



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 09:33 PM GMT

PDB ID : 5AEK  
Title : Crystal structure of the human SENP2 C548S in complex with the human SUMO1 K48M F66W  
Authors : Gallego, P.; Grana-Montes, R.; Espargaro, A.; Castillo, V.; Torrent, J.; Lange, R.; Papaleo, E.; Lindorff-Larsen, K.; Ventura, S.; Reverter, D.  
Deposited on : 2014-12-23  
Resolution : 3.00 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](mailto: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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
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) : rb-20026982

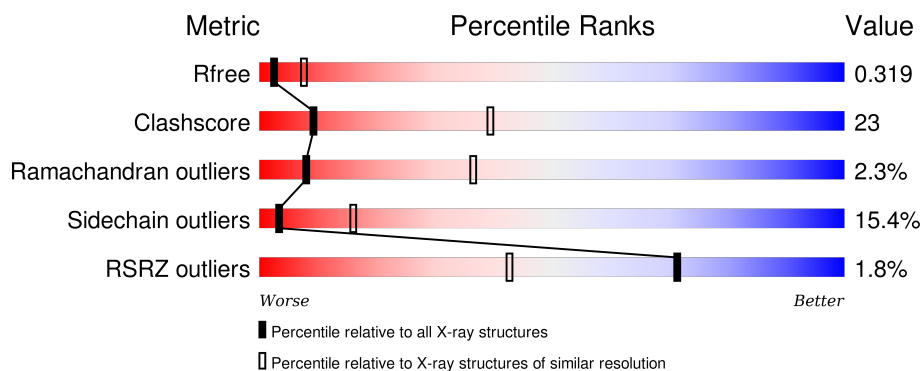
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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.






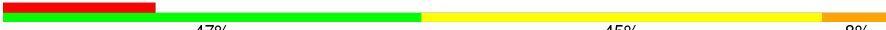






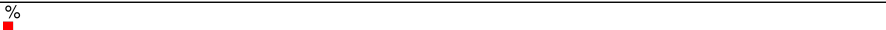

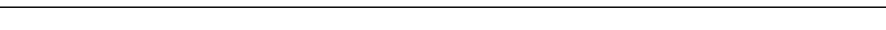
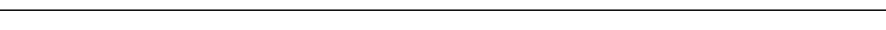





Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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 on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	224	 54% 38% 6% •
1	C	224	 55% 34% 11%
1	E	224	 40% 51% 8%
1	G	224	 57% 36% 5% •
1	I	224	 42% 44% 12% •

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	224	
1	M	224	
1	O	224	
1	Q	224	
1	S	224	
1	U	224	
1	W	224	
2	B	78	
2	D	78	
2	F	78	
2	H	78	
2	J	78	
2	L	78	
2	N	78	
2	P	78	
2	R	78	
2	T	78	
2	V	78	
2	X	78	

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 29972 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
			1865	1198	326	331	10			
1	E	224	Total	C	N	O	S	0	0	0
			1865	1198	326	331	10			
1	G	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	I	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	K	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	M	222	Total	C	N	O	S	0	0	0
			1851	1190	324	327	10			
1	O	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	Q	224	Total	C	N	O	S	0	0	0
			1865	1198	326	331	10			
1	S	224	Total	C	N	O	S	0	0	0
			1865	1198	326	331	10			
1	U	223	Total	C	N	O	S	0	0	0
			1860	1195	325	330	10			
1	W	222	Total	C	N	O	S	0	0	0
			1851	1190	324	327	10			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
C	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
E	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
G	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
I	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
K	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
M	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
O	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
Q	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
S	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
U	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62
W	548	SER	CYS	ENGINEERED MUTATION	UNP Q9HC62

- Molecule 2 is a protein called SMALL UBIQUITIN-RELATED MODIFIER 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	D	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	F	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	H	77	Total 630	C 396	N 108	O 121	S 5	0	0	0
2	J	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	L	77	Total 630	C 396	N 108	O 121	S 5	0	0	0
2	N	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	P	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	R	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	T	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	V	78	Total 639	C 401	N 109	O 124	S 5	0	0	0
2	X	78	Total 639	C 401	N 109	O 124	S 5	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
B	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
D	48	MET	LYS	ENGINEERED MUTATION	UNP P63165

