Full wwPDB X-ray Structure Validation Report ⓘ

May 13, 2020 – 03:07 pm BST

PDB ID	:	5MOM
Title	:	Crystal Structure of PCNA encoding the hypomorphic mutation S228I
Authors	:	Biasutto, A.J.; Mancini, E.J.; Green, C.M.; Wilson, R.H.C.
Deposited on	:	2016-12-14
Resolution	:	2.27 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 ⓘ symbol.

The following versions of software and data (see references ①) were used in the production of this report:

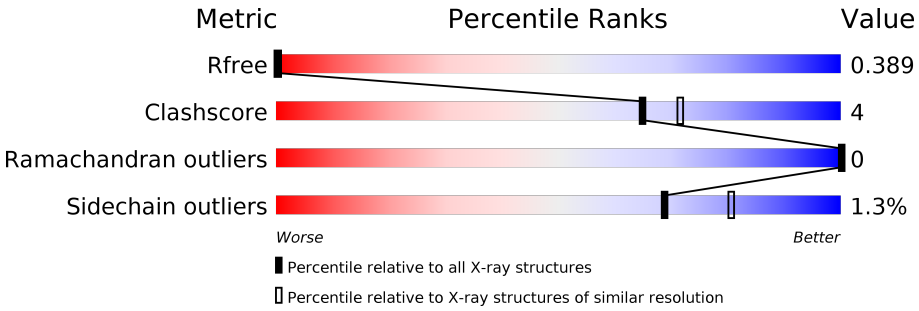
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation 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 2.27 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	6980 (2.30-2.26)
Clashscore	141614	7711 (2.30-2.26)
Ramachandran outliers	138981	7597 (2.30-2.26)
Sidechain outliers	138945	7598 (2.30-2.26)

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	258	88% 8% • •
1	B	258	88% 8% • •
1	C	258	85% 10% • •

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 5731 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 Proliferating cell nuclear antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	251	Total	C	N	O	S	0	0	0
			1869	1189	303	361	16			
1	B	249	Total	C	N	O	S	0	0	0
			1879	1193	305	365	16			
1	C	248	Total	C	N	O	S	0	0	0
			1831	1168	292	355	16			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	228	ILE	SER	engineered mutation	UNP P12004
B	228	ILE	SER	engineered mutation	UNP P12004
C	228	ILE	SER	engineered mutation	UNP P12004

- Molecule 2 is water.

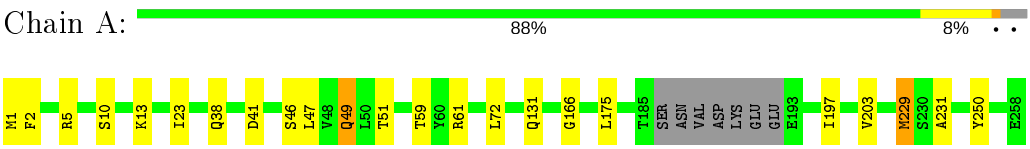
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	36	Total	O	0	0
			36	36		
2	B	75	Total	O	0	0
			75	75		
2	C	41	Total	O	0	0
			41	41		

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.

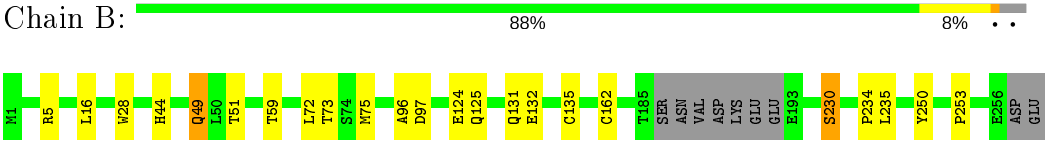
- Molecule 1: Proliferating cell nuclear antigen

Chain A:



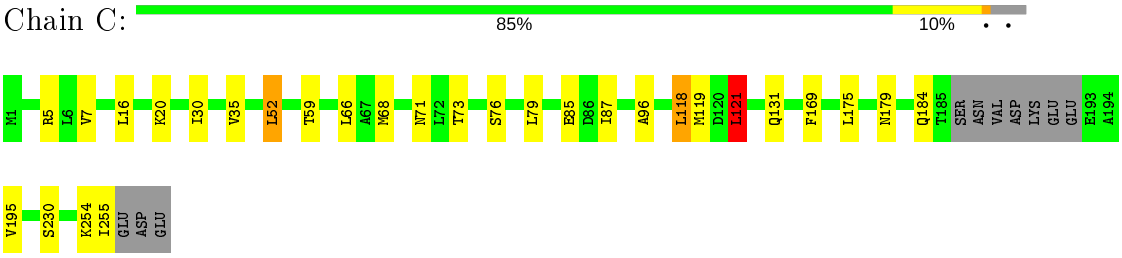
- Molecule 1: Proliferating cell nuclear antigen

Chain B:



- Molecule 1: Proliferating cell nuclear antigen

Chain C:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	162.95Å 162.95Å 140.40Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	89.07 – 2.27 89.07 – 2.27	Depositor EDS
% Data completeness (in resolution range)	100.0 (89.07-2.27) 100.0 (89.07-2.27)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.70 (at 2.27Å)	Xtriage
Refinement program	PHENIX 1.10.1_2155	Depositor
R, R_{free}	0.199 , 0.216 0.388 , 0.389	Depositor DCC
R_{free} test set	4471 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	52.6	Xtriage
Anisotropy	0.171	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 48.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.77	EDS
Total number of atoms	5731	wwPDB-VP
Average B, all atoms (Å^2)	62.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.14% 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	0.37	0/1894	0.66	1/2564 (0.0%)
1	B	0.48	0/1904	0.70	0/2575
1	C	0.38	0/1856	0.66	3/2517 (0.1%)
All	All	0.42	0/5654	0.67	4/7656 (0.1%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	175	LEU	CA-CB-CG	8.97	135.93	115.30
1	C	121	LEU	CA-CB-CG	6.05	129.21	115.30
1	C	52	LEU	CA-CB-CG	5.76	128.55	115.30
1	C	175	LEU	CA-CB-CG	5.45	127.83	115.30

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	1869	0	1851	14	0
1	B	1879	0	1884	17	0
1	C	1831	0	1798	15	0
2	A	36	0	0	3	1
2	B	75	0	0	2	0
2	C	41	0	0	2	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	5731	0	5533	46	1

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.

