W O R L D W I D E
wwPDB
P R O T E I N   D A T A   B A N K

wwPDB X-ray Structure Validation Summary Report (i)

Feb 13, 2017 – 09:24 am GMT

PDB ID	:	3UIN
Title	:	Complex between human RanGAP1-SUMO2, UBC9 and the IR1 domain from RanBP2
Authors	:	Gareau, J.R.; Reverter, D.; Lima, C.D.
Deposited on	:	2011-11-05
Resolution	:	2.60 Å(reported)

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

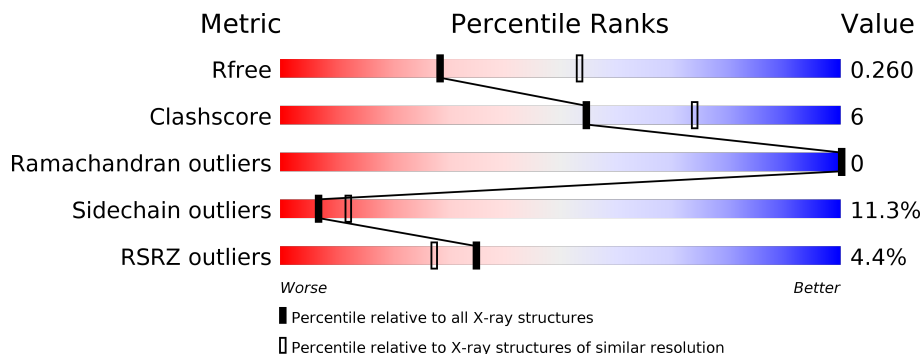|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	recalc28949

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 2.60 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	Percentile Ranks	Value
Rfree		0.260
Clashscore		6
Ramachandran outliers		0
Sidechain outliers		11.3%
RSRZ outliers		4.4%

Worse .... *Better*

▌ Percentile relative to all X-ray structures

◻ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	100719	2542 (2.60-2.60)
Clashscore	112137	2895 (2.60-2.60)
Ramachandran outliers	110173	2848 (2.60-2.60)
Sidechain outliers	110143	2848 (2.60-2.60)
RSRZ outliers	101464	2550 (2.60-2.60)

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. 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	158	3% — 82% / 15% / •••
2	B	80	4% — 61% / 33% / • •
3	C	171	3% — 74% / 14% / • / 9%
4	D	69	12% — 72% / 22% / 6%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 3651 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 SUMO-conjugating enzyme UBC9.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	157	Total	C	N	O	S	0	0	0
			1260	810	217	226	7			

- Molecule 2 is a protein called Small ubiquitin-related modifier 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	78	Total	C	N	O	S	0	0	0
			620	383	113	120	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	55	MET	VAL	SEE REMARK 999	UNP P61956

- Molecule 3 is a protein called Ran GTPase-activating protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	156	Total	C	N	O	S	0	0	0
			1204	775	199	224	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	417	SER	-	EXPRESSION TAG	UNP P46060
C	418	LEU	-	EXPRESSION TAG	UNP P46060

- Molecule 4 is a protein called E3 SUMO-protein ligase RanBP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	65	Total	C	N	O	S	0	0	0
			525	335	78	111	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	2627	SER	-	EXPRESSION TAG	UNP P49792
D	2628	LEU	-	EXPRESSION TAG	UNP P49792

- Molecule 5 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	18	Total 18	O 18	0	0
5	B	6	Total 6	O 6	0	0
5	C	14	Total 14	O 14	0	0
5	D	4	Total 4	O 4	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 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.
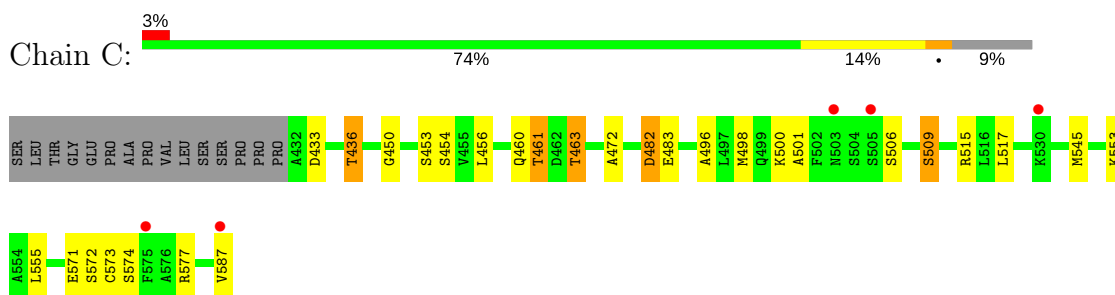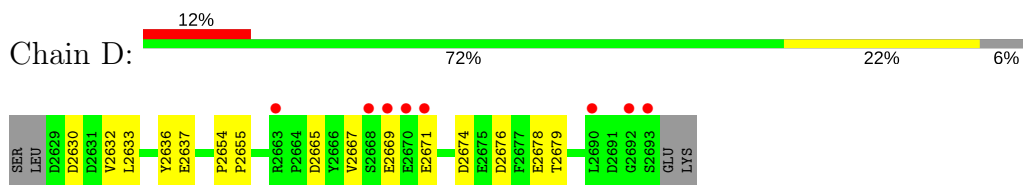
• Molecule 1: SUMO-conjugating enzyme UBC9