*Continued on next page...*

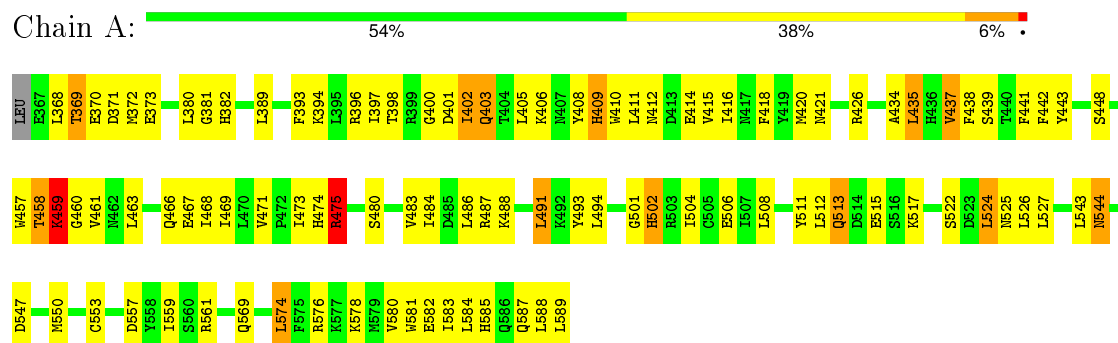
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
F	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
F	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
H	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
H	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
J	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
J	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
L	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
L	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
N	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
N	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
P	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
P	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
R	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
R	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
T	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
T	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
V	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
V	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165
X	48	MET	LYS	ENGINEERED MUTATION	UNP P63165
X	66	TRP	PHE	ENGINEERED MUTATION	UNP P63165

### 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.

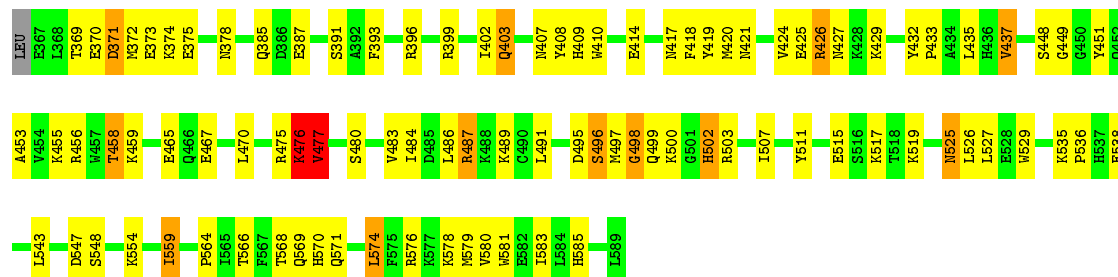
#### • Molecule 1: SENTRIN-SPECIFIC PROTEASE 2





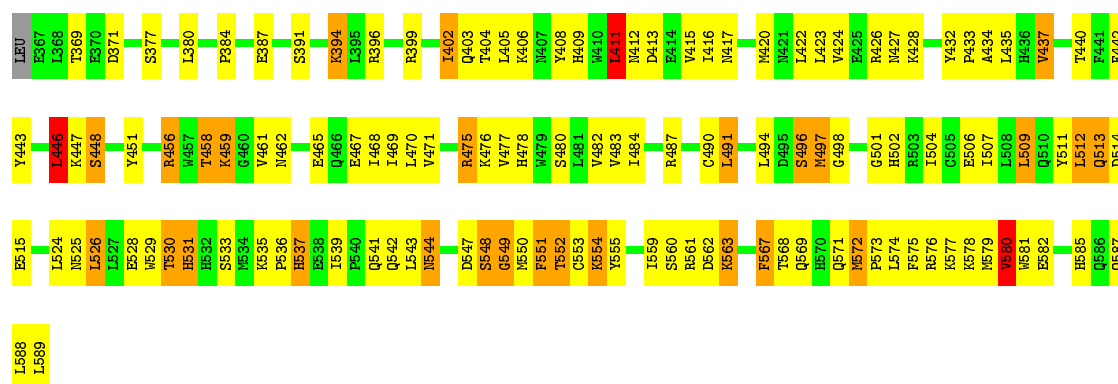
• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2

