



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 02:47 AM GMT

PDB ID : 2IO1
Title : Crystal structure of human Senp2 in complex with preSUMO-3
Authors : Reverter, D.; Lima, C.D.
Deposited on : 2006-10-09
Resolution : 2.60 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
<http://wwpdb.org/validation/2016/XrayValidationReportHelp>
with specific help available everywhere you see the ⓘ symbol.

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

MolProbity : 4.02b-467
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

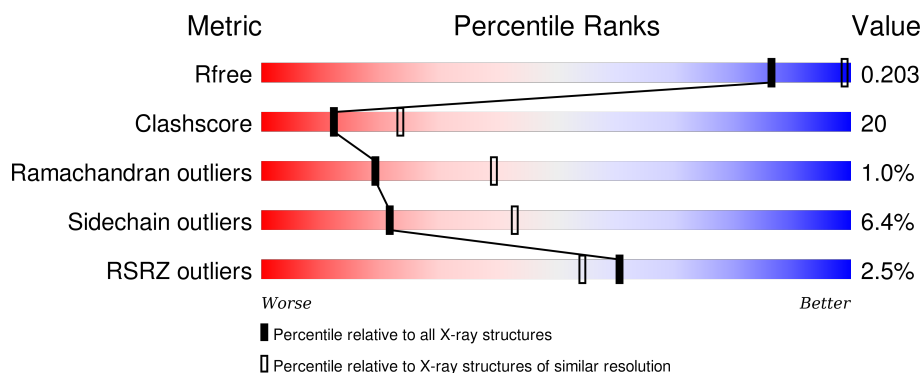
1 Overall quality at a glance ⓘ

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	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	2328 (2.60-2.60)
Clashscore	102246	2679 (2.60-2.60)
Ramachandran outliers	100387	2635 (2.60-2.60)
Sidechain outliers	100360	2635 (2.60-2.60)
RSRZ outliers	91569	2334 (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 $\leq 5\%$. 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	232	<div> <div>2%</div> <div> <div></div> <div>59%</div> <div>34%</div> <div>• •</div> </div> </div>
1	C	232	<div> <div></div> <div> <div>61%</div> <div>31%</div> <div>• •</div> </div> </div>
1	E	232	<div> <div>3%</div> <div> <div></div> <div>58%</div> <div>35%</div> <div>• •</div> </div> </div>
2	B	94	<div> <div>%</div> <div> <div></div> <div>55%</div> <div>29%</div> <div>• 14%</div> </div> </div>
2	D	94	<div> <div>3%</div> <div> <div></div> <div>56%</div> <div>29%</div> <div>• 13%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	94	 A horizontal bar chart showing the quality of chain F. The bar is divided into four segments: 5% (red), 52% (green), 30% (yellow), and 14% (grey). The segments are labeled with their respective percentages: 5%, 52%, 30%, and 14%.

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7816 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 Sentrin-specific protease 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	C	224	Total	C	N	O	S	0	0	0
			1868	1201	326	331	10			
1	E	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	358	GLY	-	CLONING ARTIFACT	UNP Q9HC62
A	359	SER	-	CLONING ARTIFACT	UNP Q9HC62
A	360	HIS	-	CLONING ARTIFACT	UNP Q9HC62
A	361	MET	-	CLONING ARTIFACT	UNP Q9HC62
A	362	ALA	-	CLONING ARTIFACT	UNP Q9HC62
A	363	SER	-	CLONING ARTIFACT	UNP Q9HC62
A	548	SER	CYS	ENGINEERED	UNP Q9HC62
C	358	GLY	-	CLONING ARTIFACT	UNP Q9HC62
C	359	SER	-	CLONING ARTIFACT	UNP Q9HC62
C	360	HIS	-	CLONING ARTIFACT	UNP Q9HC62
C	361	MET	-	CLONING ARTIFACT	UNP Q9HC62
C	362	ALA	-	CLONING ARTIFACT	UNP Q9HC62
C	363	SER	-	CLONING ARTIFACT	UNP Q9HC62
C	548	SER	CYS	ENGINEERED	UNP Q9HC62
E	358	GLY	-	CLONING ARTIFACT	UNP Q9HC62
E	359	SER	-	CLONING ARTIFACT	UNP Q9HC62
E	360	HIS	-	CLONING ARTIFACT	UNP Q9HC62
E	361	MET	-	CLONING ARTIFACT	UNP Q9HC62
E	362	ALA	-	CLONING ARTIFACT	UNP Q9HC62
E	363	SER	-	CLONING ARTIFACT	UNP Q9HC62
E	548	SER	CYS	ENGINEERED	UNP Q9HC62

