



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 01:22 PM 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 validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

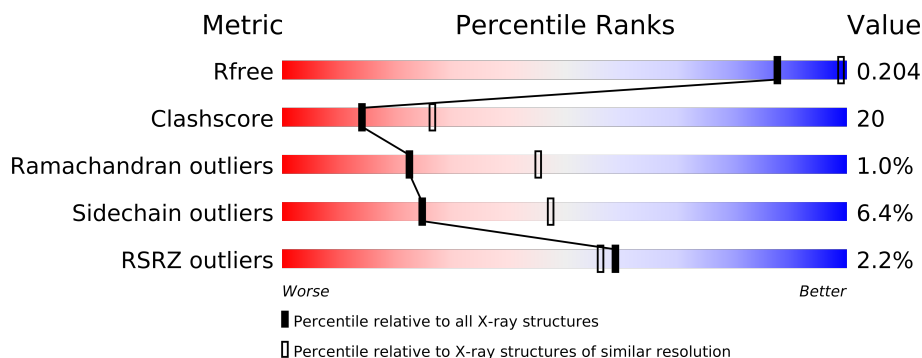
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

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))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	232	
1	C	232	
1	E	232	
2	B	94	
2	D	94	
2	F	94	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 7816 atoms, of which 0 are hydrogen and 0 are deuterium.

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



Chain B: 



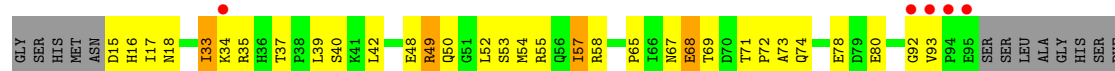
- Molecule 2: Small ubiquitin-related modifier 3 precursor

Chain D: 



- Molecule 2: Small ubiquitin-related modifier 3 precursor

Chain F: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	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
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.25 (at 2.60Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.204 , 0.230 0.205 , 0.204	Depositor DCC
$R_{free}$ test set	1953 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	36.2	Xtriage
Anisotropy	0.350	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.33 , 32.0	EDS
Estimated twinning fraction	0.029 for -k,-h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 40408 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7816	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	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.*

<sup>1</sup>Intensities estimated from amplitudes.

## 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	0.51	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 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

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

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	C	83	0	0	3	0
3	D	31	0	0	2	0
3	E	50	0	0	5	0
3	F	13	0	0	2	0
All	All	7816	0	7528	296	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 20.

The worst 5 of 296 close contacts within the same asymmetric unit are listed below.

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

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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%)	16	32
1	C	222/232 (96%)	216 (97%)	5 (2%)	1 (0%)	38	67
1	E	221/232 (95%)	211 (96%)	7 (3%)	3 (1%)	16	32
2	B	79/94 (84%)	76 (96%)	2 (2%)	1 (1%)	18	35
2	D	80/94 (85%)	76 (95%)	4 (5%)	0	100	100
2	F	79/94 (84%)	70 (89%)	8 (10%)	1 (1%)	18	35

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
All	All	902/978 (92%)	852 (94%)	41 (4%)	9 (1%)	22	45

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%)	21	39
1	C	208/214 (97%)	196 (94%)	12 (6%)	28	53
1	E	207/214 (97%)	196 (95%)	11 (5%)	32	58
2	B	72/82 (88%)	67 (93%)	5 (7%)	22	42
2	D	73/82 (89%)	68 (93%)	5 (7%)	22	43
2	F	72/82 (88%)	66 (92%)	6 (8%)	16	30
All	All	839/888 (94%)	785 (94%)	54 (6%)	25	47

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 ⓘ

There are no RNA chains 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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	223/232 (96%)	-0.28	4 (1%) 65 64	16, 39, 76, 110	0
1	C	224/232 (96%)	-0.45	0 100 100	16, 36, 60, 77	0
1	E	223/232 (96%)	-0.11	7 (3%) 47 43	21, 46, 86, 107	0
2	B	81/94 (86%)	-0.20	1 (1%) 75 77	29, 46, 77, 95	0
2	D	82/94 (87%)	-0.16	3 (3%) 39 35	32, 44, 75, 92	0
2	F	81/94 (86%)	0.47	5 (6%) 20 17	41, 67, 85, 97	0
All	All	914/978 (93%)	-0.20	20 (2%) 59 56	16, 44, 79, 110	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	536	PRO	6.6
2	F	93	VAL	5.4
2	F	94	PRO	5.1
2	F	95	GLU	4.8
1	E	476	LYS	4.5

### 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 ⓘ

There are no such residues in this entry.