Chain G: 57% 36% 5%



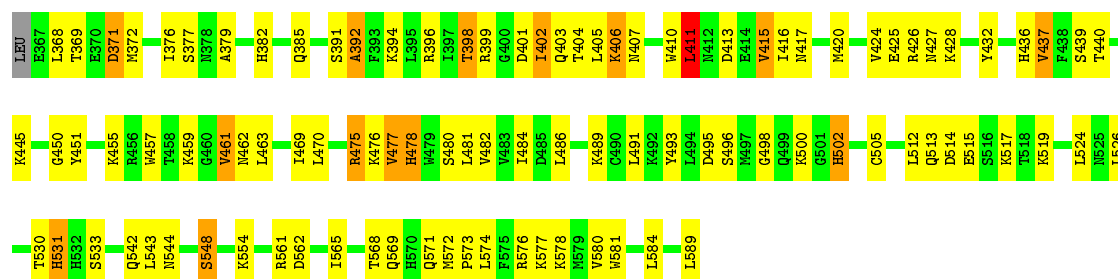
• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2

Chain I: 42% 44% 12%



• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2

Chain K: 55% 38% 6%

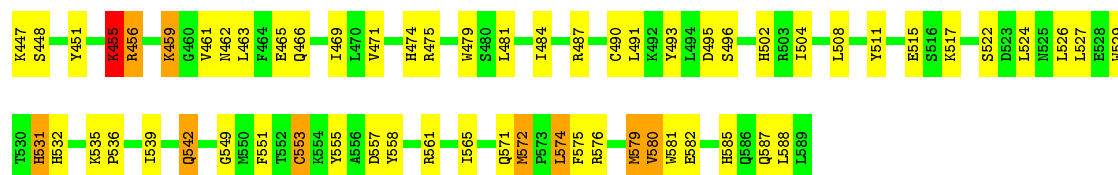


• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2

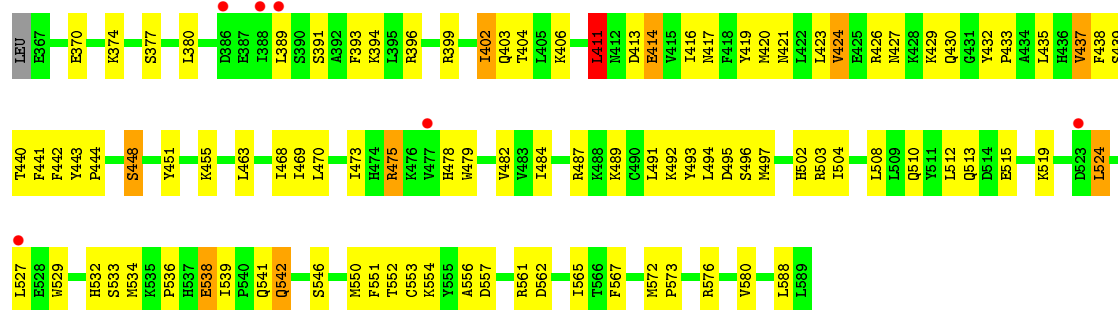
Chain M: 55% 36% 6%



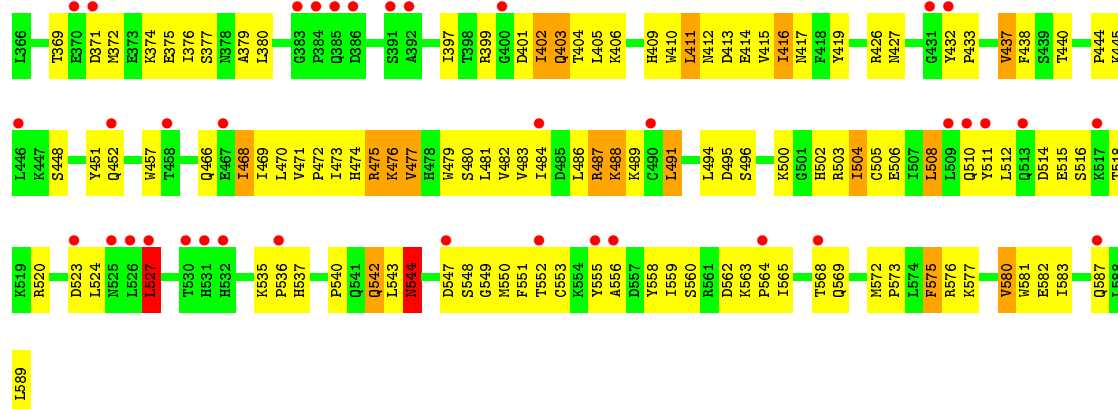




