



Full wwPDB X-ray Structure Validation Report ⓘ

May 25, 2020 – 04:59 am BST

PDB ID : 2WBH
Title : Icosahedral particle of covalent coat protein dimer of bacteriophage MS2
Authors : Plevka, P.; Tars, K.; Liljas, L.
Deposited on : 2009-02-27
Resolution : 4.70 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 ⓘ](#)) 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

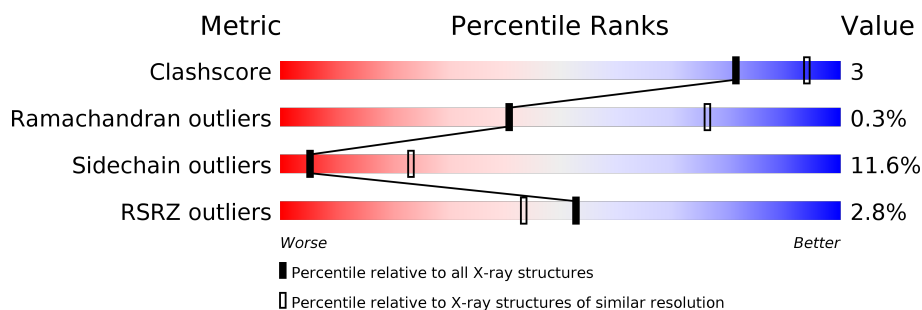
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 4.70 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	1159 (5.60-3.80)
Ramachandran outliers	138981	1094 (5.58-3.80)
Sidechain outliers	138945	1074 (5.58-3.80)
RSRZ outliers	127900	1118 (5.70-3.70)

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 $\leq 5\%$. 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	257	
1	B	257	
1	C	257	

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 2892 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 COAT PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	129	Total	C	N	O	S	0	0	0
			964	606	165	189	4			
1	B	129	Total	C	N	O	S	0	0	0
			964	606	165	189	4			
1	C	129	Total	C	N	O	S	0	0	0
			964	606	165	189	4			

- Molecule 1: COAT PROTEIN



GLY
ILE
TYR

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 3	Depositor
Cell constants a, b, c, α , β , γ	368.20 Å 368.20 Å 368.20 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 4.70 49.20 – 4.31	Depositor EDS
% Data completeness (in resolution range)	84.0 (50.00-4.70) 71.2 (49.20-4.31)	Depositor EDS
R_{merge}	0.25	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.00 (at 4.29 Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.318 , (Not available) 0.270 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	131.9	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 131.1	EDS
L-test for twinning ²	$\langle L \rangle = 0.34$, $\langle L^2 \rangle = 0.18$	Xtriage
Estimated twinning fraction	0.119 for l,-k,h	Xtriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	2892	wwPDB-VP
Average B, all atoms (Å ²)	138.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 1.59% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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	1.01	0/981	1.65	19/1337 (1.4%)
1	B	1.02	0/981	1.52	18/1337 (1.3%)
1	C	0.95	0/981	1.67	22/1337 (1.6%)
All	All	0.99	0/2943	1.62	59/4011 (1.5%)

There are no bond length outliers.