All (46) 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:C:230:SER:OG	2:C:301:HOH:O	1.80	0.91
1:A:131:GLN:OE1	2:A:301:HOH:O	1.99	0.80
1:B:16:LEU:HD21	1:B:75:MET:CE	2.19	0.72
1:B:16:LEU:HD21	1:B:75:MET:HE3	1.71	0.71
1:A:203:VAL:HG11	1:A:229:MET:HG2	1.72	0.71
1:B:73:THR:OG1	2:B:301:HOH:O	2.13	0.66
1:B:5:ARG:HB3	1:B:59:THR:HB	1.78	0.64
1:A:49:GLN:OE1	1:A:51:THR:HG23	2.00	0.62
1:C:66:LEU:HD13	1:C:68:MET:HG3	1.82	0.62
1:B:49:GLN:OE1	1:B:51:THR:HG23	2.06	0.56
1:C:96:ALA:HB1	1:C:118:LEU:HD23	1.88	0.56
1:A:41:ASP:OD1	1:A:46:SER:N	2.18	0.55
1:C:5:ARG:HB3	1:C:59:THR:HB	1.89	0.55
1:A:131:GLN:HE22	1:A:250:TYR:HE2	1.57	0.53
1:B:44:HIS:HE1	1:B:124:GLU:OE2	1.93	0.52
1:A:5:ARG:HB3	1:A:59:THR:HB	1.93	0.51
1:C:119:MET:HE2	1:C:121:LEU:HD11	1.93	0.51
1:A:1:MET:SD	1:A:2:PHE:N	2.83	0.51
1:B:132:GLU:HG3	1:B:230:SER:CB	2.41	0.50
1:C:71:ASN:HB2	1:C:119:MET:SD	2.52	0.49
1:B:124:GLU:O	1:B:125:GLN:HB2	2.11	0.49
1:A:166:GLY:HA2	1:A:197:ILE:HD13	1.94	0.49
1:B:16:LEU:CD2	1:B:75:MET:HE3	2.43	0.49
1:B:135:CYS:SG	1:B:162:CYS:SG	3.08	0.48
1:B:131:GLN:NE2	2:B:306:HOH:O	2.47	0.48
1:B:234:PRO:HA	1:B:253:PRO:HD3	1.96	0.48
1:B:132:GLU:HG3	1:B:230:SER:HB3	1.97	0.47
1:A:38:GLN:HG2	1:A:47:LEU:HD11	1.96	0.47
1:B:28:TRP:HE1	1:B:72:LEU:HD21	1.80	0.47
1:C:184:GLN:HG3	1:C:195:VAL:O	2.16	0.46
1:C:254:LYS:O	1:C:255:ILE:HG13	2.15	0.46
1:C:30:ILE:HD12	1:C:35:VAL:HG22	1.97	0.46
1:A:1:MET:HE1	1:A:61:ARG:HG2	1.96	0.46

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:131:GLN:OE1	2:C:302:HOH:O	2.21	0.45
1:A:231:ALA:CB	2:A:305:HOH:O	2.64	0.45
1:C:85:GLU:N	1:C:85:GLU:OE1	2.50	0.44
1:C:7:VAL:HA	1:C:87:ILE:HD13	1.98	0.44
1:A:10:SER:HA	1:A:13:LYS:HD3	2.00	0.43
1:A:231:ALA:HB3	2:A:305:HOH:O	2.19	0.42
1:B:235:LEU:O	1:B:250:TYR:HA	2.20	0.42
1:A:23:ILE:HG13	1:A:72:LEU:HD12	2.01	0.42
1:C:16:LEU:HG	1:C:79:LEU:HD12	2.02	0.42
1:B:96:ALA:O	1:B:97:ASP:HB2	2.20	0.41
1:C:20:LYS:HD3	1:C:76:SER:OG	2.20	0.41
1:B:28:TRP:CZ2	1:B:75:MET:CE	3.03	0.41
1:C:169:PHE:O	1:C:179:ASN:HA	2.21	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:328:HOH:O	2:C:329:HOH:O[8_554]	2.17	0.03

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	247/258 (96%)	238 (96%)	9 (4%)	0	100	100
1	B	245/258 (95%)	239 (98%)	6 (2%)	0	100	100
1	C	244/258 (95%)	240 (98%)	4 (2%)	0	100	100
All	All	736/774 (95%)	717 (97%)	19 (3%)	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	199/226 (88%)	197 (99%)	2 (1%)	76 86
1	B	206/226 (91%)	204 (99%)	2 (1%)	76 86
1	C	193/226 (85%)	189 (98%)	4 (2%)	53 68
All	All	598/678 (88%)	590 (99%)	8 (1%)	69 80

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	49	GLN
1	A	229	MET
1	B	49	GLN
1	B	230	SER
1	C	52	LEU
1	C	73	THR
1	C	118	LEU
1	C	121	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	38	GLN
1	B	44	HIS
1	B	200	ASN
1	C	184	GLN
1	C	200	ASN

5.3.3 RNA (i)

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.