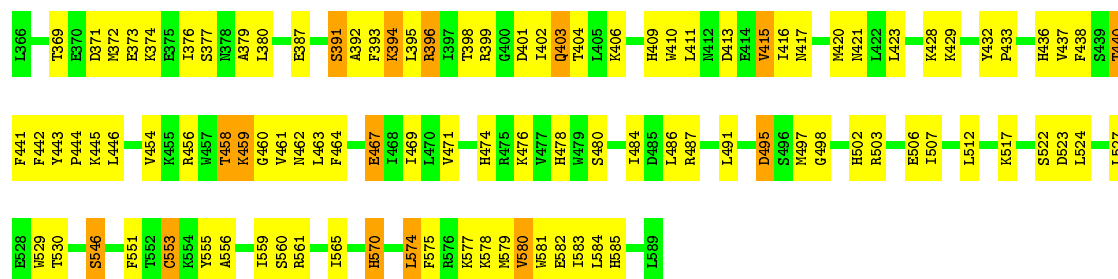
• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2



• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2



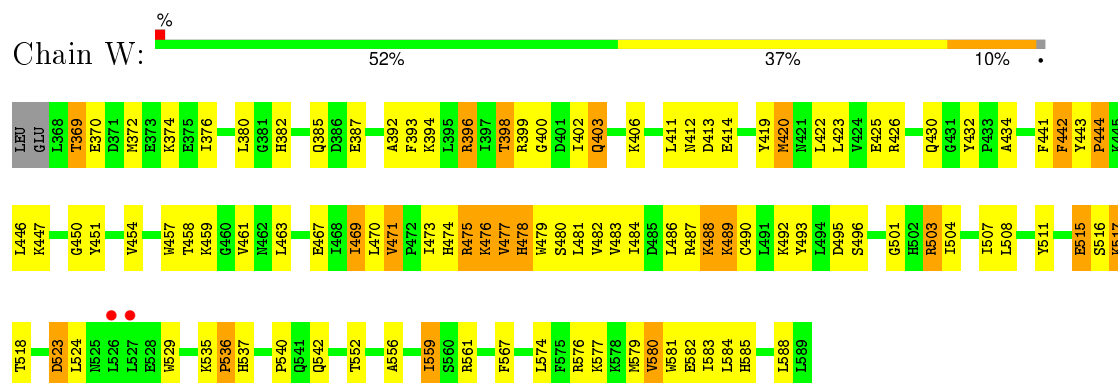
• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2



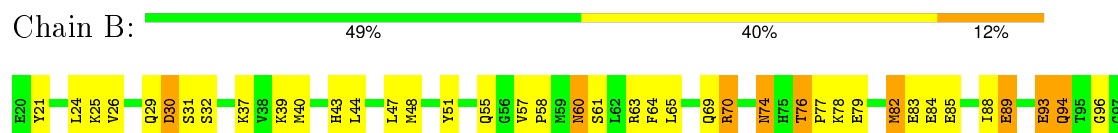
• Molecule 1: SENTRIN-SPECIFIC PROTEASE 2



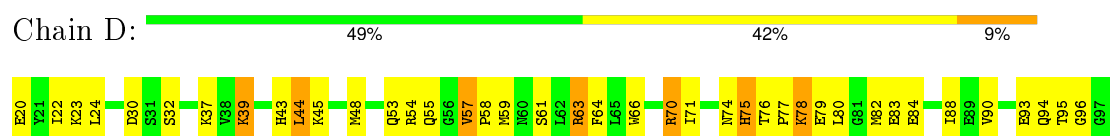
- Molecule 1: SENTRIN-SPECIFIC PROTEASE 2



