



wwPDB X-ray Structure Validation Summary Report

Feb 26, 2014 – 11:43 PM GMT

PDB ID : 3CSQ
Title : Crystal and cryoEM structural studies of a cell wall degrading enzyme in the bacteriophage phi29 tail
Authors : Xiang, Y.; Rossmann, M.G.
Deposited on : 2008-04-10
Resolution : 1.80 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

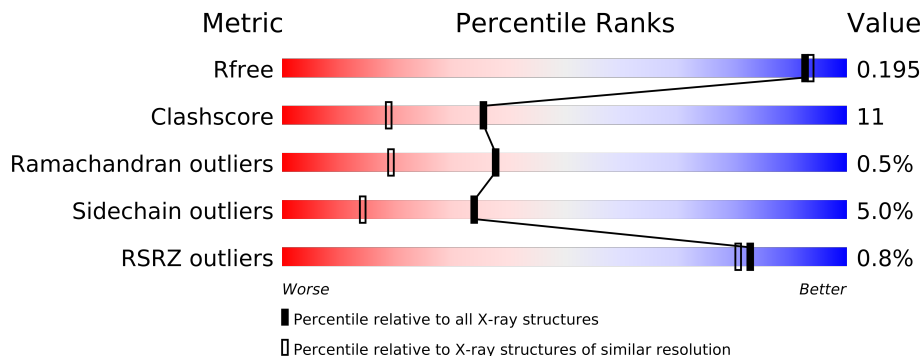
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 1.80 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}	66092	3513 (1.80-1.80)
Clashscore	79885	4461 (1.80-1.80)
Ramachandran outliers	78287	4404 (1.80-1.80)
Sidechain outliers	78261	4403 (1.80-1.80)
RSRZ outliers	66119	3515 (1.80-1.80)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	334	
1	B	334	
1	C	334	
1	D	334	

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 10684 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Morphogenesis protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	324	Total	C	N	O	S	0	3	0
			2614	1664	435	497	18			
1	B	324	Total	C	N	O	S	0	0	0
			2597	1652	433	494	18			
1	C	324	Total	C	N	O	S	0	4	0
			2616	1665	436	497	18			
1	D	326	Total	C	N	O	S	0	4	0
			2633	1676	443	496	18			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	89	ASN	ASP	SEE REMARK 999	UNP P15132
B	89	ASN	ASP	SEE REMARK 999	UNP P15132
C	89	ASN	ASP	SEE REMARK 999	UNP P15132
D	89	ASN	ASP	SEE REMARK 999	UNP P15132

- Molecule 2 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Zn	0	0
			1	1		
2	A	1	Total	Zn	0	0
			1	1		
2	D	1	Total	Zn	0	0
			1	1		
2	C	1	Total	Zn	0	0
			1	1		

- Molecule 3 is water.

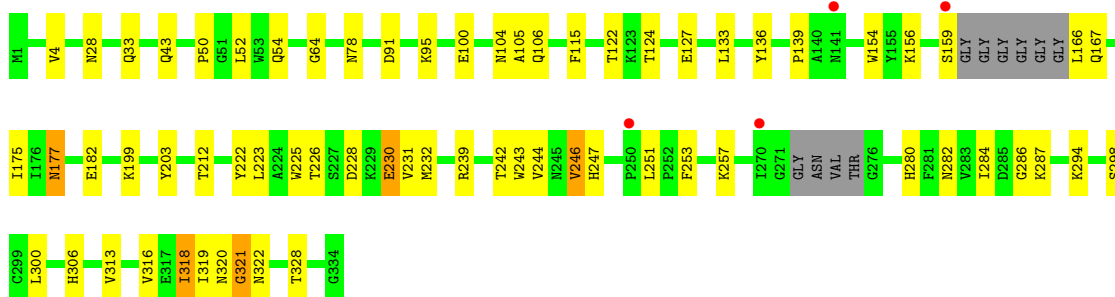
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	49	Total 49	O 49	0	0
3	B	54	Total 54	O 54	0	0
3	C	58	Total 62	O 62	0	4
3	D	55	Total 55	O 55	0	0

3 Residue-property plots

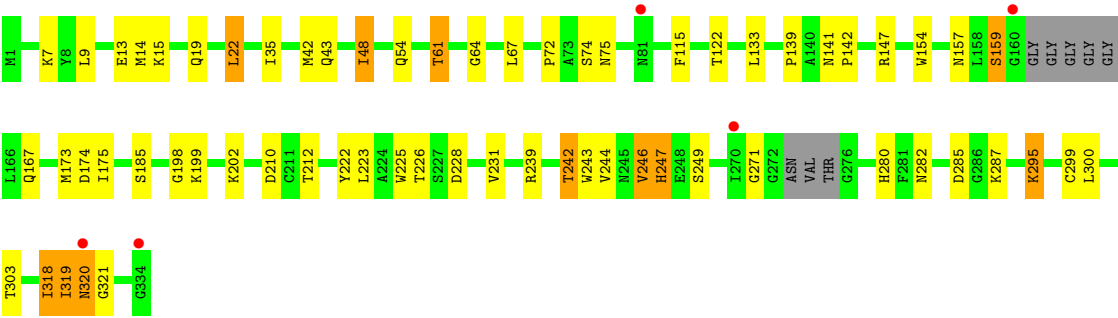
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according 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: Morphogenesis protein 1

