wwPDB X-ray Structure Validation Summary Report (i)

Sep 3, 2023 – 01:52 PM EDT

PDB ID	:	3SM4
Title	:	Crystal Structure of the K131A Mutant of Lambda Exonuclease in Complex with a 5'-Phosphorylated 14-mer/12-mer Duplex and Magnesium
Authors	:	Bell, C.E.; Zhang, J.
Deposited on	:	2011-06-27
Resolution	:	1.88 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:

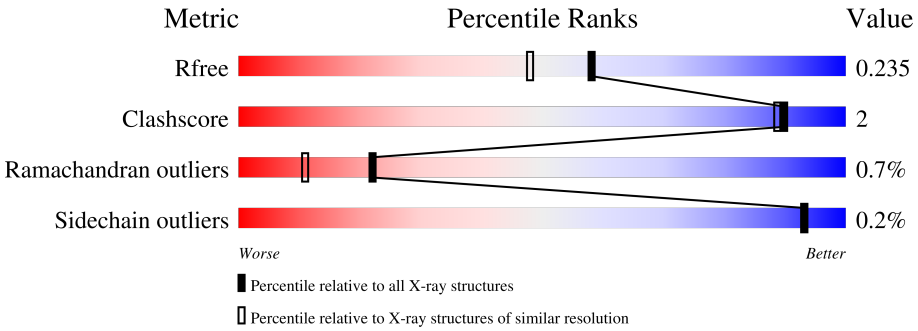
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 1.88 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	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)

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%

Mol	Chain	Length	Quality of chain
1	A	229	93%
1	B	229	96%
1	C	229	93% 7%
2	D	12	83% 17%
3	E	14	86% 14%

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 6607 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 Exonuclease.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	225	Total	C	N	O	S	0	0	0
			1808	1152	307	336	13			
1	B	228	Total	C	N	O	S	0	0	0
			1832	1166	312	340	14			
1	C	229	Total	C	N	O	S	0	0	0
			1836	1168	313	341	14			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P03697
A	-1	SER	-	expression tag	UNP P03697
A	0	HIS	-	expression tag	UNP P03697
A	131	ALA	LYS	engineered mutation	UNP P03697
B	-2	GLY	-	expression tag	UNP P03697
B	-1	SER	-	expression tag	UNP P03697
B	0	HIS	-	expression tag	UNP P03697
B	131	ALA	LYS	engineered mutation	UNP P03697
C	-2	GLY	-	expression tag	UNP P03697
C	-1	SER	-	expression tag	UNP P03697
C	0	HIS	-	expression tag	UNP P03697
C	131	ALA	LYS	engineered mutation	UNP P03697

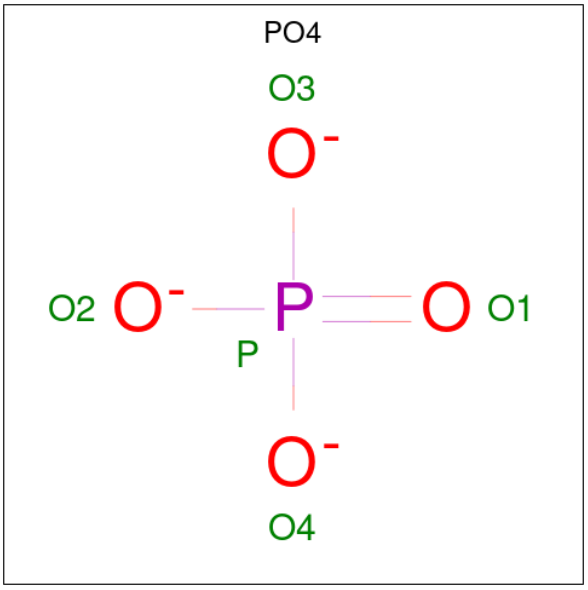
- Molecule 2 is a DNA chain called 5'-D(*TP*CP*GP*GP*TP*AP*CP*AP*GP*TP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	12	Total	C	N	O	P	0	0	0
			246	118	47	70	11			

- Molecule 3 is a DNA chain called 5'-D(P*AP*GP*CP*TP*AP*CP*TP*GP*TP*AP*CP*CP*GP*A)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	14	Total	C	N	O	P	0	0	0
			287	136	53	84	14			

- Molecule 4 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	O	P		0	0
			5	4	1			
4	C	1	Total	O	P		0	0
			5	4	1			

- Molecule 5 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Cl	0	0
			2	2		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	B	1	Total	Mg	0	0
			1	1		
6	E	1	Total	Mg	0	0
			1	1		

- Molecule 7 is water.