- Molecule 2 is a protein called Small ubiquitin-related modifier 3 precursor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	81	Total 648	C 401	N 116	O 127	S 4	0	0	0
2	D	82	Total 654	C 404	N 117	O 129	S 4	0	0	0
2	F	81	Total 648	C 401	N 116	O 127	S 4	0	0	0

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	10	GLY	-	CLONING ARTIFACT	UNP P55854
B	11	SER	-	CLONING ARTIFACT	UNP P55854
B	12	HIS	-	CLONING ARTIFACT	UNP P55854
B	13	MET	-	CLONING ARTIFACT	UNP P55854
D	10	GLY	-	CLONING ARTIFACT	UNP P55854
D	11	SER	-	CLONING ARTIFACT	UNP P55854
D	12	HIS	-	CLONING ARTIFACT	UNP P55854
D	13	MET	-	CLONING ARTIFACT	UNP P55854
F	10	GLY	-	CLONING ARTIFACT	UNP P55854
F	11	SER	-	CLONING ARTIFACT	UNP P55854
F	12	HIS	-	CLONING ARTIFACT	UNP P55854
F	13	MET	-	CLONING ARTIFACT	UNP P55854

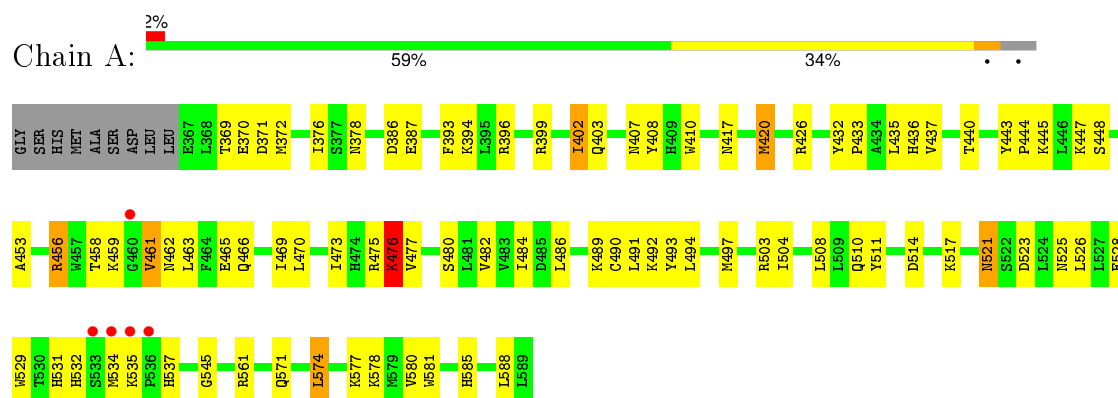
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	74	Total 74	O 74	0	0
3	B	27	Total 27	O 27	0	0
3	C	83	Total 83	O 83	0	0
3	D	31	Total 31	O 31	0	0
3	E	50	Total 50	O 50	0	0
3	F	13	Total 13	O 13	0	0

