wwPDB X-ray Structure Validation Summary Report (i)

May 23, 2020 – 06:57 am BST

PDB ID	:	5LON
Title	:	Structure of /K. lactis/ Dcp1-Dcp2 decapping complex.
Authors	:	Charenton, C.; Taverniti, V.; Gaudon-Plesse, C.; Back, R.; Seraphin, B.; Graille, M.
Deposited on	:	2016-08-09
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
https://www.wwpdb.org/validation/2017/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:

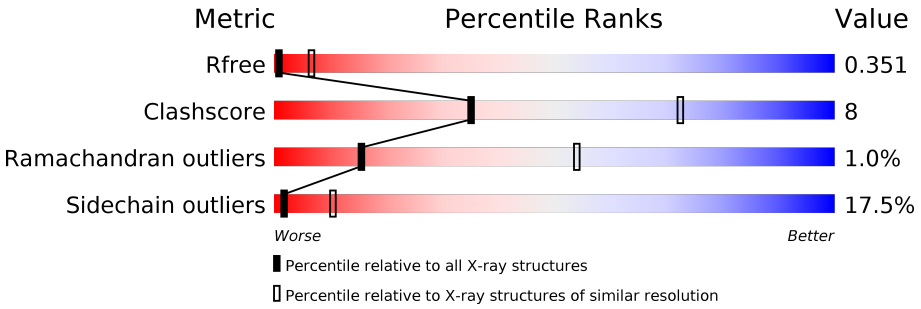
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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}	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)

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 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%

Mol	Chain	Length	Quality of chain
1	A	281	58% 18% • 21%
2	B	188	55% 34% • 7%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 3242 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 KLLA0F23980p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	221	Total	C	N	O	S	0	0	0
			1807	1177	298	326	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	276	HIS	-	expression tag	UNP Q6CIU1
A	277	HIS	-	expression tag	UNP Q6CIU1
A	278	HIS	-	expression tag	UNP Q6CIU1
A	279	HIS	-	expression tag	UNP Q6CIU1
A	280	HIS	-	expression tag	UNP Q6CIU1
A	281	HIS	-	expression tag	UNP Q6CIU1

- Molecule 2 is a protein called KLLA0E01827p.

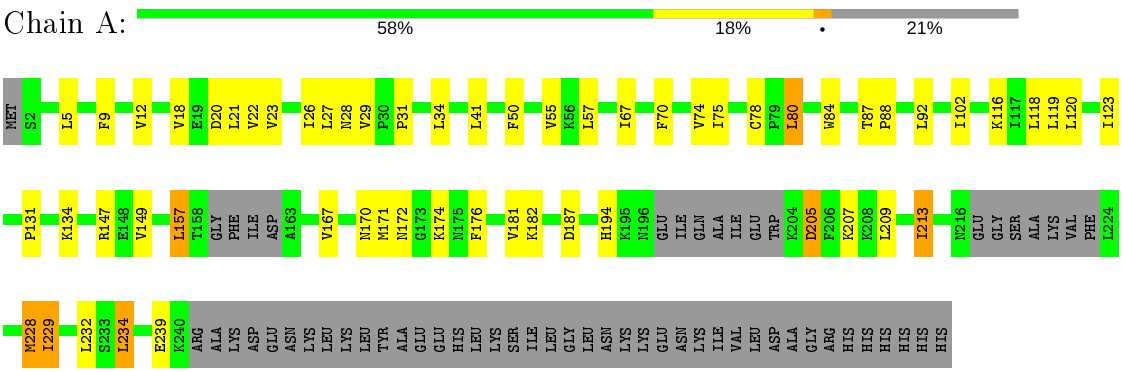
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	174	Total	C	N	O	S	0	0	0
			1435	923	241	268	3			

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 the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. 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. 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.