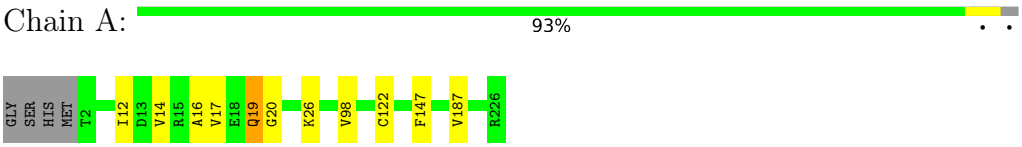
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	A	154	Total O 154 154	0	0
7	B	184	Total O 184 184	0	0
7	C	182	Total O 182 182	0	0
7	D	38	Total O 38 38	0	0
7	E	26	Total O 26 26	0	0

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. 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.

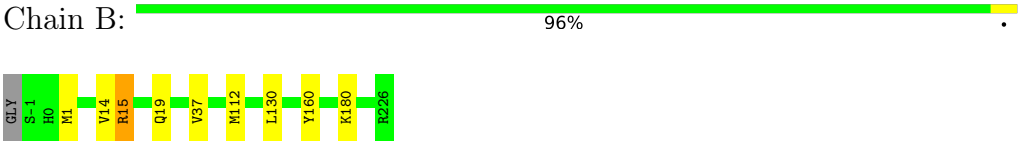
• Molecule 1: Exonuclease

Chain A:



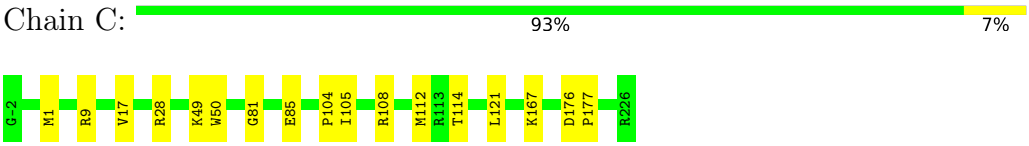
• Molecule 1: Exonuclease

Chain B:



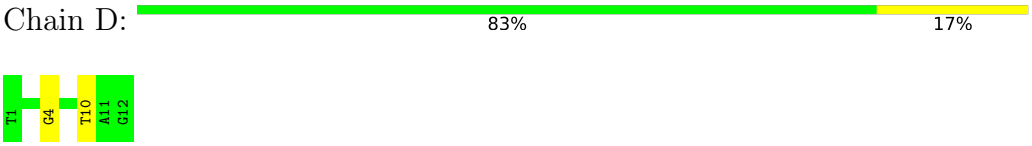
• Molecule 1: Exonuclease

Chain C:



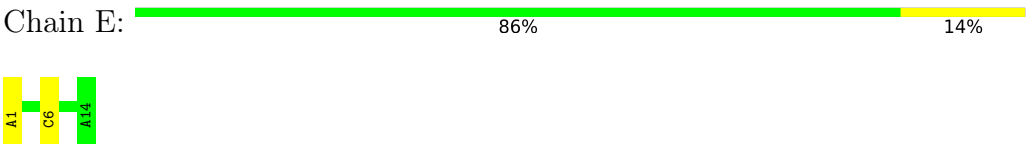
• Molecule 2: 5'-D(*TP*CP*GP*GP*TP*AP*CP*AP*GP*TP*AP*G)-3'

Chain D:



• Molecule 3: 5'-D(P*AP*GP*CP*TP*AP*CP*TP*GP*TP*AP*CP*CP*GP*A)-3'

