W O R L D W I D E
wwPDB
PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Jan 3, 2024 – 06:28 pm GMT

|		
PDB ID	:	5A9B
Title	:	Crystal structure of Bombyx mori CPV1 polyhedra base domain deleted mutant
Authors	:	Ji, X.; Axford, D.; Owen, R.; Evans, G.; Ginn, H.M.; Sutton, G.; Stuart, D.I.
Deposited on	:	2015-07-17
Resolution	:	1.88 Å(reported)

1 Overall quality at a glance ⓘ

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.
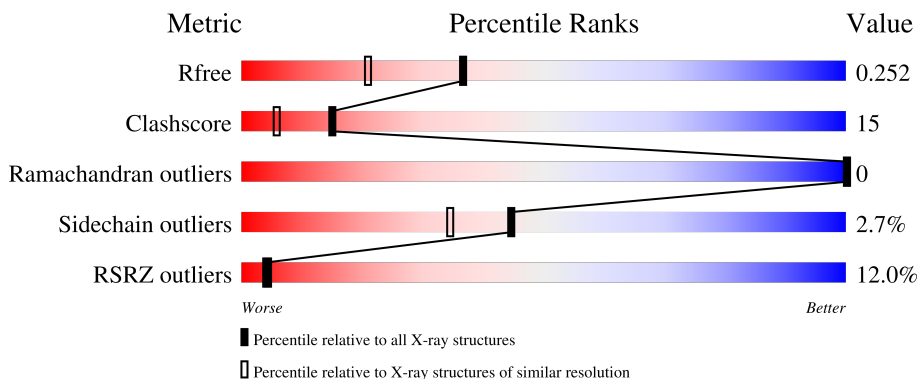
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	9470 (1.90-1.86)
Clashscore	141614	10282 (1.90-1.86)
Ramachandran outliers	138981	10152 (1.90-1.86)
Sidechain outliers	138945	10152 (1.90-1.86)
RSRZ outliers	127900	9303 (1.90-1.86)

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% 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	218	11% / 74% / 21% · ·

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 1787 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 POLYHEDRIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	209	Total	C	N	O	S	0	1	0
			1686	1069	295	318	4			

There are 38 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	29	SER	ASN	conflict	UNP P11041
A	67	VAL	ALA	conflict	UNP P11041
A	?	-	TYR	deletion	UNP P11041
A	?	-	ARG	deletion	UNP P11041
A	?	-	GLU	deletion	UNP P11041
A	?	-	GLY	deletion	UNP P11041
A	?	-	GLN	deletion	UNP P11041
A	?	-	HIS	deletion	UNP P11041
A	?	-	ASN	deletion	UNP P11041
A	?	-	ASP	deletion	UNP P11041
A	?	-	SER	deletion	UNP P11041
A	?	-	TYR	deletion	UNP P11041
A	?	-	ASP	deletion	UNP P11041
A	?	-	GLU	deletion	UNP P11041
A	?	-	TYR	deletion	UNP P11041
A	?	-	GLU	deletion	UNP P11041
A	?	-	VAL	deletion	UNP P11041
A	?	-	ASN	deletion	UNP P11041
A	?	-	GLN	deletion	UNP P11041
A	?	-	SER	deletion	UNP P11041
A	?	-	ILE	deletion	UNP P11041
A	?	-	TYR	deletion	UNP P11041
A	?	-	TYR	deletion	UNP P11041
A	?	-	PRO	deletion	UNP P11041
A	?	-	ASN	deletion	UNP P11041
A	?	-	GLY	deletion	UNP P11041
A	?	-	GLY	deletion	UNP P11041

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ASP	deletion	UNP P11041
A	?	-	ALA	deletion	UNP P11041
A	?	-	ARG	deletion	UNP P11041
A	?	-	LYS	deletion	UNP P11041
A	?	-	PHE	deletion	UNP P11041
A	?	-	HIS	deletion	UNP P11041
A	?	-	SER	deletion	UNP P11041
A	?	-	ASN	deletion	UNP P11041
A	101	GLY	-	insertion	UNP P11041
A	102	SER	-	insertion	UNP P11041
A	103	GLY	-	insertion	UNP P11041

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	101	Total 101	O 101	0	0

3 Residue-property plots ⓘ

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: POLYHEDRIN
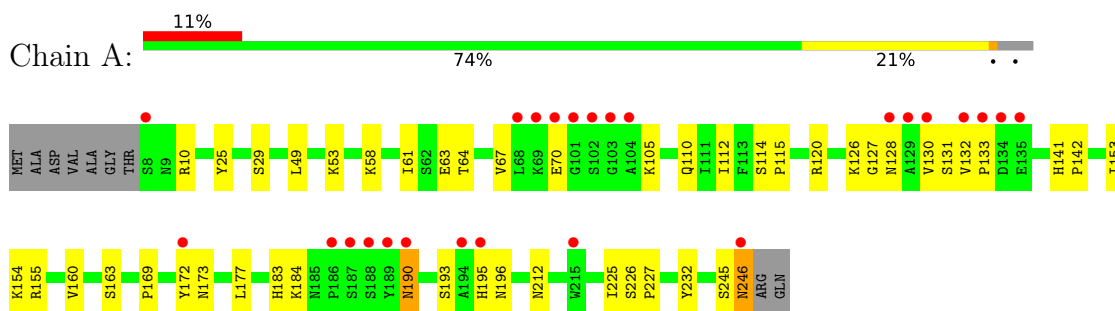
4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	I 2 3	Depositor
Cell constants a, b, c, α, β, γ	103.86Å 103.86Å 103.86Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.40 – 1.88 42.40 – 1.88	Depositor EDS
% Data completeness (in resolution range)	98.4 (42.40-1.88) 98.4 (42.40-1.88)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.22 (at 1.88Å)	Xtriage
Refinement program	PHENIX (PHENIX.REFINE)	Depositor
R, R_{free}	0.189 , 0.251 0.188 , 0.252	Depositor DCC
R_{free} test set	758 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	25.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.40 , 64.2	EDS
L-test for twinning[2]	$< |L| > = 0.47, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.060 for -l,-k,-h	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	1787	wwPDB-VP
Average B, all atoms (Å^2)	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.38% 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	0.26	0/1733	0.44	0/2350

There are no bond length outliers.

There are no bond angle outliers.

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	1686	0	1622	50	0
2	A	101	0	0	5	1
All	All	1787	0	1622	50	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 15.

The worst 5 of 50 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:A:127:GLY:CA	1:A:133:PRO:HG3	1.49	1.42
1:A:127:GLY:CA	1:A:133:PRO:CG	2.12	1.28
1:A:127:GLY:HA3	1:A:133:PRO:CG	1.77	1.14
1:A:127:GLY:HA3	1:A:133:PRO:HG3	1.25	1.12
1:A:127:GLY:HA2	1:A:133:PRO:CG	1.89	1.02

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:2059:HOH:O	2:A:2095:HOH:O[24_554]	2.02	0.18

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	208/218 (95%)	201 (97%)	7 (3%)	0	100	100

There are no Ramachandran outliers to report.

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	185/190 (97%)	180 (97%)	5 (3%)	44	34

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	58	LYS
1	A	120	ARG
1	A	190	ASN
1	A	195	HIS
1	A	246	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 monosaccharides 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 (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	209/218 (95%)	0.38	25 (11%)	4	4	16, 29, 80, 105	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	133	PRO	7.9
1	A	132	VAL	7.4
1	A	186	PRO	7.0
1	A	129	ALA	6.9
1	A	134	ASP	6.7

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.