wwPDB X-ray Structure Validation Summary Report (i)

Jun 19, 2024 – 07:55 AM EDT

PDB ID	:	4L0K
Title	:	Crystal structure of a type II restriction endonuclease
Authors	:	Zhuo, W.; Ge, J.; Yang, M.
Deposited on	:	2013-05-31
Resolution	:	2.33 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:

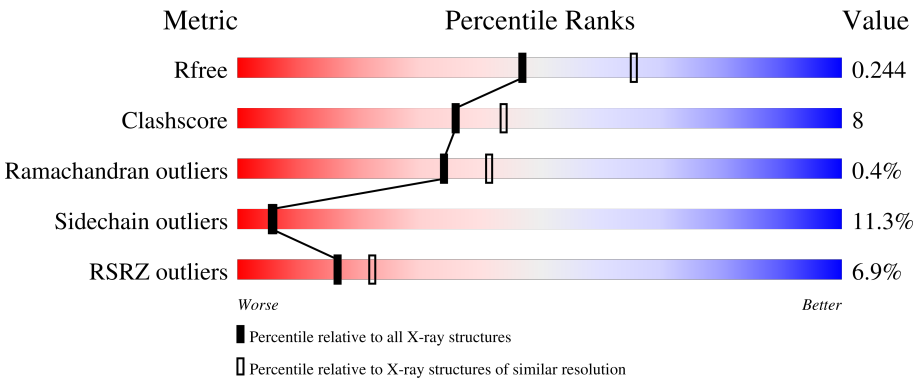
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.33 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	5974 (2.34-2.30)
Clashscore	141614	6604 (2.34-2.30)
Ramachandran outliers	138981	6523 (2.34-2.30)
Sidechain outliers	138945	6523 (2.34-2.30)
RSRZ outliers	127900	5855 (2.34-2.30)

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	227	2%, 72%, 16%, • •, 8%
1	B	227	10%, 69%, 16%, •, 11%
1	C	227	11%, 69%, 16%, • •, 10%
1	D	227	2%, 72%, 14%, 5%, •, 9%

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 6903 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 DraIII.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	208	1653	1059	283	297	14	0	0	0
1	B	202	1595	1022	272	287	14	0	0	0
1	C	204	1627	1045	279	290	13	0	0	0
1	D	207	1656	1062	284	296	14	0	0	0

- Molecule 2 is water.

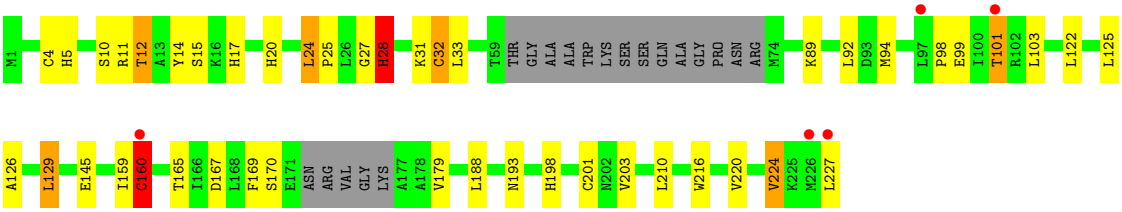
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
2	A	93	93	93	0	0
2	B	92	92	92	0	0
2	C	93	93	93	0	0
2	D	94	94	94	0	0

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.

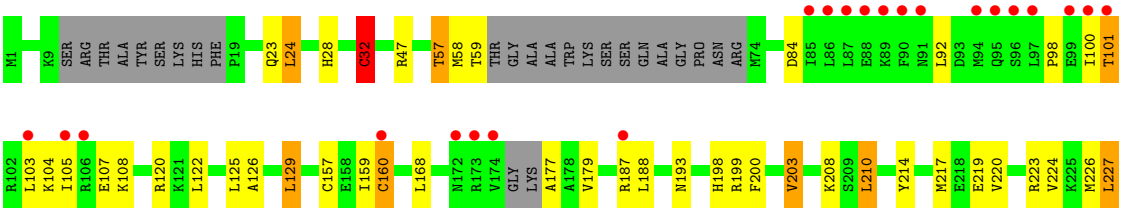
- Molecule 1: DraIII

Chain A:



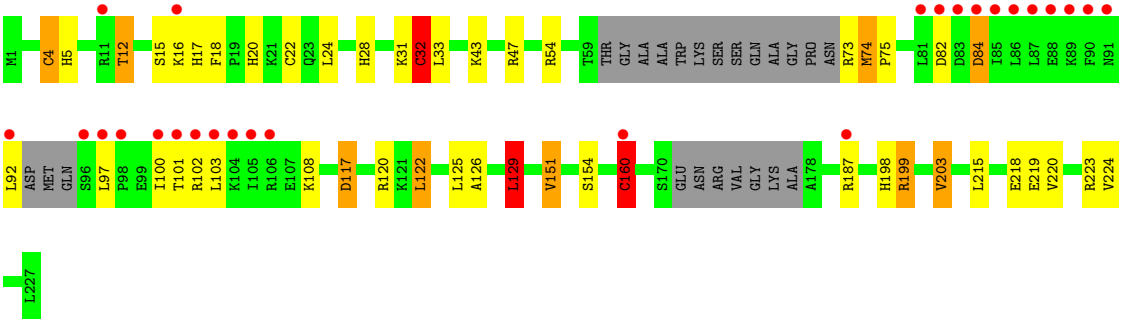
- Molecule 1: DraIII

Chain B:



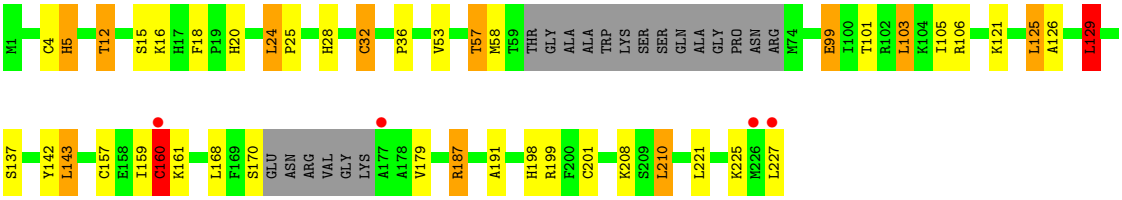
- Molecule 1: DraIII

Chain C:



- Molecule 1: DraIII

Chain D:


M1
C4
H5
T12
S15
K16
H17
F18
P19
H20
L24
P25
H28
C32
P36
V53
T57
M58
T59
THR
GLY
ALA
ALA
TRP
LYS
SER
SER
GLN
ALA
GLY
PRO
ASN
ARG
M74
E99
I100
T101
R102
L103
K104
I105
R106
K121
L125
A126
L129
S137
Y142
L143
C157
E158
I159
C160
K161
L168
F169
S170
GLU
ASN
ARG
VAL
GLY
LYS
A177
A178
V179
R187
A191
H198
R199
F200
C201
K208
S209
L210
L221
K225
M226
L227

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	58.40Å 119.06Å 82.11Å 90.00° 92.89° 90.00°	Depositor
Resolution (Å)	35.72 – 2.33 35.72 – 2.33	Depositor EDS
% Data completeness (in resolution range)	98.3 (35.72-2.33) 98.3 (35.72-2.33)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.05	Depositor
$< I/\sigma(I) >$ [1]	3.09 (at 2.34Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.2_1309)	Depositor
R, R_{free}	0.192 , 0.243 0.195 , 0.244	Depositor DCC
R_{free} test set	2389 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	39.3	Xtriage
Anisotropy	0.377	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 57.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.037 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6903	wwPDB-VP
Average B, all atoms (Å^2)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.30% 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 (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	0.69	1/1691 (0.1%)	0.60	0/2287
1	B	0.53	1/1629 (0.1%)	0.66	1/2201 (0.0%)
1	C	0.58	2/1664 (0.1%)	0.63	2/2248 (0.1%)
1	D	0.56	1/1694 (0.1%)	0.60	1/2288 (0.0%)
All	All	0.59	5/6678 (0.1%)	0.62	4/9024 (0.0%)

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	160	CYS	CB-SG	15.30	2.08	1.82
1	D	160	CYS	CB-SG	13.40	2.05	1.82
1	C	160	CYS	CB-SG	10.30	1.99	1.82
1	B	32	CYS	CB-SG	9.97	1.99	1.82
1	C	32	CYS	CB-SG	6.90	1.94	1.82

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	227	LEU	CB-CG-CD2	-5.91	100.96	111.00
1	D	129	LEU	CA-CB-CG	5.42	127.76	115.30
1	C	129	LEU	CA-CB-CG	5.39	127.70	115.30
1	C	4	CYS	CA-CB-SG	5.18	123.33	114.00

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	1653	0	1644	25	0
1	B	1595	0	1588	29	0
1	C	1627	0	1628	27	0
1	D	1656	0	1664	27	0
2	A	93	0	0	2	0
2	B	92	0	0	5	0
2	C	93	0	0	5	0
2	D	94	0	0	1	0
All	All	6903	0	6524	105	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 105 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:D:160:CYS:SG	1:D:160:CYS:CB	2.05	1.45
1:A:160:CYS:SG	1:A:160:CYS:CB	2.08	1.39
1:D:160:CYS:SG	1:D:198:HIS:ND1	2.19	1.15
1:B:28:HIS:NE2	1:B:32:CYS:SG	2.19	1.14
1:A:160:CYS:SG	1:A:198:HIS:ND1	2.22	1.11

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	202/227 (89%)	194 (96%)	6 (3%)	2 (1%)	15 17
1	B	194/227 (86%)	187 (96%)	7 (4%)	0	100 100
1	C	196/227 (86%)	192 (98%)	4 (2%)	0	100 100
1	D	201/227 (88%)	196 (98%)	4 (2%)	1 (0%)	29 35

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles
All	All	793/908 (87%)	769 (97%)	21 (3%)	3 (0%)	34 41

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	5	HIS
1	A	28	HIS
1	A	27	GLY

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	181/197 (92%)	159 (88%)	22 (12%)	5 5
1	B	174/197 (88%)	159 (91%)	15 (9%)	10 12
1	C	179/197 (91%)	155 (87%)	24 (13%)	4 3
1	D	183/197 (93%)	163 (89%)	20 (11%)	6 7
All	All	717/788 (91%)	636 (89%)	81 (11%)	6 6

5 of 81 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	160	CYS
1	D	125	LEU
1	C	199	ARG
1	D	24	LEU
1	D	160	CYS

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, 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	208/227 (91%)	-0.11	5 (2%) 59 66	24, 39, 73, 80	0
1	B	202/227 (88%)	0.40	22 (10%) 5 8	24, 39, 87, 98	0
1	C	204/227 (89%)	0.54	26 (12%) 3 5	24, 39, 108, 119	0
1	D	207/227 (91%)	-0.14	4 (1%) 66 73	23, 41, 69, 93	0
All	All	821/908 (90%)	0.17	57 (6%) 16 22	23, 40, 85, 119	0

The worst 5 of 57 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	92	LEU	10.2
1	C	87	LEU	9.7
1	C	101	THR	8.4
1	C	105	ILE	6.8
1	B	100	ILE	6.5

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.