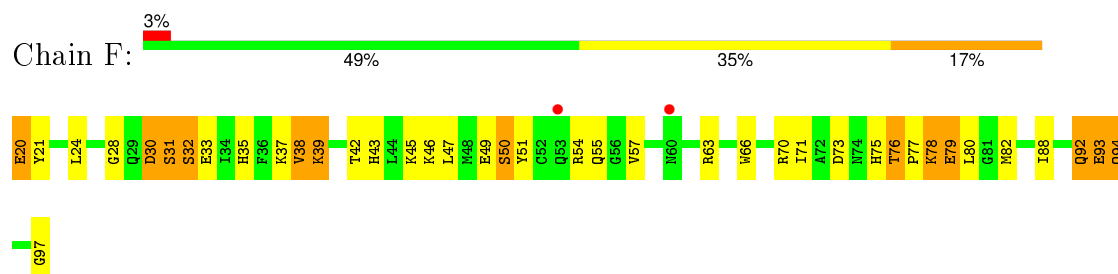
- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1

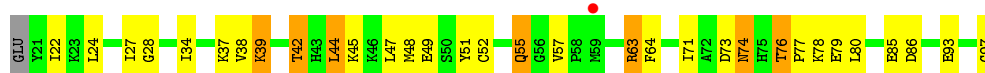


- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1

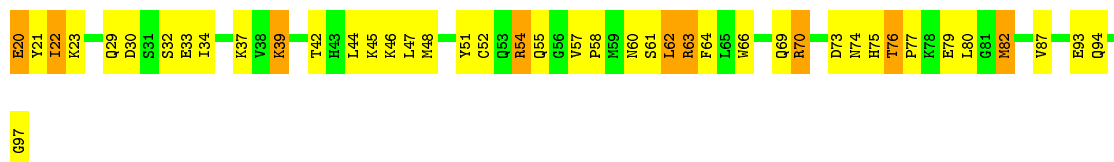


- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1





- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



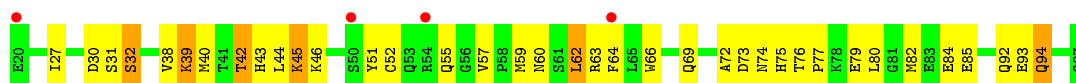
- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



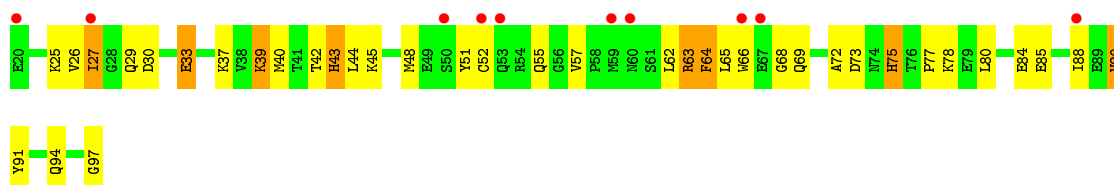
- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



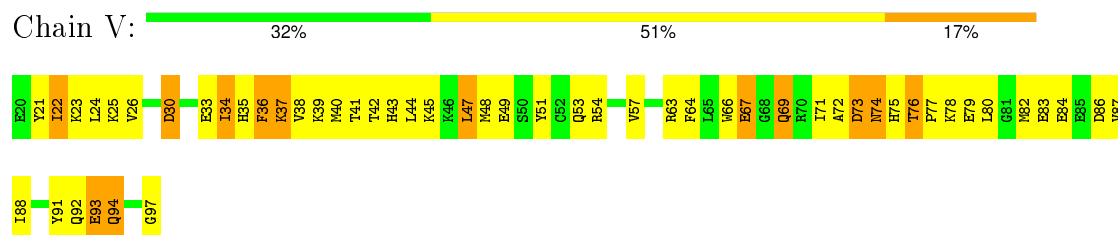
- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



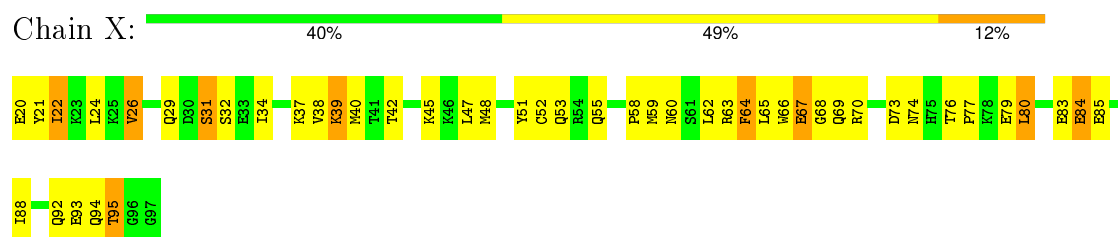
- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



