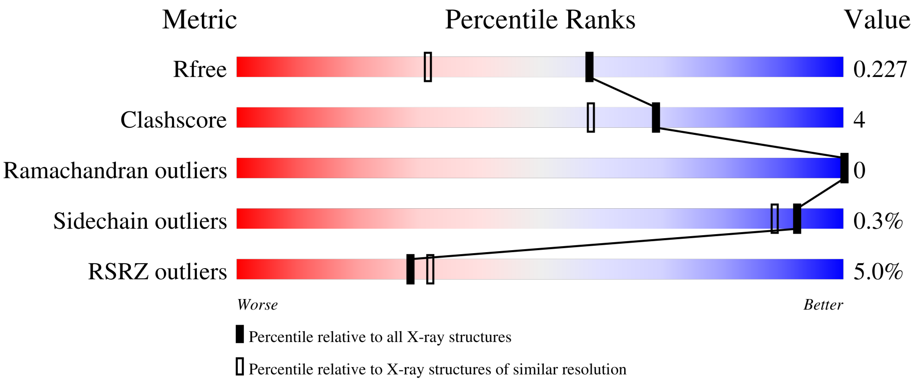
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 1.71 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	5722 (1.74-1.70)
Clashscore	141614	6152 (1.74-1.70)
Ramachandran outliers	138981	6051 (1.74-1.70)
Sidechain outliers	138945	6051 (1.74-1.70)
RSRZ outliers	127900	5629 (1.74-1.70)

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	191	3% / 77% 12% 11%
1	B	191	6% / 83% 6% 12%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3020 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 Oplophorus-luciferin 2-monooxygenase catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	170	Total 1348	C 872	N 224	O 249	S 3	0	0	0
1	B	169	Total 1340	C 866	N 223	O 248	S 3	0	0	0

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-21	MET	-	expression tag	UNP Q9GV45
A	-20	ASN	-	expression tag	UNP Q9GV45
A	-19	HIS	-	expression tag	UNP Q9GV45
A	-18	LYS	-	expression tag	UNP Q9GV45
A	-17	VAL	-	expression tag	UNP Q9GV45
A	-16	HIS	-	expression tag	UNP Q9GV45
A	-15	HIS	-	expression tag	UNP Q9GV45
A	-14	HIS	-	expression tag	UNP Q9GV45
A	-13	HIS	-	expression tag	UNP Q9GV45
A	-12	HIS	-	expression tag	UNP Q9GV45
A	-11	HIS	-	expression tag	UNP Q9GV45
A	-10	MET	-	expression tag	UNP Q9GV45
A	-9	GLU	-	expression tag	UNP Q9GV45
A	-8	LEU	-	expression tag	UNP Q9GV45
A	-7	GLY	-	expression tag	UNP Q9GV45
A	-6	THR	-	expression tag	UNP Q9GV45
A	-5	LEU	-	expression tag	UNP Q9GV45
A	-4	GLU	-	expression tag	UNP Q9GV45
A	-3	GLY	-	expression tag	UNP Q9GV45
A	-2	SER	-	expression tag	UNP Q9GV45
A	-1	GLU	-	expression tag	UNP Q9GV45
A	0	PHE	-	expression tag	UNP Q9GV45
A	4	GLU	ALA	engineered mutation	UNP Q9GV45
A	11	ARG	GLN	engineered mutation	UNP Q9GV45
A	18	LEU	GLN	engineered mutation	UNP Q9GV45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	27	VAL	LEU	engineered mutation	UNP Q9GV45
A	33	ASN	ALA	engineered mutation	UNP Q9GV45
A	43	ARG	LYS	engineered mutation	UNP Q9GV45
A	44	ILE	VAL	engineered mutation	UNP Q9GV45
A	54	ILE	ALA	engineered mutation	UNP Q9GV45
A	68	ASP	PHE	engineered mutation	UNP Q9GV45
A	72	GLN	LEU	engineered mutation	UNP Q9GV45
A	75	LYS	MET	engineered mutation	UNP Q9GV45
A	90	VAL	ILE	engineered mutation	UNP Q9GV45
A	115	GLU	PRO	engineered mutation	UNP Q9GV45
A	124	LYS	GLN	engineered mutation	UNP Q9GV45
A	138	ILE	TYR	engineered mutation	UNP Q9GV45
A	166	ARG	ASN	engineered mutation	UNP Q9GV45
B	-21	MET	-	expression tag	UNP Q9GV45
B	-20	ASN	-	expression tag	UNP Q9GV45
B	-19	HIS	-	expression tag	UNP Q9GV45
B	-18	LYS	-	expression tag	UNP Q9GV45
B	-17	VAL	-	expression tag	UNP Q9GV45
B	-16	HIS	-	expression tag	UNP Q9GV45
B	-15	HIS	-	expression tag	UNP Q9GV45
B	-14	HIS	-	expression tag	UNP Q9GV45
B	-13	HIS	-	expression tag	UNP Q9GV45
B	-12	HIS	-	expression tag	UNP Q9GV45
B	-11	HIS	-	expression tag	UNP Q9GV45
B	-10	MET	-	expression tag	UNP Q9GV45
B	-9	GLU	-	expression tag	UNP Q9GV45
B	-8	LEU	-	expression tag	UNP Q9GV45
B	-7	GLY	-	expression tag	UNP Q9GV45
B	-6	THR	-	expression tag	UNP Q9GV45
B	-5	LEU	-	expression tag	UNP Q9GV45
B	-4	GLU	-	expression tag	UNP Q9GV45
B	-3	GLY	-	expression tag	UNP Q9GV45
B	-2	SER	-	expression tag	UNP Q9GV45
B	-1	GLU	-	expression tag	UNP Q9GV45
B	0	PHE	-	expression tag	UNP Q9GV45
B	4	GLU	ALA	engineered mutation	UNP Q9GV45
B	11	ARG	GLN	engineered mutation	UNP Q9GV45
B	18	LEU	GLN	engineered mutation	UNP Q9GV45
B	27	VAL	LEU	engineered mutation	UNP Q9GV45
B	33	ASN	ALA	engineered mutation	UNP Q9GV45
B	43	ARG	LYS	engineered mutation	UNP Q9GV45
B	44	ILE	VAL	engineered mutation	UNP Q9GV45