All (59) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	56	ARG	NE-CZ-NH2	-20.57	110.01	120.30
1	C	11	ASP	CB-CA-C	-10.57	89.25	110.40
1	C	11	ASP	N-CA-CB	10.18	128.92	110.60
1	A	56	ARG	NE-CZ-NH1	9.16	124.88	120.30
1	C	32	TRP	CE2-CD2-CG	-8.66	100.37	107.30
1	C	56	ARG	NE-CZ-NH2	-8.63	115.98	120.30
1	B	32	TRP	CD1-CG-CD2	8.20	112.86	106.30
1	C	49	ARG	NE-CZ-NH2	-8.07	116.27	120.30
1	A	32	TRP	CD1-CG-CD2	8.06	112.75	106.30
1	C	32	TRP	CG-CD2-CE3	7.92	141.03	133.90
1	B	82	TRP	CD1-CG-CD2	7.86	112.59	106.30
1	C	32	TRP	CD1-CG-CD2	7.83	112.56	106.30
1	B	83	ARG	NE-CZ-NH2	-7.78	116.41	120.30
1	A	82	TRP	CD1-CG-CD2	7.76	112.51	106.30
1	C	82	TRP	CD1-CG-CD2	7.66	112.42	106.30
1	A	32	TRP	CE2-CD2-CG	-7.60	101.22	107.30
1	C	32	TRP	CB-CG-CD1	-7.40	117.38	127.00
1	B	32	TRP	CE2-CD2-CG	-7.40	101.38	107.30
1	C	108	MET	CG-SD-CE	-7.32	88.49	100.20
1	A	2	SER	N-CA-C	-7.24	91.44	111.00
1	B	48	VAL	CB-CA-C	-7.13	97.86	111.40
1	C	82	TRP	CE2-CD2-CG	-7.08	101.64	107.30
1	A	82	TRP	CE2-CD2-CG	-7.07	101.64	107.30
1	C	58	TYR	CB-CG-CD2	-6.86	116.88	121.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	82	TRP	CE2-CD2-CG	-6.64	101.99	107.30
1	C	20	VAL	N-CA-CB	-6.62	96.92	111.50
1	B	20	VAL	N-CA-CB	-6.46	97.28	111.50
1	B	2	SER	N-CA-C	-6.35	93.84	111.00
1	B	33	ILE	CA-CB-CG1	-6.29	99.05	111.00
1	A	49	ARG	NE-CZ-NH1	6.28	123.44	120.30
1	A	20	VAL	N-CA-CB	-6.21	97.85	111.50
1	C	10	VAL	CA-C-N	6.03	130.47	117.20
1	A	79	VAL	N-CA-CB	-5.99	98.32	111.50
1	A	32	TRP	CB-CG-CD1	-5.96	119.25	127.00
1	B	32	TRP	CG-CD1-NE1	-5.91	104.19	110.10
1	B	50	GLN	CA-CB-CG	-5.84	100.56	113.40
1	B	82	TRP	CG-CD1-NE1	-5.75	104.35	110.10
1	C	11	ASP	CA-CB-CG	5.65	125.83	113.40
1	A	55	ASN	CB-CA-C	-5.64	99.11	110.40
1	A	32	TRP	CG-CD1-NE1	-5.62	104.48	110.10
1	B	38	ARG	NE-CZ-NH1	5.62	123.11	120.30
1	B	20	VAL	CB-CA-C	5.62	122.07	111.40
1	C	56	ARG	NE-CZ-NH1	5.56	123.08	120.30
1	C	20	VAL	CB-CA-C	5.53	121.90	111.40
1	B	32	TRP	CG-CD2-CE3	5.50	138.85	133.90
1	C	101	CYS	CA-CB-SG	-5.44	104.21	114.00
1	A	35	SER	N-CA-CB	-5.43	102.36	110.50
1	B	57	LYS	CB-CA-C	-5.43	99.55	110.40
1	A	38	ARG	CG-CD-NE	-5.42	100.41	111.80
1	B	32	TRP	CB-CG-CD1	-5.41	119.96	127.00
1	A	32	TRP	CG-CD2-CE3	5.37	138.74	133.90
1	C	49	ARG	N-CA-CB	-5.32	101.03	110.60
1	B	6	GLN	N-CA-CB	-5.27	101.11	110.60
1	C	6	GLN	N-CA-CB	-5.26	101.14	110.60
1	A	56	ARG	NH1-CZ-NH2	5.19	125.11	119.40
1	A	82	TRP	CB-CG-CD1	-5.15	120.30	127.00
1	A	48	VAL	CB-CA-C	-5.14	101.63	111.40
1	C	84	SER	N-CA-CB	5.14	118.21	110.50
1	C	48	VAL	CB-CA-C	-5.00	101.89	111.40

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	964	0	964	7	0
1	B	964	0	964	4	0
1	C	964	0	964	6	0
All	All	2892	0	2892	15	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 3.

