wwPDB X-ray Structure Validation Summary Report (i)

Mar 23, 2024 – 01:29 PM EDT

PDB ID	:	1Z5S
Title	:	Crystal structure of a complex between UBC9, SUMO-1, RANGAP1 and NUP358/RANBP2
Authors	:	Reverter, D.; Lima, C.D.
Deposited on	:	2005-03-19
Resolution	:	3.01 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 (1)) were used in the production of this report:

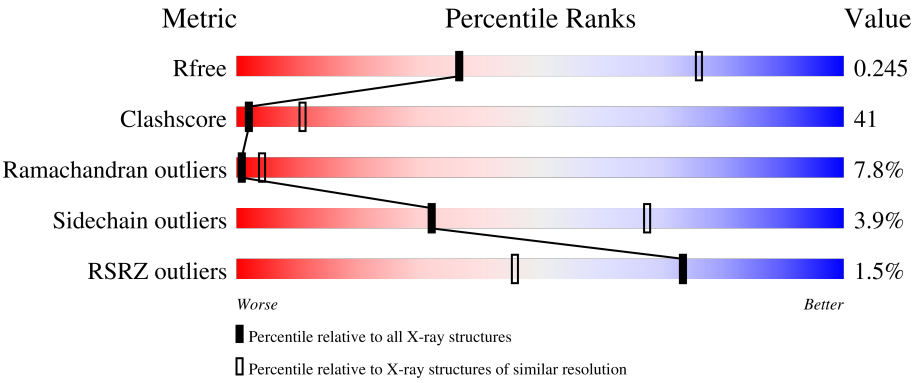
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36.1

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 3.01 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	2399 (3.04-3.00)
Clashscore	141614	2734 (3.04-3.00)
Ramachandran outliers	138981	2640 (3.04-3.00)
Sidechain outliers	138945	2643 (3.04-3.00)
RSRZ outliers	127900	2287 (3.04-3.00)

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	158	46% 49%
2	B	82	30% 59% 6% 5%
3	C	172	34% 49% 8% 9%
4	D	83	8% 17% 58% 22%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 3592 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 Ubiquitin-conjugating enzyme E2 I.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	156	Total	C	N	O	S	0	0	0
			1243	800	214	222	7			

- Molecule 2 is a protein called Ubiquitin-like protein SMT3C.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	78	Total	C	N	O	S	0	0	0
			632	398	109	121	4			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	16	MET	-	cloning artifact	UNP P63165
B	17	GLY	-	cloning artifact	UNP P63165

- 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
			1192	766	196	224	6			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	416	SER	-	cloning artifact	UNP P46060
C	417	LEU	-	cloning artifact	UNP P46060

- Molecule 4 is a protein called Ran-binding protein 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	65	Total	C	N	O	S	0	0	0
			497	322	73	101	1			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	2629	SER	-	cloning artifact	UNP P49792
D	2630	LEU	-	cloning artifact	UNP P49792

- Molecule 5 is water.

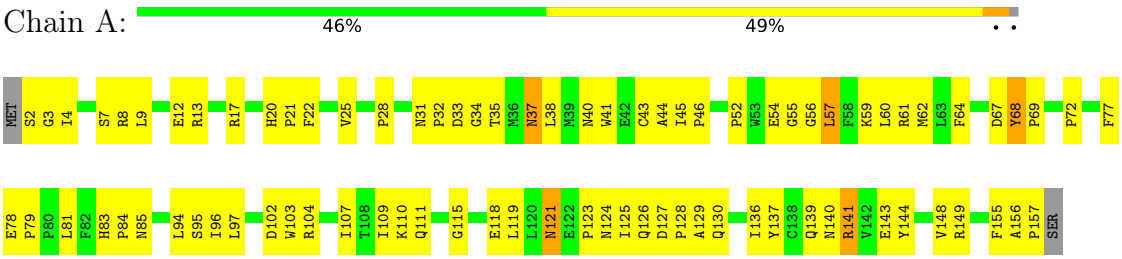
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	7	Total O 7 7	0	0
5	B	4	Total O 4 4	0	0
5	C	8	Total O 8 8	0	0
5	D	9	Total O 9 9	0	0

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.