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	54	ILE	ALA	engineered mutation	UNP Q9GV45
B	68	ASP	PHE	engineered mutation	UNP Q9GV45
B	72	GLN	LEU	engineered mutation	UNP Q9GV45
B	75	LYS	MET	engineered mutation	UNP Q9GV45
B	90	VAL	ILE	engineered mutation	UNP Q9GV45
B	115	GLU	PRO	engineered mutation	UNP Q9GV45
B	124	LYS	GLN	engineered mutation	UNP Q9GV45
B	138	ILE	TYR	engineered mutation	UNP Q9GV45
B	166	ARG	ASN	engineered mutation	UNP Q9GV45

- Molecule 2 is water.

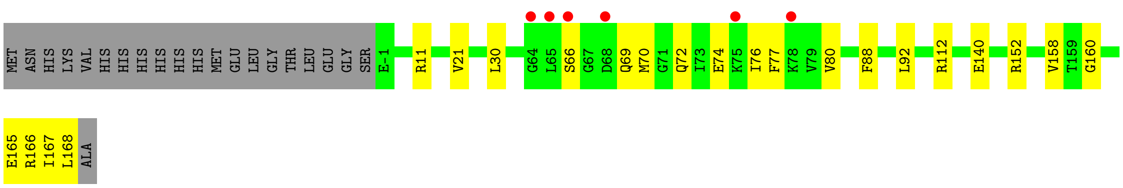
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	171	Total O 171 171	0	0
2	B	161	Total O 161 161	0	0

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.

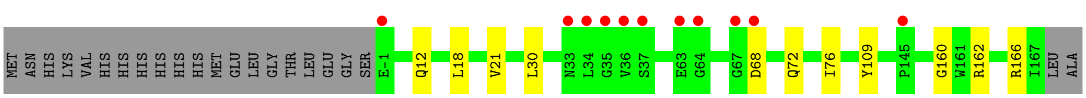
- Molecule 1: Oplophorus-luciferin 2-monooxygenase catalytic subunit

Chain A:

- Molecule 1: Oplophorus-luciferin 2-monooxygenase catalytic subunit

Chain B:

4 Data and refinement statistics

Property	Value	Source
Space group	P 42 21 2	Depositor
Cell constants a, b, c, α, β, γ	79.07Å 79.07Å 123.42Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.65 – 1.71 48.65 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.65-1.71) 99.6 (48.65-1.65)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.94 (at 1.65Å)	Xtriage
Refinement program	PHENIX 1.9_1690	Depositor
R, R_{free}	0.201 , 0.226 0.203 , 0.227	Depositor DCC
R_{free} test set	2000 reflections (4.20%)	wwPDB-VP
Wilson B-factor (Å^2)	24.2	Xtriage
Anisotropy	0.097	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.37 , 43.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3020	wwPDB-VP
Average B, all atoms (Å^2)	30.0	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 59.89 % 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 1.6497e-05. 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	0.44	0/1378	0.66	1/1871 (0.1%)
1	B	0.39	0/1370	0.60	0/1860
All	All	0.41	0/2748	0.63	1/3731 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	112	ARG	NE-CZ-NH1	-5.03	117.78	120.30

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	1348	0	1345	14	0
1	B	1340	0	1334	6	0
2	A	171	0	0	2	0
2	B	161	0	0	1	0
All	All	3020	0	2679	20	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 4.

The worst 5 of 20 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:166:ARG:NH2	2:B:202:HOH:O	2.25	0.70
1:A:72:GLN:NE2	2:A:201:HOH:O	2.25	0.69
1:A:11:ARG:NH1	1:A:165:GLU:OE2	2.30	0.65
1:A:167:ILE:O	1:A:168:LEU:HB2	1.99	0.62
1:B:18:LEU:HD12	1:B:162:ARG:CZ	2.33	0.59

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	168/191 (88%)	162 (96%)	6 (4%)	0	100	100
1	B	167/191 (87%)	160 (96%)	7 (4%)	0	100	100
All	All	335/382 (88%)	322 (96%)	13 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	148/166 (89%)	148 (100%)	0	100	100
1	B	147/166 (89%)	146 (99%)	1 (1%)	84	76
All	All	295/332 (89%)	294 (100%)	1 (0%)	92	89

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	109	TYR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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, 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	170/191 (89%)	0.21	6 (3%)	44	48	15, 24, 54, 73	0
1	B	169/191 (88%)	0.26	11 (6%)	18	21	17, 30, 54, 65	0
All	All	339/382 (88%)	0.23	17 (5%)	28	32	15, 27, 55, 73	0

The worst 5 of 17 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	65	LEU	3.9
1	A	68	ASP	3.6
1	B	36	VAL	3.6
1	B	64	GLY	3.5
1	B	-1	GLU	3.4

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.