Full wwPDB X-ray Structure Validation Report (i)

Oct 10, 2021 – 04:44 PM EDT

PDB ID	:	3GS2
Title	:	Ring1B C-terminal domain/Cbx7 Cbox Complex
Authors	:	Wang, R.; Taylor, A.B.; Kim, C.A.
Deposited on	:	2009-03-26
Resolution	:	1.70 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 (i) symbol.

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

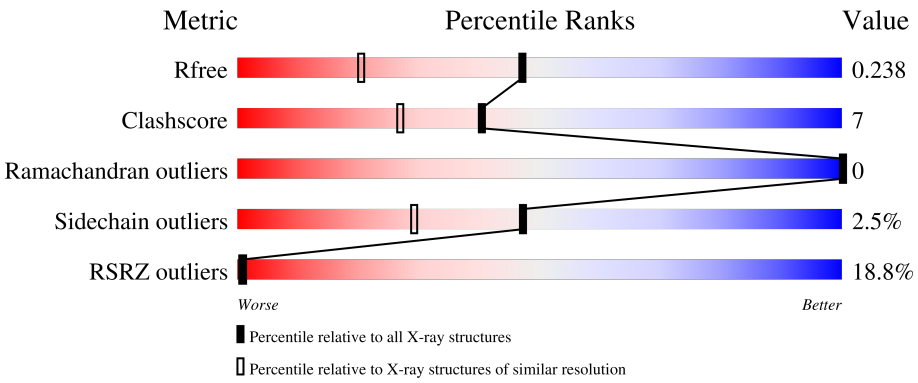
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.23.2
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 1.70 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))
R_{free}	130704	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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 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	111	11%; 87%; 11%; •
1	C	111	26%; 72%; 24%; • •
2	B	30	20%; 90%; 10%
2	D	30	17%; 90%; 7%; •

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 2415 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 E3 ubiquitin-protein ligase RING2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	109	888	564	143	178	3	0	4	0
1	C	108	872	555	142	172	3	0	2	0

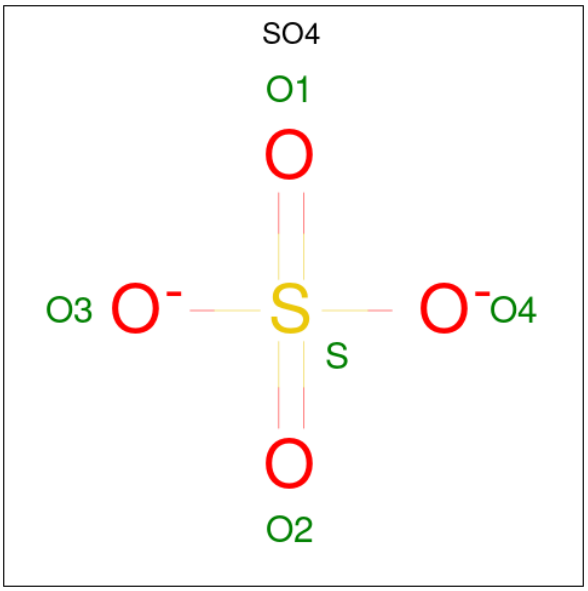
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	306	ASP	ASN	engineered mutation	UNP Q99496
C	306	ASP	ASN	engineered mutation	UNP Q99496

- Molecule 2 is a protein called Chromobox protein homolog 7.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	B	30	234	144	41	49	0	0	0
2	D	29	228	141	40	47	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0

- Molecule 4 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Zn 1 1	0	0

- Molecule 5 is water.

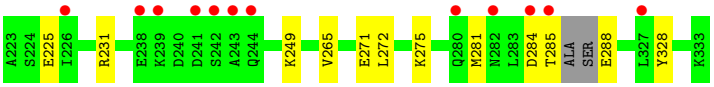
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	93	Total O 93 93	0	0
5	B	21	Total O 21 21	0	0
5	C	49	Total O 49 49	0	0
5	D	19	Total O 19 19	0	0

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 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.