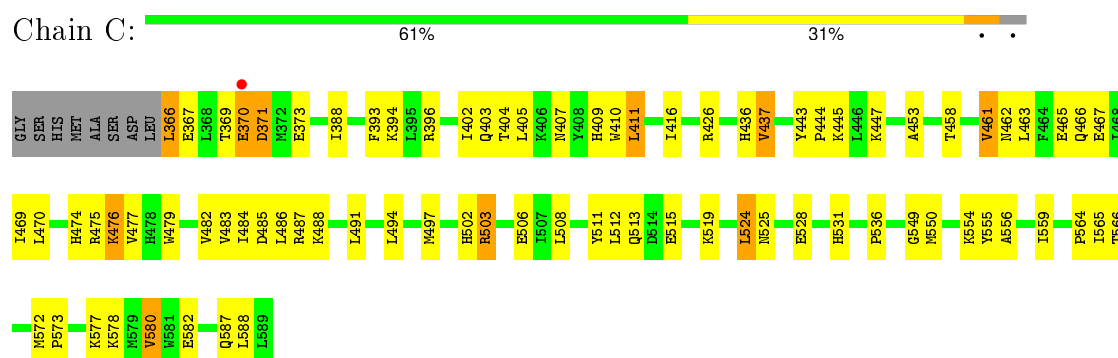
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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.

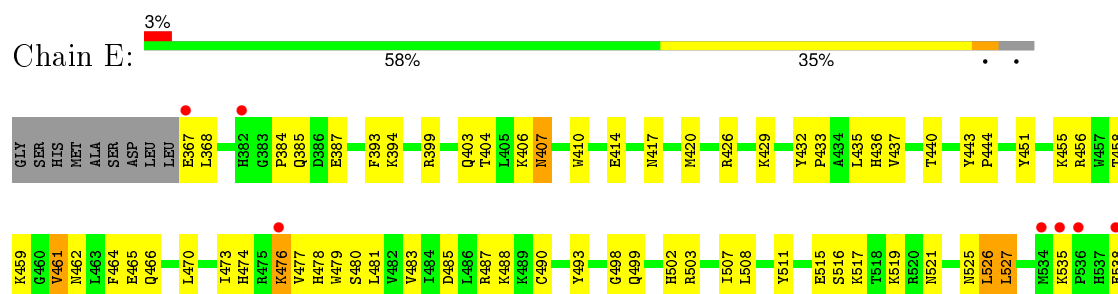
• Molecule 1: Sentrin-specific protease 2



• Molecule 1: Sentrin-specific protease 2



• Molecule 1: Sentrin-specific protease 2

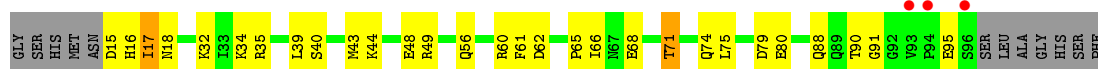




- Molecule 2: Small ubiquitin-related modifier 3 precursor



- Molecule 2: Small ubiquitin-related modifier 3 precursor



- Molecule 2: Small ubiquitin-related modifier 3 precursor



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	141.98Å 143.36Å 134.12Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.60 20.08 – 2.60	Depositor EDS
% Data completeness (in resolution range)	91.9 (20.00-2.60) 92.0 (20.08-2.60)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.25 (at 2.60Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.204 , 0.230 0.205 , 0.203	Depositor DCC
R_{free} test set	1953 reflections (5.05%)	DCC
Wilson B-factor (Å ²)	36.2	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.32 , 50.8	EDS
Estimated twinning fraction	0.029 for -k,-h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 40408 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	7816	wwPDB-VP
Average B, all atoms (Å ²)	46.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.67% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ for acentric reflections are 0.5, 0.375 respectively for untwinned datasets, and 0.333, 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	0.51	0/1906	0.68	0/2567
1	C	0.51	0/1914	0.70	0/2578
1	E	0.48	0/1906	0.65	0/2567
2	B	0.49	0/658	0.63	0/884
2	D	0.47	0/664	0.63	0/892
2	F	0.41	0/658	0.61	0/884
All	All	0.49	0/7706	0.66	0/10372

