wwPDB X-ray Structure Validation Summary Report (i)

Oct 18, 2023 – 01:25 AM EDT

PDB ID	:	2B2G
Title	:	MS2 Wild-type RNA stemloop complexed with an N87S mutant MS2 capsid
Authors	:	Horn, W.T.; Tars, K.; Grahn, E.; Helgstrand, C.; Baron, A.J.; Lago, H.; Adams, C.J.; Peabody, D.S.; Phillips, S.E.V.; Stonehouse, N.J.; Liljas, L.; Stockley, P.G.
Deposited on	:	2005-09-19
Resolution	:	3.02 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 types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

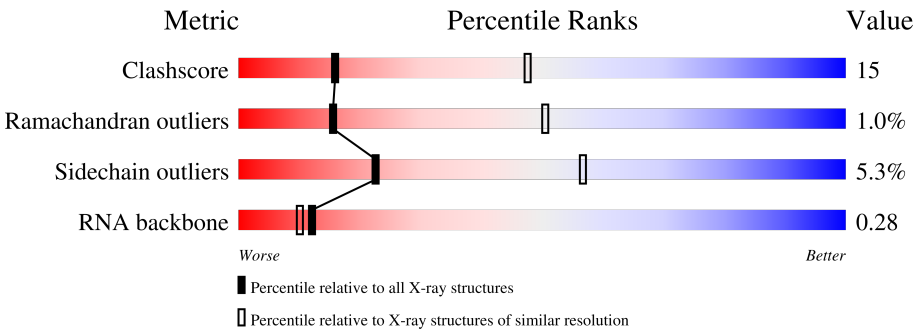
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 3.02 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	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RNA backbone	3102	1066 (3.30-2.74)

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	R	19	11% 42% 21% 11% 16%
1	S	19	5% 32% 42% 21%
2	A	129	72% 26% •
2	B	129	72% 24% •
2	C	129	83% 16% •

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 3726 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 RNA chain called 5'-R(*AP*CP*AP*UP*GP*AP*GP*GP*AP*UP*UP*AP*CP*CP*CP*AP*UP*GP*U)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	R	16	342	153	62	111	16	0	0	0
1	S	15	319	143	57	104	15	0	0	0

- Molecule 2 is a protein called Coat protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	129	962	605	164	189	4	0	0	0
2	B	129	962	605	164	189	4	0	0	0
2	C	129	962	605	164	189	4	0	0	0

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	87	SER	ASN	engineered mutation	UNP P03612
B	87	SER	ASN	engineered mutation	UNP P03612
C	87	SER	ASN	engineered mutation	UNP P03612

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	R	4	4	4	0	0
3	S	17	17	17	0	0
3	A	81	81	81	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	B	20	20	20	0	0
			Total	O		
3	C	57	57	57	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'-R(*AP*CP*AP*UP*GP*AP*GP*GP*AP*UP*UP*AP*CP*CP*CP*AP*UP*GP*U)-3'

Chain R: 11% 42% 21% 11% 16%

A C A302 U303 G304 A305 G306 G307 A308 U309 U310 A311 C312 C313 C314 A315 U316 G317 U

• Molecule 1: 5'-R(*AP*CP*AP*UP*GP*AP*GP*GP*AP*UP*UP*AP*CP*CP*CP*AP*UP*GP*U)-3'

Chain S: 5% 32% 42% 21%

A C A402 U403 G404 A405 G406 G407 A408 U409 U410 A411 C412 C413 C414 A415 U416 G U

• Molecule 2: Coat protein

Chain A: 72% 26% •

A1 S2 N3 F4 V8 L9 V10 V18 T19 V20 S23 N24 F25 E31 W32 I33 K43 V44 T45 V48 R49 Q50 S51 S52 A53 Q54 K61 V79 A80 L86 F95 A96 T97 N98 S99 L103 I104 V105 Q109 L112 P117 I128 Y129

• Molecule 2: Coat protein

Chain B: 72% 24% •

A1 S2 N3 F4 V10 T15 V18 T19 V20 A21 P22 W32 I33 S34 S35 N36 S37 R38 V48 R49 Q50 S51 S52 A53 Q54 K57 V62 E63 K66 L77 W82 R83 L86 E89 V105 M108 Q109 L112 K113 D114 P117 N125