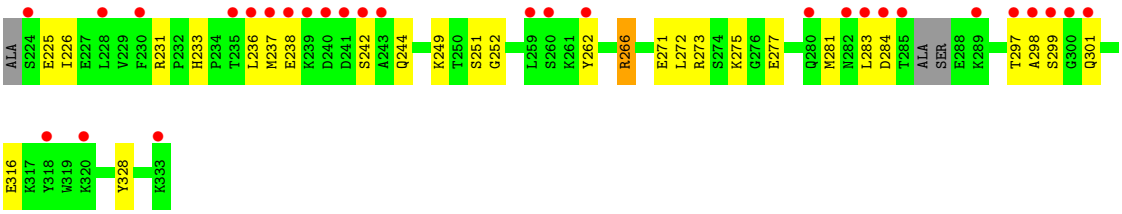
• Molecule 1: E3 ubiquitin-protein ligase RING2

Chain A:



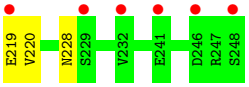
• Molecule 1: E3 ubiquitin-protein ligase RING2

Chain C:



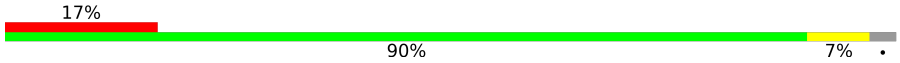
• Molecule 2: Chromobox protein homolog 7

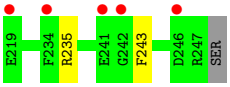
Chain B:



• Molecule 2: Chromobox protein homolog 7

Chain D:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	44.76Å 50.55Å 124.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	21.05 – 1.70 21.05 – 1.70	Depositor EDS
% Data completeness (in resolution range)	95.1 (21.05-1.70) 95.1 (21.05-1.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$< I/\sigma(I) >$ [1]	2.09 (at 1.69Å)	Xtriage
Refinement program	PHENIX	Depositor
R, R_{free}	0.206 , 0.243 0.201 , 0.238	Depositor DCC
R_{free} test set	1585 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	24.7	Xtriage
Anisotropy	0.300	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 47.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2415	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.63% 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)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, 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	0.34	0/916	0.53	0/1235
1	C	0.31	0/894	0.51	0/1207
2	B	0.34	0/236	0.49	0/319
2	D	0.33	0/230	0.49	0/311
All	All	0.33	0/2276	0.52	0/3072

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	888	0	887	11	0
1	C	872	0	868	21	0
2	B	234	0	222	2	0
2	D	228	0	217	2	0
3	A	5	0	0	1	0
3	B	5	0	0	0	0
4	A	1	0	0	0	0
5	A	93	0	0	1	0
5	B	21	0	0	4	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
5	C	49	0	0	5	0
5	D	19	0	0	2	0
All	All	2415	0	2194	32	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 7.

All (32) 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:297:THR:HG22	5:C:173:HOH:O	1.82	0.79
1:A:275:LYS:HD3	5:A:168:HOH:O	1.97	0.65
1:A:272:LEU:HD12	1:C:272:LEU:HD12	1.81	0.63
1:A:281:MET:HG2	5:D:44:HOH:O	2.01	0.61
1:C:249:LYS:HE3	2:D:235:ARG:HH21	1.66	0.60
5:B:170:HOH:O	1:C:277:GLU:HG2	2.05	0.57
1:C:233:HIS:O	1:C:237:MET:HG3	2.07	0.55
2:B:228:ASN:ND2	5:B:127:HOH:O	2.41	0.53
1:C:226:ILE:HG13	1:C:252:GLY:HA2	1.91	0.52
5:B:170:HOH:O	1:C:277:GLU:HA	2.10	0.51
3:A:1:SO4:O4	1:C:273:ARG:NH2	2.42	0.51
1:C:249:LYS:HE2	2:D:243:PHE:CE1	2.45	0.50
1:C:262:TYR:CZ	1:C:266:ARG:HG3	2.46	0.50
1:C:275:LYS:HZ3	1:C:284:ASP:CG	2.16	0.49
1:A:284:ASP:O	1:A:285:THR:HB	2.13	0.49
1:C:283:LEU:N	5:C:133:HOH:O	2.42	0.48
1:A:225:GLU:OE2	1:A:249:LYS:HD3	2.13	0.48
1:C:281:MET:HG2	5:C:171:HOH:O	2.13	0.48
1:A:231:ARG:O	1:A:328:TYR:HA	2.13	0.48
1:C:275:LYS:NZ	1:C:284:ASP:OD1	2.48	0.47
1:C:233:HIS:CE1	1:C:236:LEU:HD13	2.51	0.46
1:C:298:ALA:HA	1:C:299:SER:HA	1.55	0.46
1:C:242:SER:OG	1:C:244:GLN:HG2	2.15	0.46
1:A:284:ASP:CG	1:A:285:THR:H	2.20	0.45
1:C:231:ARG:O	1:C:328:TYR:HA	2.15	0.45
1:A:271:GLU:OE2	1:A:275:LYS:HE2	2.17	0.44
2:B:219:GLU:HG2	2:B:220:VAL:H	1.83	0.43
5:B:170:HOH:O	1:C:277:GLU:CG	2.65	0.43
1:A:288:GLU:HG3	5:C:145:HOH:O	2.19	0.42
1:A:265:VAL:HG22	1:C:283:LEU:HD21	2.02	0.42
1:C:316:GLU:HG2	5:C:174:HOH:O	2.20	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:281:MET:CG	5:D:44:HOH:O	2.64	0.40