All (15) 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:10:VAL:HG12	1:C:12:ASN:HB2	1.79	0.64
1:A:97:THR:HG22	1:A:99:SER:H	1.63	0.63
1:A:4:PHE:O	1:C:1:ALA:HB3	1.99	0.62
1:B:20:VAL:HG13	1:B:32:TRP:HB3	1.83	0.59
1:C:20:VAL:HG13	1:C:32:TRP:HB3	1.89	0.54
1:A:20:VAL:HG13	1:A:32:TRP:HB3	1.88	0.54
1:B:105:VAL:O	1:B:109:GLN:HG2	2.08	0.54
1:A:97:THR:HG22	1:A:99:SER:N	2.25	0.51
1:A:60:ILE:HB	1:A:88:MET:HG3	1.93	0.51
1:A:105:VAL:O	1:A:109:GLN:HG2	2.15	0.46
1:C:60:ILE:HB	1:C:88:MET:HG3	1.98	0.45
1:C:20:VAL:HG13	1:C:32:TRP:CB	2.46	0.45
1:A:37:SER:HA	1:C:96:ALA:O	2.20	0.42
1:B:62:VAL:HB	1:B:86:LEU:HB3	2.02	0.41
1:B:33:ILE:HD12	1:B:38:ARG:CZ	2.50	0.41

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	127/257 (49%)	123 (97%)	4 (3%)	0	100	100
1	B	127/257 (49%)	123 (97%)	3 (2%)	1 (1%)	19	60
1	C	127/257 (49%)	123 (97%)	4 (3%)	0	100	100
All	All	381/771 (49%)	369 (97%)	11 (3%)	1 (0%)	41	76

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	2	SER

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	106/211 (50%)	91 (86%)	15 (14%)	3	18
1	B	106/211 (50%)	96 (91%)	10 (9%)	8	30
1	C	106/211 (50%)	94 (89%)	12 (11%)	6	24
All	All	318/633 (50%)	281 (88%)	37 (12%)	5	23

All (37) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	SER
1	A	4	PHE
1	A	6	GLN
1	A	8	VAL
1	A	19	THR
1	A	20	VAL
1	A	25	PHE
1	A	33	ILE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	48	VAL
1	A	50	GLN
1	A	52	SER
1	A	79	VAL
1	A	109	GLN
1	A	126	SER
1	A	128	ILE
1	B	4	PHE
1	B	20	VAL
1	B	33	ILE
1	B	35	SER
1	B	38	ARG
1	B	48	VAL
1	B	54	GLN
1	B	77	LEU
1	B	103	LEU
1	B	109	GLN
1	C	11	ASP
1	C	20	VAL
1	C	23	SER
1	C	37	SER
1	C	38	ARG
1	C	49	ARG
1	C	66	LYS
1	C	76	GLU
1	C	84	SER
1	C	87	ASN
1	C	109	GLN
1	C	114	ASP

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

Mol	Chain	Res	Type
1	A	40	GLN
1	A	109	GLN
1	B	36	ASN

5.3.3 RNA ⓘ

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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

There are no ligands in this entry.

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](#)

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	129/257 (50%)	0.21	4 (3%)	49 39	136, 136, 136, 136	0
1	B	129/257 (50%)	0.16	2 (1%)	72 63	140, 140, 140, 140	0
1	C	129/257 (50%)	0.23	5 (3%)	39 32	140, 140, 140, 140	0
All	All	387/771 (50%)	0.20	11 (2%)	53 43	136, 140, 140, 140	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	ALA	4.0
1	A	129	TYR	3.2
1	A	128	ILE	3.2
1	C	4	PHE	2.8
1	C	111	LEU	2.7
1	A	43	LYS	2.4
1	B	66	LYS	2.3
1	C	108	MET	2.2
1	B	4	PHE	2.2
1	C	104	ILE	2.0
1	C	63	GLU	2.0

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 carbohydrates 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.