- Molecule 2: SMALL UBIQUITIN-RELATED MODIFIER 1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	113.72Å 119.32Å 199.84Å 90.00° 89.67° 90.00°	Depositor
Resolution (Å)	47.00 – 3.00 46.76 – 3.00	Depositor EDS
% Data completeness (in resolution range)	97.7 (47.00-3.00) 94.0 (46.76-3.00)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 3.01Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.257 , 0.326 0.255 , 0.319	Depositor DCC
$R_{free}$ test set	3167 reflections (3.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	65.2	Xtriage
Anisotropy	0.318	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 56.7	EDS
Estimated twinning fraction	0.000 for k,h,-l 0.000 for -k,-h,-l 0.023 for h,-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 100933 reflections	Xtriage
$F_o, F_c$ correlation	0.89	EDS
Total number of atoms	29972	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.03 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.4011e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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

<sup>2</sup> Theoretical 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

### 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	0.72	1/1906 (0.1%)	0.82	1/2567 (0.0%)
1	C	0.60	1/1911 (0.1%)	0.73	1/2574 (0.0%)
1	E	0.63	0/1911	0.74	0/2574
1	G	0.70	0/1906	0.79	0/2567
1	I	0.65	0/1906	0.78	1/2567 (0.0%)
1	K	0.71	0/1906	0.81	1/2567 (0.0%)
1	M	0.73	2/1897 (0.1%)	0.81	3/2555 (0.1%)
1	O	0.57	0/1906	0.73	1/2567 (0.0%)
1	Q	0.61	0/1911	0.77	3/2574 (0.1%)
1	S	0.69	1/1911 (0.1%)	0.80	0/2574
1	U	0.70	0/1906	0.78	1/2567 (0.0%)
1	W	0.58	0/1897	0.69	0/2555
2	B	0.75	0/650	0.84	0/869
2	D	0.62	0/650	0.72	0/869
2	F	0.77	0/650	0.80	0/869
2	H	0.69	0/641	0.76	0/857
2	J	0.65	0/650	0.76	0/869
2	L	0.66	0/641	0.81	0/857
2	N	0.78	0/650	0.82	0/869
2	P	0.57	0/650	0.69	0/869
2	R	0.68	0/650	0.68	0/869
2	T	0.76	0/650	0.81	1/869 (0.1%)
2	V	0.63	0/650	0.73	0/869
2	X	0.71	0/650	0.82	0/869
All	All	0.67	5/30656 (0.0%)	0.77	13/41212 (0.0%)

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
2	P	0	1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	553	CYS	CB-SG	-6.13	1.71	1.82
1	S	553	CYS	CB-SG	-5.63	1.72	1.81
1	C	553	CYS	CB-SG	-5.21	1.73	1.81
1	M	553	CYS	CB-SG	-5.19	1.73	1.81
1	M	394	LYS	CD-CE	5.01	1.63	1.51

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	Q	452	GLN	OE1-CD-NE2	-6.55	106.83	121.90
1	U	411	LEU	CA-CB-CG	6.36	129.93	115.30
1	Q	452	GLN	CG-CD-NE2	5.91	130.87	116.70
1	M	574	LEU	CA-CB-CG	5.83	128.72	115.30
1	O	411	LEU	CA-CB-CG	5.81	128.66	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	P	30	ASP	Peptide

## 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	1860	0	1870	85	0
1	C	1865	0	1872	77	0
1	E	1865	0	1872	153	0
1	G	1860	0	1870	73	0
1	I	1860	0	1870	115	0
1	K	1860	0	1870	74	0
1	M	1851	0	1864	68	1
1	O	1860	0	1870	66	0
1	Q	1865	0	1872	101	0
1	S	1865	0	1872	79	1
1	U	1860	0	1870	83	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	W	1851	0	1864	82	0
2	B	639	0	629	31	0
2	D	639	0	629	39	0
2	F	639	0	629	49	0
2	H	630	0	623	18	0
2	J	639	0	629	33	0
2	L	630	0	623	20	0
2	N	639	0	629	26	0
2	P	639	0	629	23	0
2	R	639	0	629	33	0
2	T	639	0	629	28	0
2	V	639	0	629	52	0
2	X	639	0	629	44	0
All	All	29972	0	29972	1372	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 23.

