wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 01:17 AM EDT

PDB ID	:	1SX5
Title	:	K38A EcoRV bound to cleaved DNA and Mn2+: P1 crystal form
Authors	:	Horton, N.C.; Perona, J.J.
Deposited on	:	2004-03-30
Resolution	:	1.50 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 (i)) were used in the production of this report:

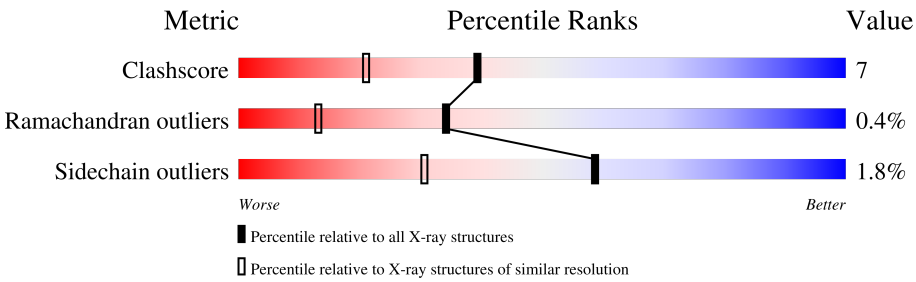
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.50 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))
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	C	6	33% 50% 17%
1	E	6	67% 17% 17%
2	D	5	20% 40% 20% 20%
2	F	5	40% 40% 20%
3	A	244	85% 15%
3	B	244	84% 12% • •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 4693 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 DNA chain called 5'-D(*AP*AP*AP*GP*AP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	6	123	60	27	31	5	0	0	0
1	E	6	123	60	27	31	5	0	0	0

- Molecule 2 is a DNA chain called 5'-D(P*AP*TP*CP*TP*T)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	D	5	101	49	14	33	5	0	0	0
2	F	5	101	49	14	33	5	0	0	0

- Molecule 3 is a protein called Type II restriction enzyme EcoRV.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	A	244	1946	1254	319	372	1	0	0	0
3	B	238	1918	1239	315	363	1	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	38	ALA	LYS	engineered mutation	UNP P04390
B	38	ALA	LYS	engineered mutation	UNP P04390

- Molecule 4 is MANGANESE (II) ION (three-letter code: MN) (formula: Mn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
4	A	5	5	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Mn		
4	B	4	4	4	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
5	C	16	16	16	0	0
			Total	O		
5	D	14	14	14	0	0
			Total	O		
5	E	12	12	12	0	0
			Total	O		
5	F	3	3	3	0	0
			Total	O		
5	A	174	174	174	0	0
			Total	O		
5	B	153	153	153	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.

Note EDS was not executed.

• Molecule 1: 5'-D(*AP*AP*AP*GP*AP*T)-3'

Chain C: 33% 50% 17%

• Molecule 1: 5'-D(*AP*AP*AP*GP*AP*T)-3'

Chain E: 67% 17% 17%

• Molecule 2: 5'-D(P*AP*TP*CP*TP*T)-3'

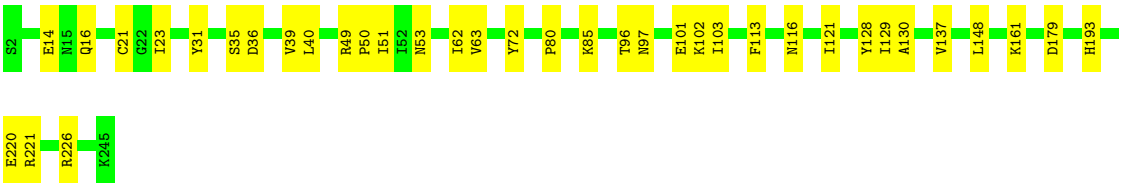
Chain D: 20% 40% 20% 20%

• Molecule 2: 5'-D(P*AP*TP*CP*TP*T)-3'

Chain F: 40% 40% 20%

• Molecule 3: Type II restriction enzyme EcoRV

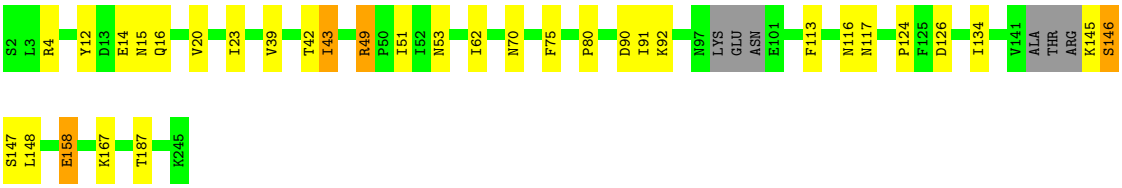
Chain A: 85% 15%



• Molecule 3: Type II restriction enzyme EcoRV

Chain B:

84%
12%


S2
L3
R4
Y12
D13
E14
N15
Q16
V20
I23
V39
T42
I43
R49
P50
I51
I52
N53
I62
N70
F75
P80
D90
I91
K92
N97
LYS
GLU
ASN
E101
F113
N116
N117
P124
F125
D126
I134
V141
ALA
THR
ARG
K145
S146
S147
L148
E158
K167
T187
K245

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	47.71Å 48.63Å 63.54Å 96.68° 108.94° 107.33°	Depositor
Resolution (Å)	20.00 – 1.50	Depositor
% Data completeness (in resolution range)	93.0 (20.00-1.50)	Depositor
R_{merge}	0.07	Depositor
R_{sym}	0.07	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.214 , 0.254	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	4693	wwPDB-VP
Average B, all atoms (Å^2)	23.0	wwPDB-VP

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: MN

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	C	0.97	0/139	1.22	1/213 (0.5%)
1	E	1.23	0/139	1.22	0/213
2	D	1.90	3/111 (2.7%)	1.68	3/167 (1.8%)
2	F	1.13	0/111	1.24	1/167 (0.6%)
3	A	0.94	1/1997 (0.1%)	0.93	2/2720 (0.1%)
3	B	1.03	0/1967	0.95	3/2671 (0.1%)
All	All	1.03	4/4464 (0.1%)	1.00	10/6151 (0.2%)

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
1	C	0	2
1	E	0	1
2	D	0	2
2	F	0	1
All	All	0	6

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	7	DA	OP3-P	-7.50	1.52	1.61
2	D	7	DA	N9-C4	-6.62	1.33	1.37
2	D	7	DA	C5-C6	-5.90	1.35	1.41
3	A	128	TYR	CE1-CZ	5.01	1.45	1.38

The worst 5 of 10 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	35	SER	N-CA-C	6.86	129.53	111.00
2	D	7	DA	O4'-C1'-N9	-6.79	103.25	108.00
3	B	49	ARG	NE-CZ-NH2	-6.61	117.00	120.30
3	A	179	ASP	CB-CG-OD1	6.35	124.02	118.30
2	D	7	DA	OP1-P-OP2	-6.10	110.46	119.60

There are no chirality outliers.

5 of 6 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	2	DA	Sidechain
1	C	6	DT	Sidechain
2	D	10	DT	Sidechain
2	D	7	DA	Sidechain
1	E	2	DA	Sidechain

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	C	123	0	69	2	0
1	E	123	0	69	3	0
2	D	101	0	59	4	0
2	F	101	0	59	1	0
3	A	1946	0	1820	29	0
3	B	1918	0	1813	24	0
4	A	5	0	0	0	0
4	B	4	0	0	0	0
5	A	174	0	0	3	0
5	B	153	0	0	6	0
5	C	16	0	0	0	0
5	D	14	0	0	1	0
5	E	12	0	0	0	0
5	F	3	0	0	0	0
All	All	4693	0	3889	60	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 7.

The worst 5 of 60 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)
3:A:148:LEU:HD13	3:B:53:ASN:ND2	1.92	0.84
3:A:96:THR:HG21	5:A:510:HOH:O	1.82	0.79
3:B:92:LYS:HG2	5:B:561:HOH:O	1.84	0.76
3:B:117:ASN:HD21	3:B:126:ASP:H	1.33	0.76
3:A:21:CYS:HB3	3:A:161:LYS:NZ	2.03	0.73

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
3	A	242/244 (99%)	232 (96%)	10 (4%)	0	100 100
3	B	232/244 (95%)	221 (95%)	9 (4%)	2 (1%)	17 3
All	All	474/488 (97%)	453 (96%)	19 (4%)	2 (0%)	34 13

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	B	146	SER
3	B	187	THR

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	
3	A	200/219 (91%)	197 (98%)	3 (2%)	65	39
3	B	199/219 (91%)	195 (98%)	4 (2%)	55	25
All	All	399/438 (91%)	392 (98%)	7 (2%)	59	30

5 of 7 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	B	43	ILE
3	B	124	PRO
3	B	167	LYS
3	B	158	GLU
3	A	97	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	B	97	ASN
3	B	116	ASN
3	B	117	ASN
3	A	238	ASN
3	A	97	ASN

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 9 ligands modelled in this entry, 9 are monoatomic - leaving 0 for Mogul analysis.

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 (i)

6.1 Protein, DNA and RNA chains (i)

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains (i)

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates (i)

EDS was not executed - this section is therefore empty.

6.4 Ligands (i)

EDS was not executed - this section is therefore empty.

6.5 Other polymers (i)

EDS was not executed - this section is therefore empty.