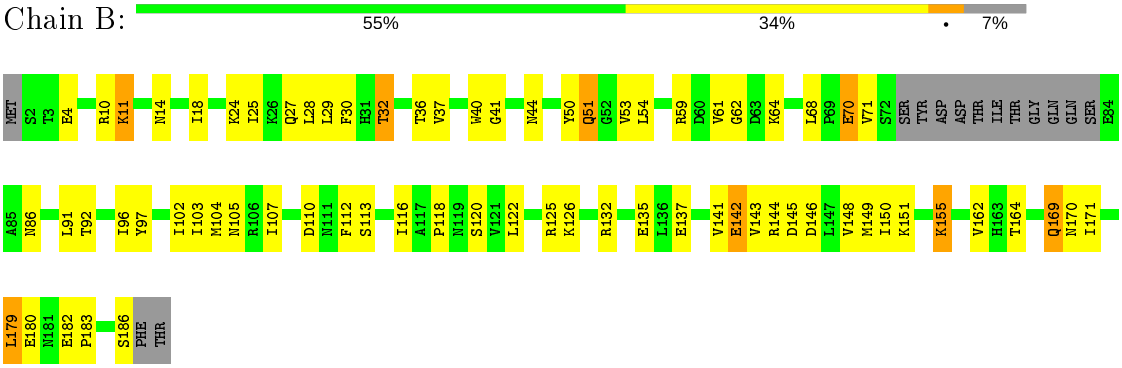
- Molecule 1: KLLA0F23980p

Chain A:



- Molecule 2: KLLA0E01827p

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 4 3 2	Depositor
Cell constants a, b, c, α, β, γ	216.69Å 216.69Å 216.69Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.45 – 3.50 46.20 – 3.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.45-3.50) 100.0 (46.20-3.50)	Depositor EDS
R_{merge}	0.04	Depositor
R_{sym}	0.26	Depositor
$< I/\sigma(I) >$ [1]	1.20 (at 3.48Å)	Xtriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.250 , 0.312 0.299 , 0.351	Depositor DCC
R_{free} test set	543 reflections (4.82%)	wwPDB-VP
Wilson B-factor (Å^2)	147.6	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 268.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	3242	wwPDB-VP
Average B, all atoms (Å^2)	209.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.17% 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	0.51	0/1848	0.70	0/2495
2	B	0.52	0/1468	0.75	1/1994 (0.1%)
All	All	0.51	0/3316	0.72	1/4489 (0.0%)

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
2	B	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	146	ASP	CB-CG-OD2	5.23	123.00	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	B	70	GLU	Peptide

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	1807	0	1836	22	0
2	B	1435	0	1429	32	0
All	All	3242	0	3265	53	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 53 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:61:VAL:HG21	2:B:96:ILE:HA	1.70	0.73
2:B:50:TYR:HA	2:B:105:ASN:HD21	1.55	0.71
1:A:102:ILE:HB	1:A:174:LYS:HG2	1.73	0.68
1:A:21:LEU:HD21	1:A:50:PHE:HD1	1.58	0.66
1:A:41:LEU:HD22	1:A:92:LEU:HB3	1.79	0.63

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	213/281 (76%)	185 (87%)	27 (13%)	1 (0%)	29 68
2	B	170/188 (90%)	151 (89%)	16 (9%)	3 (2%)	8 41
All	All	383/469 (82%)	336 (88%)	43 (11%)	4 (1%)	15 54

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	71	VAL
2	B	155	LYS
2	B	41	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	31	PRO

5.3.2 Protein sidechains

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	206/257 (80%)	172 (84%)	34 (16%)	2 13
2	B	159/172 (92%)	129 (81%)	30 (19%)	1 8
All	All	365/429 (85%)	301 (82%)	64 (18%)	2 10

5 of 64 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	209	LEU
2	B	10	ARG
2	B	164	THR
1	A	213	ILE
1	A	234	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	170	ASN
2	B	169	GLN
2	B	27	GLN
1	A	64	ASN
1	A	175	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

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

Unable to reproduce the depositors R factor - this section is therefore empty.

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

Unable to reproduce the depositors R factor - this section is therefore empty.

6.3 Carbohydrates

Unable to reproduce the depositors R factor - this section is therefore empty.

6.4 Ligands

Unable to reproduce the depositors R factor - this section is therefore empty.

6.5 Other polymers

Unable to reproduce the depositors R factor - this section is therefore empty.