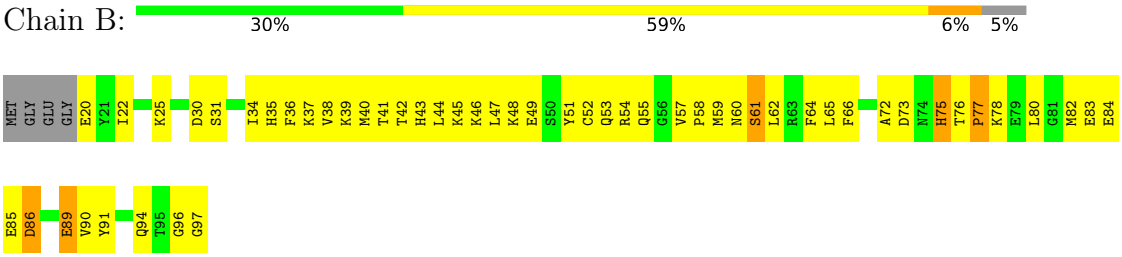
• Molecule 1: Ubiquitin-conjugating enzyme E2 I

Chain A:



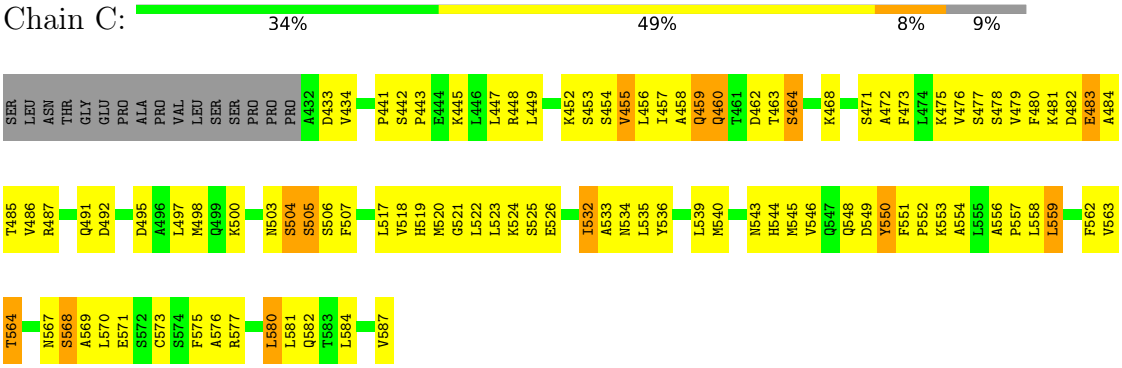
• Molecule 2: Ubiquitin-like protein SMT3C

Chain B:



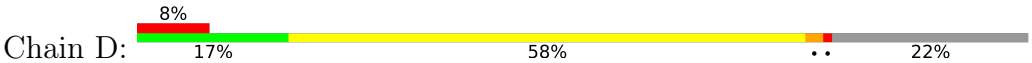
• Molecule 3: Ran GTPase-activating protein 1

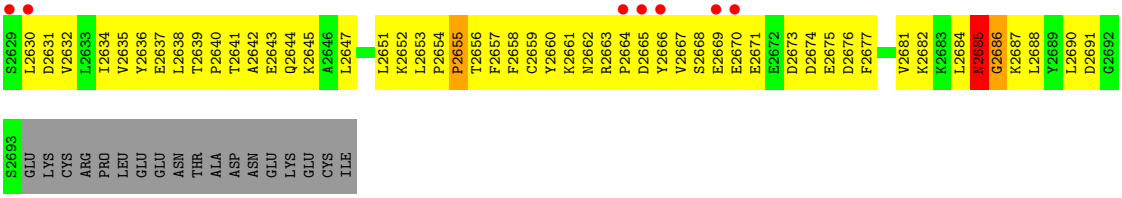
Chain C:



• Molecule 4: Ran-binding protein 2

Chain D:




S2629 L2630 D2631 V2632 L2633 I2634 V2635 Y2636 E2637 L2638 T2639 P2640 T2641 A2642 E2643 Q2644 K2645 A2646 L2647 L2651 K2652 L2653 P2654 P2655 T2656 F2657 F2658 C2659 Y2660 K2661 N2662 R2663 P2664 D2665 Y2666 V2667 S2668 E2669 E2670 E2671 E2672 D2673 D2674 E2675 D2676 F2677 V2681 K2682 K2683 L2684 N2685 G2686 K2687 L2688 Y2689 L2690 D2691 G2692
S2693 GLU LYS CYS ARG PRO LEU GLU GLU GLU ASN THR ALA ASP ASN GLU LYS GLU CYS ILE

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 32 2 1	Depositor
Cell constants a, b, c, α, β, γ	157.12Å 157.12Å 59.61Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	29.69 – 3.01 29.69 – 3.01	Depositor EDS
% Data completeness (in resolution range)	96.5 (29.69-3.01) 96.6 (29.69-3.01)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.34 (at 3.00Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.247 , 0.290 0.246 , 0.245	Depositor DCC
R_{free} test set	832 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å^2)	86.8	Xtriage
Anisotropy	0.302	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.27 , 42.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.052 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3592	wwPDB-VP
Average B, all atoms (Å^2)	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.14% 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

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	0.36	0/1280	0.60	0/1739
2	B	0.31	0/642	0.58	0/858
3	C	0.34	0/1215	0.56	0/1649
4	D	0.32	0/506	0.60	0/688
All	All	0.34	0/3643	0.59	0/4934

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

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	1243	0	1227	85	0
2	B	632	0	628	68	0
3	C	1192	0	1205	107	0
4	D	497	0	475	54	0
5	A	7	0	0	0	0
5	B	4	0	0	0	0
5	C	8	0	0	1	0
5	D	9	0	0	1	0
All	All	3592	0	3535	289	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 41.

The worst 5 of 289 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:534:ASN:HD21	3:C:536:TYR:HB3	1.25	0.99
4:D:2641:THR:HB	4:D:2644:GLN:HG3	1.48	0.95
2:B:38:VAL:HG13	4:D:2632:VAL:HG22	1.57	0.86
2:B:46:LYS:HE3	4:D:2630:LEU:HD12	1.56	0.84
3:C:546:VAL:HG12	3:C:587:VAL:HG11	1.59	0.83

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	154/158 (98%)	124 (80%)	26 (17%)	4 (3%)	5 26
2	B	76/82 (93%)	59 (78%)	11 (14%)	6 (8%)	1 4
3	C	154/172 (90%)	104 (68%)	33 (21%)	17 (11%)	0 2
4	D	63/83 (76%)	44 (70%)	11 (18%)	8 (13%)	0 1
All	All	447/495 (90%)	331 (74%)	81 (18%)	35 (8%)	1 4

5 of 35 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	35	THR
2	B	61	SER
3	C	504	SER
3	C	505	SER
4	D	2665	ASP

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	132/137 (96%)	127 (96%)	5 (4%)	33	68
2	B	70/73 (96%)	67 (96%)	3 (4%)	29	64
3	C	133/150 (89%)	128 (96%)	5 (4%)	33	68
4	D	52/76 (68%)	50 (96%)	2 (4%)	33	68
All	All	387/436 (89%)	372 (96%)	15 (4%)	32	67

5 of 15 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	89	GLU
4	D	2655	PRO
3	C	460	GLN
4	D	2685	ASN
3	C	564	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 12 such sidechains are listed below:

Mol	Chain	Res	Type
3	C	548	GLN
3	C	567	ASN
4	D	2685	ASN
3	C	578	HIS
1	A	140	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)

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

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, 95^{th} 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	156/158 (98%)	-0.42	0 100 100	47, 73, 92, 98	0
2	B	78/82 (95%)	-0.09	0 100 100	69, 111, 144, 157	0
3	C	156/172 (90%)	-0.42	0 100 100	63, 91, 115, 122	0
4	D	65/83 (78%)	0.14	7 (10%) 5 2	56, 97, 143, 149	0
All	All	455/495 (91%)	-0.28	7 (1%) 73 46	47, 86, 137, 157	0

The worst 5 of 7 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	D	2629	SER	3.9
4	D	2630	LEU	3.8
4	D	2665	ASP	2.7
4	D	2664	PRO	2.7
4	D	2666	TYR	2.3

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 monosaccharides 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.