Chain E:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65	Depositor
Cell constants a, b, c, α, β, γ	78.39Å 78.39Å 247.53Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	67.89 – 1.88 67.89 – 1.88	Depositor EDS
% Data completeness (in resolution range)	99.8 (67.89-1.88) 99.9 (67.89-1.88)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.16 (at 1.88Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.160 , 0.186 0.209 , 0.235	Depositor DCC
R_{free} test set	3584 reflections (5.15%)	wwPDB-VP
Wilson B-factor (Å^2)	23.6	Xtriage
Anisotropy	0.402	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 36.7	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	0.054 for h,-h-k,-l	Xtriage
Reported twinning fraction	0.563 for H, K, L 0.437 for -H-K, K, -L	Depositor
Outliers	0 of 69569 reflections	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	6607	wwPDB-VP
Average B, all atoms (Å^2)	24.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.84% 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

Bond lengths and bond angles in the following residue types are not validated in this section: MG, PO4, CL

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	0.33	0/1853	0.46	0/2505
1	B	0.34	0/1878	0.46	1/2538 (0.0%)
1	C	0.34	0/1882	0.46	0/2543
2	D	0.62	0/276	1.29	2/425 (0.5%)
3	E	0.86	1/321 (0.3%)	1.15	2/491 (0.4%)
All	All	0.40	1/6210 (0.0%)	0.59	5/8502 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	E	1	DA	OP3-P	-10.78	1.48	1.61

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	4	DG	O4'-C4'-C3'	-6.00	102.10	104.50
3	E	1	DA	O4'-C1'-N9	-5.85	103.90	108.00
1	B	130	LEU	CA-CB-CG	5.56	128.08	115.30
3	E	6	DC	O4'-C1'-N1	5.23	111.66	108.00
2	D	10	DT	O4'-C1'-N1	5.19	111.63	108.00

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	1808	0	1752	4	0
1	B	1832	0	1776	6	0
1	C	1836	0	1779	10	0
2	D	246	0	137	0	0
3	E	287	0	158	0	0
4	A	5	0	0	0	0
4	C	5	0	0	0	0
5	A	2	0	0	1	0
6	B	1	0	0	0	0
6	E	1	0	0	0	0
7	A	154	0	0	0	0
7	B	184	0	0	0	0
7	C	182	0	0	1	0
7	D	38	0	0	0	0
7	E	26	0	0	0	0
All	All	6607	0	5602	21	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 2.

The worst 5 of 21 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)
5:A:228:CL:CL	5:A:229:CL:CL	2.82	0.72
1:B:14:VAL:O	1:B:15:ARG:HB2	1.89	0.71
1:C:9:ARG:HD2	1:C:112:MET:SD	2.44	0.57
1:C:108:ARG:HB3	1:C:114:THR:OG1	2.05	0.56
1:B:180:LYS:HD3	7:C:331:HOH:O	2.10	0.52

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	223/229 (97%)	218 (98%)	2 (1%)	3 (1%)	12 3
1	B	226/229 (99%)	223 (99%)	1 (0%)	2 (1%)	17 7
1	C	227/229 (99%)	226 (100%)	1 (0%)	0	100 100
All	All	676/687 (98%)	667 (99%)	4 (1%)	5 (1%)	22 11

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	19	GLN
1	A	16	ALA
1	B	15	ARG
1	A	19	GLN
1	A	17	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	190/193 (98%)	189 (100%)	1 (0%)	88 88
1	B	193/193 (100%)	193 (100%)	0	100 100
1	C	193/193 (100%)	193 (100%)	0	100 100
All	All	576/579 (100%)	575 (100%)	1 (0%)	93 93

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

Mol	Chain	Res	Type
1	A	14	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	25	HIS
1	B	0	HIS
1	B	38	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)

Of 6 ligands modelled in this entry, 4 are monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	PO4	A	227	-	4,4,4	1.04	0	6,6,6	0.54	0
4	PO4	C	227	-	4,4,4	1.07	0	6,6,6	0.65	0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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

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.