The worst 5 of 1372 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:I:496:SER:HA	1:I:555:TYR:OH	1.45	1.16
1:Q:520:ARG:HH21	1:Q:524:LEU:HD21	1.12	1.09
1:E:420:MET:HB3	1:E:437:VAL:HG21	1.31	1.09
1:E:426:ARG:HH11	1:E:426:ARG:HG3	1.18	1.05
1:G:375:GLU:HA	1:G:378:ASN:HD22	1.24	1.03

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:M:377:SER:OG	1:S:429:LYS:O[1_544]	2.19	0.01

## 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	221/224 (99%)	199 (90%)	19 (9%)	3 (1%)	14	51
1	C	222/224 (99%)	187 (84%)	30 (14%)	5 (2%)	8	36
1	E	222/224 (99%)	185 (83%)	31 (14%)	6 (3%)	6	32
1	G	221/224 (99%)	192 (87%)	21 (10%)	8 (4%)	4	24
1	I	221/224 (99%)	172 (78%)	38 (17%)	11 (5%)	3	15
1	K	221/224 (99%)	183 (83%)	30 (14%)	8 (4%)	4	24
1	M	220/224 (98%)	189 (86%)	27 (12%)	4 (2%)	11	45
1	O	221/224 (99%)	198 (90%)	20 (9%)	3 (1%)	14	51
1	Q	222/224 (99%)	188 (85%)	29 (13%)	5 (2%)	8	36
1	S	222/224 (99%)	195 (88%)	23 (10%)	4 (2%)	11	45
1	U	221/224 (99%)	194 (88%)	26 (12%)	1 (0%)	34	76
1	W	220/224 (98%)	189 (86%)	24 (11%)	7 (3%)	5	27
2	B	76/78 (97%)	68 (90%)	7 (9%)	1 (1%)	15	53
2	D	76/78 (97%)	61 (80%)	11 (14%)	4 (5%)	2	14
2	F	76/78 (97%)	65 (86%)	10 (13%)	1 (1%)	15	53
2	H	75/78 (96%)	69 (92%)	6 (8%)	0	100	100
2	J	76/78 (97%)	70 (92%)	5 (7%)	1 (1%)	15	53
2	L	75/78 (96%)	69 (92%)	5 (7%)	1 (1%)	15	53
2	N	76/78 (97%)	69 (91%)	5 (7%)	2 (3%)	7	33
2	P	76/78 (97%)	68 (90%)	8 (10%)	0	100	100
2	R	76/78 (97%)	63 (83%)	12 (16%)	1 (1%)	15	53
2	T	76/78 (97%)	71 (93%)	3 (4%)	2 (3%)	7	33
2	V	76/78 (97%)	66 (87%)	9 (12%)	1 (1%)	15	53
2	X	76/78 (97%)	67 (88%)	6 (8%)	3 (4%)	4	21
All	All	3564/3624 (98%)	3077 (86%)	405 (11%)	82 (2%)	8	36

5 of 82 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	E	459	LYS
2	F	31	SER

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	G	502	HIS
1	G	570	HIS
1	I	502	HIS