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	1860	0	1870	77	1
1	C	1868	0	1881	76	0
1	E	1860	0	1870	72	0
2	B	648	0	634	30	0
2	D	654	0	639	24	1
2	F	648	0	634	44	0
3	A	74	0	0	1	0
3	B	27	0	0	1	0
3	C	83	0	0	3	0
3	D	31	0	0	2	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	E	50	0	0	5	0
3	F	13	0	0	2	0
All	All	7816	0	7528	296	1

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 20.

The worst 5 of 296 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:367:GLU:HB2	1:E:367:GLU:N	1.60	1.12
1:E:476:LYS:HD3	1:E:477:VAL:HG13	1.30	1.12
1:A:369:THR:HG22	1:A:371:ASP:H	1.23	1.04
2:F:71:THR:HG22	2:F:74:GLN:HG3	1.39	1.03
2:D:71:THR:HG22	2:D:74:GLN:H	1.28	0.98

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:475:ARG:NH2	2:D:48:GLU:O[4_555]	2.17	0.03

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	221/232 (95%)	203 (92%)	15 (7%)	3 (1%)	14	28
1	C	222/232 (96%)	216 (97%)	5 (2%)	1 (0%)	34	60
1	E	221/232 (95%)	211 (96%)	7 (3%)	3 (1%)	14	28
2	B	79/94 (84%)	76 (96%)	2 (2%)	1 (1%)	15	30

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	80/94 (85%)	76 (95%)	4 (5%)	0	100	100
2	F	79/94 (84%)	70 (89%)	8 (10%)	1 (1%)	15	30
All	All	902/978 (92%)	852 (94%)	41 (4%)	9 (1%)	19	39

5 of 9 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	49	ARG
1	A	476	LYS
1	C	476	LYS
1	E	499	GLN
1	A	521	ASN

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	207/214 (97%)	192 (93%)	15 (7%)	18	35
1	C	208/214 (97%)	196 (94%)	12 (6%)	25	49
1	E	207/214 (97%)	196 (95%)	11 (5%)	28	53
2	B	72/82 (88%)	67 (93%)	5 (7%)	19	38
2	D	73/82 (89%)	68 (93%)	5 (7%)	20	39
2	F	72/82 (88%)	66 (92%)	6 (8%)	14	27
All	All	839/888 (94%)	785 (94%)	54 (6%)	22	43

5 of 54 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	402	ILE
1	C	524	LEU
2	F	57	ILE
1	C	411	LEU
1	C	461	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 30 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	403	GLN
1	C	510	GLN
1	E	587	GLN
1	C	409	HIS
1	C	521	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 carbohydrates 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 [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	223/232 (96%)	-0.26	5 (2%) 65 59	16, 39, 76, 110	0
1	C	224/232 (96%)	-0.42	1 (0%) 93 91	16, 36, 60, 77	0
1	E	223/232 (96%)	-0.08	8 (3%) 46 38	21, 46, 86, 107	0
2	B	81/94 (86%)	-0.14	1 (1%) 81 77	29, 46, 77, 95	0
2	D	82/94 (87%)	-0.15	3 (3%) 45 37	32, 44, 75, 92	0
2	F	81/94 (86%)	0.52	5 (6%) 24 18	41, 67, 85, 97	0
All	All	914/978 (93%)	-0.16	23 (2%) 61 54	16, 44, 79, 110	0

The worst 5 of 23 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	536	PRO	6.7
2	F	93	VAL	5.7
2	F	95	GLU	5.0
2	F	94	PRO	4.9
1	E	476	LYS	4.5

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](#)

There are no ligands in this entry.

6.5 Other polymers [i](#)

There are no such residues in this entry.