Chain A: 



Chain D: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	53.84Å 133.96Å 85.72Å 90.00° 89.99° 90.00°	Depositor
Resolution (Å)	19.90 – 1.80 19.70 – 1.80	Depositor EDS
% Data completeness (in resolution range)	97.6 (19.90-1.80) 97.5 (19.70-1.80)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	8.84 (at 1.80Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, R_{free}	0.238 , 0.282 0.181 , 0.195	Depositor DCC
R_{free} test set	5487 reflections (5.28%)	DCC
Wilson B-factor (Å ²)	15.4	Xtriage
Anisotropy	0.012	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.41 , 21.0	EDS
Estimated twinning fraction	0.450 for h,-k,-l	Xtriage
L-test for twinning	$\langle L \rangle = 0.42$, $\langle L^2 \rangle = 0.24$	Xtriage
Outliers	0 of 109340 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	10684	wwPDB-VP
Average B, all atoms (Å ²)	18.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 48.74 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 8.2421e-05.*

¹Intensities estimated from amplitudes.

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: ZN

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	0/2694	0.76	1/3656 (0.0%)
1	B	0.72	0/2668	0.80	3/3622 (0.1%)
1	C	0.68	0/2699	0.74	0/3663
1	D	0.70	0/2716	0.74	0/3682
All	All	0.70	0/10777	0.76	4/14623 (0.0%)

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	B	0	1
1	D	0	1
All	All	0	2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	321	GLY	N-CA-C	-9.52	89.31	113.10
1	A	321	GLY	N-CA-C	-6.14	97.74	113.10
1	B	320	ASN	C-N-CA	5.70	134.26	122.30
1	B	239	ARG	NE-CZ-NH2	5.11	122.85	120.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	320	ASN	Peptide
1	D	320	ASN	Peptide

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2614	0	2492	46	0
1	B	2597	0	2467	57	0
1	C	2616	0	2497	54	0
1	D	2633	0	2525	64	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	49	0	0	5	0
3	B	54	0	0	2	0
3	C	62	0	0	6	0
3	D	55	0	0	5	0
All	All	10684	0	9981	215	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 11.

The worst 5 of 215 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:D:174:ASP:OD2	1:D:175:ILE:HD12	1.46	1.15
1:C:246:VAL:HG22	1:C:280:HIS:HB3	1.35	1.07
1:B:42:MET:HE3	1:B:67:LEU:HB3	1.40	1.04
1:B:42:MET:CE	1:B:67:LEU:HB3	1.87	1.03
1:D:42:MET:HE2	1:D:48:ILE:HG23	1.39	1.02

There are no symmetry-related clashes.

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	321/334 (96%)	305 (95%)	15 (5%)	1 (0%)	50	31
1	B	318/334 (95%)	303 (95%)	13 (4%)	2 (1%)	33	15
1	C	322/334 (96%)	311 (97%)	9 (3%)	2 (1%)	33	15
1	D	324/334 (97%)	305 (94%)	17 (5%)	2 (1%)	33	15
All	All	1285/1336 (96%)	1224 (95%)	54 (4%)	7 (0%)	38	19

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	295	LYS
1	D	247	HIS
1	A	247	HIS
1	B	58	GLU
1	C	247	HIS

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	283/283 (100%)	268 (95%)	15 (5%)	32	13
1	B	280/283 (99%)	268 (96%)	12 (4%)	40	19
1	C	284/283 (100%)	268 (94%)	16 (6%)	30	11
1	D	284/283 (100%)	270 (95%)	14 (5%)	35	15
All	All	1131/1132 (100%)	1074 (95%)	57 (5%)	34	14

5 of 57 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	328	THR
1	C	156	LYS
1	D	242	THR
1	B	331	TRP
1	C	106	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 41 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	81	ASN
1	B	282	ASN
1	D	167	GLN
1	B	83	GLN
1	B	143	ASN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

Of 4 ligands modelled in this entry, 4 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.

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

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	324/334 (97%)	-0.01	4 (1%) 75 72	9, 18, 27, 34	0
1	B	324/334 (97%)	-0.00	2 (0%) 86 85	8, 17, 26, 37	0
1	C	324/334 (97%)	-0.02	0 100 100	10, 17, 26, 34	0
1	D	326/334 (97%)	0.02	5 (1%) 70 66	9, 17, 28, 32	0
All	All	1298/1336 (97%)	-0.00	11 (0%) 83 81	8, 17, 28, 37	0

The worst 5 of 11 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	334	GLY	4.1
1	D	334	GLY	3.9
1	D	160	GLY	3.4
1	A	270	ILE	3.4
1	D	81	ASN	2.7

6.2 Non-standard residues in protein, DNA, RNA chains

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	ZN	C	335	1/1	0.07	-1.15	25,25,25,25	0
2	ZN	D	335	1/1	0.07	-1.45	28,28,28,28	0
2	ZN	A	335	1/1	0.04	-2.47	28,28,28,28	0
2	ZN	B	335	1/1	0.06	-2.72	23,23,23,23	0

6.5 Other polymers ⓘ

There are no such residues in this entry.