### 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	207/208 (100%)	181 (87%)	26 (13%)	5	24
1	C	207/208 (100%)	167 (81%)	40 (19%)	2	9
1	E	207/208 (100%)	176 (85%)	31 (15%)	3	17
1	G	207/208 (100%)	188 (91%)	19 (9%)	11	40
1	I	207/208 (100%)	172 (83%)	35 (17%)	2	13
1	K	207/208 (100%)	184 (89%)	23 (11%)	8	29
1	M	206/208 (99%)	173 (84%)	33 (16%)	3	15
1	O	207/208 (100%)	181 (87%)	26 (13%)	5	24
1	Q	207/208 (100%)	181 (87%)	26 (13%)	5	24
1	S	207/208 (100%)	179 (86%)	28 (14%)	5	20
1	U	207/208 (100%)	184 (89%)	23 (11%)	8	29
1	W	206/208 (99%)	175 (85%)	31 (15%)	3	17
2	B	71/71 (100%)	56 (79%)	15 (21%)	1	6
2	D	71/71 (100%)	62 (87%)	9 (13%)	5	23
2	F	71/71 (100%)	56 (79%)	15 (21%)	1	6
2	H	70/71 (99%)	55 (79%)	15 (21%)	1	6
2	J	71/71 (100%)	50 (70%)	21 (30%)	0	2
2	L	70/71 (99%)	60 (86%)	10 (14%)	4	19
2	N	71/71 (100%)	55 (78%)	16 (22%)	1	5
2	P	71/71 (100%)	57 (80%)	14 (20%)	1	9
2	R	71/71 (100%)	57 (80%)	14 (20%)	1	9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	T	71/71 (100%)	60 (84%)	11 (16%)	3	16
2	V	71/71 (100%)	55 (78%)	16 (22%)	1	5
2	X	71/71 (100%)	56 (79%)	15 (21%)	1	6
All	All	3332/3348 (100%)	2820 (85%)	512 (15%)	3	16

5 of 512 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	K	402	ILE
1	M	542	GLN
1	W	402	ILE
1	K	475	ARG
1	M	373	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 48 such sidechains are listed below:

Mol	Chain	Res	Type
1	K	531	HIS
1	O	417	ASN
1	W	385	GLN
2	L	29	GLN
1	M	510	GLN

### 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 [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, 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	223/224 (99%)	-0.28	0 100 100	26, 40, 65, 77	0
1	C	224/224 (100%)	-0.10	1 (0%) 93 80	42, 60, 84, 88	0
1	E	224/224 (100%)	-0.10	2 (0%) 85 64	35, 58, 88, 98	0
1	G	223/224 (99%)	-0.31	0 100 100	23, 43, 72, 85	0
1	I	223/224 (99%)	-0.11	0 100 100	25, 56, 84, 89	0
1	K	223/224 (99%)	-0.26	0 100 100	22, 42, 69, 73	0
1	M	222/224 (99%)	-0.17	1 (0%) 91 76	25, 43, 67, 95	0
1	O	223/224 (99%)	0.16	6 (2%) 58 28	57, 74, 92, 101	0
1	Q	224/224 (100%)	1.02	37 (16%) 2 1	71, 94, 110, 114	0
1	S	224/224 (100%)	-0.27	0 100 100	24, 44, 67, 74	0
1	U	223/224 (99%)	-0.21	0 100 100	24, 45, 69, 77	0
1	W	222/224 (99%)	-0.01	2 (0%) 85 64	40, 63, 89, 94	0
2	B	78/78 (100%)	-0.23	0 100 100	37, 47, 60, 64	0
2	D	78/78 (100%)	0.02	0 100 100	51, 66, 76, 77	0
2	F	78/78 (100%)	-0.05	2 (2%) 59 29	35, 50, 64, 69	0
2	H	77/78 (98%)	0.05	1 (1%) 79 53	32, 66, 83, 85	0
2	J	78/78 (100%)	-0.02	0 100 100	54, 60, 72, 74	0
2	L	77/78 (98%)	-0.14	0 100 100	35, 46, 61, 65	0
2	N	78/78 (100%)	-0.05	0 100 100	39, 48, 70, 72	0
2	P	78/78 (100%)	0.47	4 (5%) 32 12	69, 87, 101, 103	0
2	R	78/78 (100%)	0.95	10 (12%) 5 2	88, 99, 113, 117	0
2	T	78/78 (100%)	-0.08	0 100 100	29, 50, 64, 65	0
2	V	78/78 (100%)	0.15	0 100 100	35, 77, 89, 89	0
2	X	78/78 (100%)	-0.08	0 100 100	46, 57, 70, 72	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
All	All	3612/3624 (99%)	-0.02	66 (1%) 71 43	22, 56, 96, 117	0

The worst 5 of 66 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	Q	530	THR	7.0
1	Q	525	ASN	5.0
1	M	369	THR	4.9
2	R	53	GLN	4.7
1	Q	526	LEU	4.0

## 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.