Chain A:
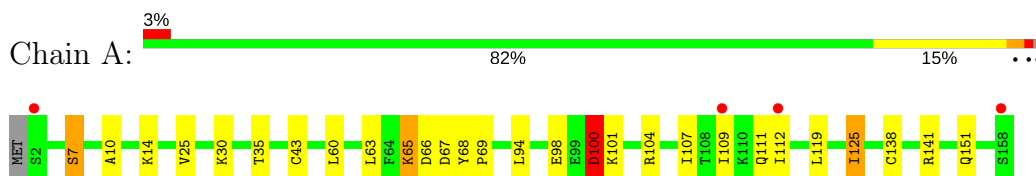
• Molecule 2: Small ubiquitin-related modifier 2

Chain B:
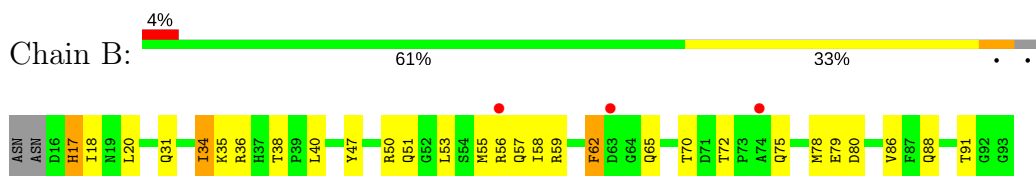
• Molecule 3: Ran GTPase-activating protein 1

Chain C:

• Molecule 4: E3 SUMO-protein ligase RanBP2

Chain D:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	155.41Å 155.41Å 57.93Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	24.41 – 2.60	Depositor
	24.41 – 2.60	EDS
% Data completeness	99.2 (24.41-2.60)	Depositor
(in resolution range)	99.2 (24.41-2.60)	EDS
R_{merge}	0.04	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.66 (at 2.60Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.2_869)	Depositor
R, R_{free}	0.211 , 0.264	Depositor
	0.208 , 0.260	DCC
R_{free} test set	1263 reflections (5.09%)	DCC
Wilson B-factor (Å^2)	59.0	Xtriage
Anisotropy	0.042	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 35.5	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.039 for -h,-k,l	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	3651	wwPDB-VP
Average B, all atoms (Å^2)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.90% 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)

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	0.54	0/1297	0.63	0/1758
2	B	0.46	0/629	0.57	0/843
3	C	0.45	0/1227	0.57	0/1661
4	D	0.49	0/535	0.58	0/725
All	All	0.49	0/3688	0.59	0/4987

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	A	0	1
3	C	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	100	ASP	Peptide
3	C	572	SER	Peptide

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	1260	0	1258	13	0
2	B	620	0	607	13	0
3	C	1204	0	1238	12	0
4	D	525	0	505	7	0
5	A	18	0	0	3	0
5	B	6	0	0	0	0
5	C	14	0	0	0	0
5	D	4	0	0	1	0
All	All	3651	0	3608	43	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 6.

The worst 5 of 43 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:A:111:GLN:NE2	5:A:215:HOH:O	2.17	0.75
1:A:138:CYS:SG	5:A:204:HOH:O	2.46	0.74
2:B:72:THR:HG23	2:B:75:GLN:H	1.59	0.68
1:A:104:ARG:H	1:A:107:ILE:HD12	1.59	0.67
2:B:50:ARG:NH1	4:D:2637:GLU:OE2	2.28	0.66

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	155/158 (98%)	152 (98%)	3 (2%)	0	100	100
2	B	76/80 (95%)	72 (95%)	4 (5%)	0	100	100
3	C	154/171 (90%)	146 (95%)	8 (5%)	0	100	100
4	D	63/69 (91%)	61 (97%)	2 (3%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	448/478 (94%)	431 (96%)	17 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	136/137 (99%)	123 (90%)	13 (10%)	10	18
2	B	68/71 (96%)	53 (78%)	15 (22%)	1	2
3	C	136/149 (91%)	124 (91%)	12 (9%)	12	22
4	D	59/63 (94%)	54 (92%)	5 (8%)	12	24
All	All	399/420 (95%)	354 (89%)	45 (11%)	7	12

5 of 45 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	58	ILE
2	B	86	VAL
4	D	2632	VAL
2	B	62	PHE
2	B	88	GLN

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

5.3.3 RNA ⓘ

There are no RNA molecules 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 ⓘ

There are no ligands in this entry.

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	157/158 (99%)	-0.03	4 (2%) 58	50	50, 64, 86, 114	0
2	B	78/80 (97%)	0.11	3 (3%) 41	33	53, 92, 120, 131	0
3	C	156/171 (91%)	0.11	5 (3%) 48	40	51, 80, 106, 121	0
4	D	65/69 (94%)	0.30	8 (12%) 5	2	54, 81, 133, 149	0
All	All	456/478 (95%)	0.09	20 (4%) 35	27	50, 74, 113, 149	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	2669	GLU	4.2
1	A	2	SER	3.8
4	D	2670	GLU	3.8
1	A	158	SER	3.5
3	C	505	SER	3.2

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 ⓘ

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.