Y129

• Molecule 2: Coat protein

Chain C: 83% 16% •

A1
V10
T15
G16
D17
V18
T19
V20
W32
N36
S37
R38
R49
R56
K61
Q70
E76
S87
V105
Q109
G110
L111
D114
I118
N125
Y129

4 Data and refinement statistics (i)

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

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, α, β, γ	287.45Å 287.45Å 651.65Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	30.00 – 3.02	Depositor
% Data completeness (in resolution range)	(Not available) (30.00-3.02)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	0.19	Depositor
Refinement program	CNS	Depositor
R, R_{free}	0.202 , 0.209	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	3726	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

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	R	2.10	6/382 (1.6%)	2.02	15/593 (2.5%)
1	S	3.10	34/356 (9.6%)	2.76	42/552 (7.6%)
2	A	0.39	0/979	0.69	1/1334 (0.1%)
2	B	0.38	0/979	0.70	0/1334
2	C	0.37	0/979	0.66	0/1334
All	All	1.23	40/3675 (1.1%)	1.29	58/5147 (1.1%)

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	R	0	1
1	S	0	1
All	All	0	2

The worst 5 of 40 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	R	305	A	N1-C2	-21.16	1.15	1.34
1	R	305	A	C5-C4	-19.24	1.25	1.38
1	R	305	A	N3-C4	14.10	1.43	1.34
1	R	305	A	N9-C8	-13.77	1.26	1.37
1	S	404	G	O3'-P	13.62	1.77	1.61

The worst 5 of 58 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	R	305	A	C2-N3-C4	-24.10	98.55	110.60
1	R	305	A	N1-C2-N3	16.93	137.77	129.30
1	S	409	U	O4'-C1'-N1	15.71	120.77	108.20
1	S	408	A	P-O3'-C3'	13.65	136.08	119.70

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	S	410	U	C4'-C3'-C2'	12.04	114.64	102.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	R	305	A	Sidechain
1	S	404	G	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	R	342	0	173	17	0
1	S	319	0	162	22	0
2	A	962	0	963	31	0
2	B	962	0	963	29	0
2	C	962	0	963	12	0
3	A	81	0	0	4	0
3	B	20	0	0	2	0
3	C	57	0	0	3	0
3	R	4	0	0	0	0
3	S	17	0	0	0	0
All	All	3726	0	3224	97	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 15.

The worst 5 of 97 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:S:404:G:H5'	1:S:405:A:OP1	1.71	0.89
1:S:404:G:H5'	1:S:405:A:P	2.22	0.79
2:A:79:VAL:HG22	2:A:80:ALA:H	1.52	0.75
2:C:111:LEU:HD11	2:C:118:ILE:HD12	1.71	0.73
2:A:128:ILE:HA	2:B:3:ASN:HD21	1.57	0.70

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	
2	A	127/129 (98%)	119 (94%)	6 (5%)	2 (2%)	9	38
2	B	127/129 (98%)	121 (95%)	5 (4%)	1 (1%)	19	55
2	C	127/129 (98%)	113 (89%)	13 (10%)	1 (1%)	19	55
All	All	381/387 (98%)	353 (93%)	24 (6%)	4 (1%)	15	50

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	B	2	SER
2	A	2	SER
2	A	51	SER
2	C	37	SER

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	
2	A	106/106 (100%)	99 (93%)	7 (7%)	16	47
2	B	106/106 (100%)	99 (93%)	7 (7%)	16	47
2	C	106/106 (100%)	103 (97%)	3 (3%)	43	76
All	All	318/318 (100%)	301 (95%)	17 (5%)	22	57

5 of 17 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	C	38	ARG
2	C	114	ASP
2	B	3	ASN
2	B	38	ARG
2	B	48	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	6	GLN
2	B	36	ASN
2	C	12	ASN
2	B	54	GLN
2	B	3	ASN

5.3.3 RNA (i)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	R	15/19 (78%)	6 (40%)	2 (13%)
1	S	15/19 (78%)	8 (53%)	9 (60%)
All	All	30/38 (78%)	14 (46%)	11 (36%)

5 of 14 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	R	305	A
1	R	307	G
1	R	309	U
1	R	310	U
1	R	311	A

5 of 11 RNA pucker outliers are listed below:

Mol	Chain	Res	Type
1	S	409	U
1	S	410	U
1	S	415	A
1	S	411	A
1	S	404	G

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)

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)

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	S	2

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	S	404:G	O3'	405:A	P	1.77
1	S	405:A	O3'	406:G	P	1.76

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.