wwPDB X-ray Structure Validation Summary Report (i)

May 29, 2020 – 08:55 am BST

PDB ID	:	5XYN
Title	:	The crystal structure of Csm2-Psy3-Shu1-Shu2 complex from budding yeast
Authors	:	Zhang, S.; Zhang, T.; Ding, J.
Deposited on	:	2017-07-09
Resolution	:	3.30 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 following versions of software and data (see references (1)) were used in the production of this report:

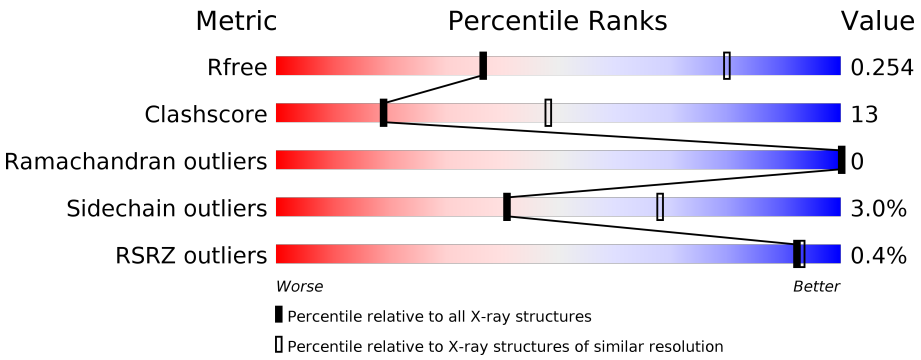
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 <=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	244	62% 22% • 14%
2	B	213	% 58% 28% • 13%
3	C	150	72% 24% • •
4	D	223	69% 23% 8%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6028 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 Platinum sensitivity protein 3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	210	1679	1097	277	297	8	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	expression tag	UNP Q12318
A	0	SER	-	expression tag	UNP Q12318

- Molecule 2 is a protein called Chromosome segregation in meiosis protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	186	1490	964	237	283	6	0	0	0

- Molecule 3 is a protein called Suppressor of HU sensitivity involved in recombination protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	146	1163	750	203	206	4	0	0	0

- Molecule 4 is a protein called Suppressor of hydroxyurea sensitivity protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	205	1695	1106	261	321	7	0	0	0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

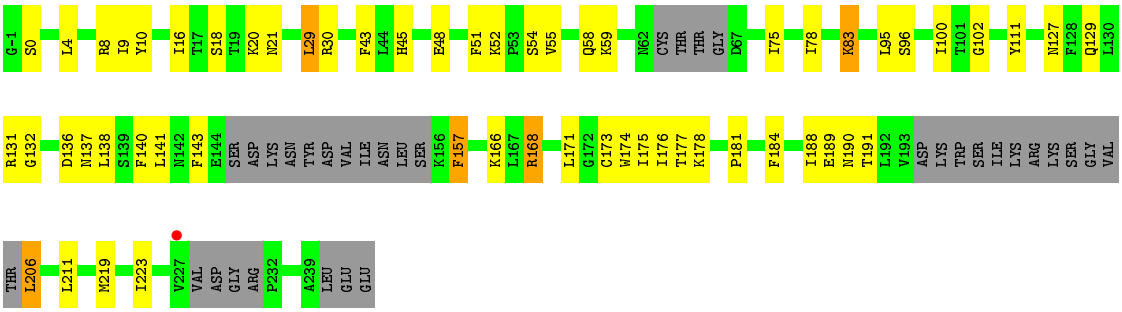
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
5	D	1	1	1	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA 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: Platinum sensitivity protein 3

Chain A:

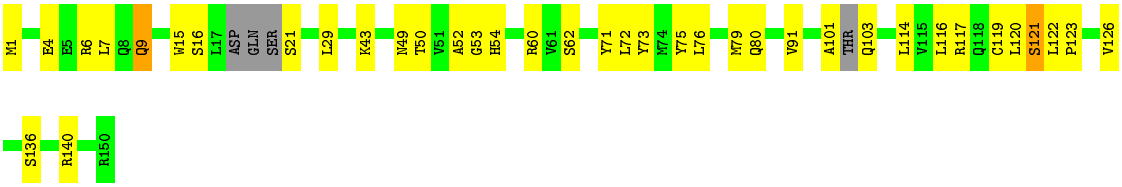


- Molecule 2: Chromosome segregation in meiosis protein 2

Chain B:

- Molecule 3: Suppressor of HU sensitivity involved in recombination protein 1

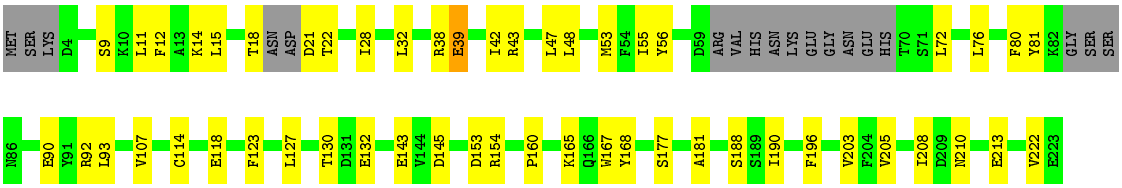
Chain C:



- Molecule 4: Suppressor of hydroxyurea sensitivity protein 2

Chain D:

69%
23%
8%


MET SER LYS D4 S9 K10 L11 F12 A13 K14 L15 T18 ASN ASP D21 T22 I28 L32 R38 E39 I42 R43 L47 L48 M53 F54 I55 Y56 D59 ARG VAL HIS ASN LYS GLU GLY ASN GLU HIS T70 S71 L72 L76 F80 Y81 K82 GLY SER SER
N86 E90 Y91 R92 I93 V107 C114 E118 F123 L127 T130 D131 E132 E143 V144 D145 D153 R154 P160 K165 Q166 W167 Y168 S177 A181 S188 S189 I190 F196 V203 F204 V205 I208 D209 N210 E213 V222 E223

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 43 21 2	Depositor
Cell constants a, b, c, α, β, γ	174.24Å 174.24Å 102.87Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	44.33 – 3.30 44.29 – 3.30	Depositor EDS
% Data completeness (in resolution range)	98.0 (44.33-3.30) 98.0 (44.29-3.30)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.09 (at 3.32Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, R_{free}	0.217 , 0.254 0.216 , 0.254	Depositor DCC
R_{free} test set	1223 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å^2)	53.0	Xtriage
Anisotropy	0.024	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 40.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6028	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.07% 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: 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	0.53	0/1709	0.69	0/2307
2	B	0.44	0/1518	0.61	0/2063
3	C	0.71	0/1188	0.75	0/1614
4	D	0.48	0/1733	0.65	0/2349
All	All	0.54	0/6148	0.67	0/8333

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	1679	0	1692	53	0
2	B	1490	0	1469	46	0
3	C	1163	0	1176	39	0
4	D	1695	0	1636	41	0
5	D	1	0	0	0	0
All	All	6028	0	5973	156	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 13.

The worst 5 of 156 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)
3:C:101:ALA:C	3:C:103:GLN:CB	2.36	0.94
3:C:117:ARG:HH11	3:C:117:ARG:HG2	1.36	0.90
2:B:33:VAL:HG12	2:B:109:PRO:HG2	1.55	0.89
3:C:117:ARG:O	3:C:121:SER:OG	1.92	0.88
2:B:3:TYR:HA	2:B:6:LEU:HD12	1.56	0.85

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	200/244 (82%)	195 (98%)	5 (2%)	0	100	100
2	B	180/213 (84%)	171 (95%)	9 (5%)	0	100	100
3	C	140/150 (93%)	135 (96%)	5 (4%)	0	100	100
4	D	197/223 (88%)	190 (96%)	7 (4%)	0	100	100
All	All	717/830 (86%)	691 (96%)	26 (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	183/228 (80%)	177 (97%)	6 (3%)	38	66
2	B	166/202 (82%)	158 (95%)	8 (5%)	25	56
3	C	127/135 (94%)	124 (98%)	3 (2%)	49	73
4	D	193/211 (92%)	190 (98%)	3 (2%)	62	79
All	All	669/776 (86%)	649 (97%)	20 (3%)	41	68

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	64	ARG
2	B	123	PHE
3	C	121	SER
2	B	37	LEU
2	B	45	SER

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

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 carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 1 ligands modelled in this entry, 1 is 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.

No monomer is involved in short contacts.

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	210/244 (86%)	-0.23	1 (0%)	91	91	8, 33, 73, 84	0
2	B	186/213 (87%)	-0.13	2 (1%)	80	81	24, 53, 78, 105	0
3	C	146/150 (97%)	-0.42	0	100	100	3, 14, 47, 68	0
4	D	205/223 (91%)	-0.06	0	100	100	18, 39, 73, 104	0
All	All	747/830 (90%)	-0.20	3 (0%)	92	93	3, 38, 74, 105	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	98	ALA	3.2
1	A	227	VAL	2.7
2	B	107	THR	2.2

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)

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
5	ZN	D	301	1/1	0.99	0.11	21,21,21,21	0

6.5 Other polymers (i)

There are no such residues in this entry.