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	
1	A	109/111 (98%)	107 (98%)	2 (2%)	0	100	100
1	C	106/111 (96%)	100 (94%)	6 (6%)	0	100	100
2	B	28/30 (93%)	27 (96%)	1 (4%)	0	100	100
2	D	27/30 (90%)	26 (96%)	1 (4%)	0	100	100
All	All	270/282 (96%)	260 (96%)	10 (4%)	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	101/98 (103%)	101 (100%)	0	100	100
1	C	98/98 (100%)	91 (93%)	7 (7%)	14	3
2	B	25/25 (100%)	25 (100%)	0	100	100
2	D	24/25 (96%)	24 (100%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	248/246 (101%)	241 (97%)	7 (3%)	47 25

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	225	GLU
1	C	238	GLU
1	C	251[A]	SER
1	C	251[B]	SER
1	C	266	ARG
1	C	271	GLU
1	C	301	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	228	ASN

5.3.3 RNA (i)

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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 3 ligands modelled in this entry, 1 is monoatomic - leaving 2 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
3	SO4	B	2	-	4,4,4	0.11	0	6,6,6	0.15	0
3	SO4	A	1	-	4,4,4	0.16	0	6,6,6	0.14	0

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.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1	SO4	1	0

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	109/111 (98%)	0.50	12 (11%)	5	6	14, 22, 56, 66	0
1	C	108/111 (97%)	1.29	29 (26%)	0	0	18, 29, 63, 74	0
2	B	30/30 (100%)	0.78	6 (20%)	1	0	16, 26, 50, 58	0
2	D	29/30 (96%)	0.84	5 (17%)	1	1	20, 27, 48, 62	0
All	All	276/282 (97%)	0.88	52 (18%)	1	1	14, 27, 61, 74	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	300	GLY	9.4
1	C	285	THR	8.3
1	C	299	SER	8.1
1	C	283	LEU	6.6
1	C	282	ASN	5.4
1	C	237	MET	5.0
2	B	248	SER	4.8
1	C	333	LYS	4.8
1	C	240	ASP	4.6
1	C	224	SER	4.2
1	C	284	ASP	4.0
2	D	246	ASP	4.0
1	A	282	ASN	3.8
1	A	284	ASP	3.7
1	C	238	GLU	3.7
1	A	242[A]	SER	3.7
1	C	298	ALA	3.6
1	C	239	LYS	3.5
1	C	236	LEU	3.5
1	A	285	THR	3.3
1	A	241	ASP	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	244	GLN	3.0
2	D	234	PHE	3.0
1	A	239	LYS	2.9
2	B	219	GLU	2.9
1	C	259	LEU	2.9
2	B	241	GLU	2.9
1	C	235	THR	2.8
2	B	229	SER	2.7
1	C	243	ALA	2.6
2	B	246	ASP	2.6
1	C	262	TYR	2.6
1	C	280	GLN	2.6
1	C	228	LEU	2.5
1	C	242	SER	2.5
1	C	241	ASP	2.5
1	A	243	ALA	2.4
1	C	230	PHE	2.4
2	D	242	GLY	2.4
2	D	219	GLU	2.4
1	A	327	LEU	2.3
1	A	280	GLN	2.3
1	A	226	ILE	2.3
1	C	289	LYS	2.2
1	C	320	LYS	2.2
1	A	238	GLU	2.2
1	C	318	TYR	2.2
1	C	297	THR	2.1
1	C	260	SER	2.1
1	C	301	GLN	2.1
2	D	241	GLU	2.0
2	B	232	VAL	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 monosaccharides in this entry.

6.4 Ligands (i)

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. 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	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	SO4	B	2	5/5	0.91	0.14	25,26,34,41	5
3	SO4	A	1	5/5	0.96	0.08	29,30,34,36	5
4	ZN	A	334	1/1	0.99	0.05	22,22,22,22	1

6.5 Other polymers (i)

There are no such residues in this entry.