



wwPDB X-ray Structure Validation Summary Report ⓘ

May 14, 2020 – 10:11 am BST

PDB ID : 5X0W
Title : Molecular mechanism for the binding between Sharpin and HOIP
Authors : Liu, J.; Li, F.; Cheng, X.; Pan, L.
Deposited on : 2017-01-23
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

A user guide is available at

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
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) : 2.11

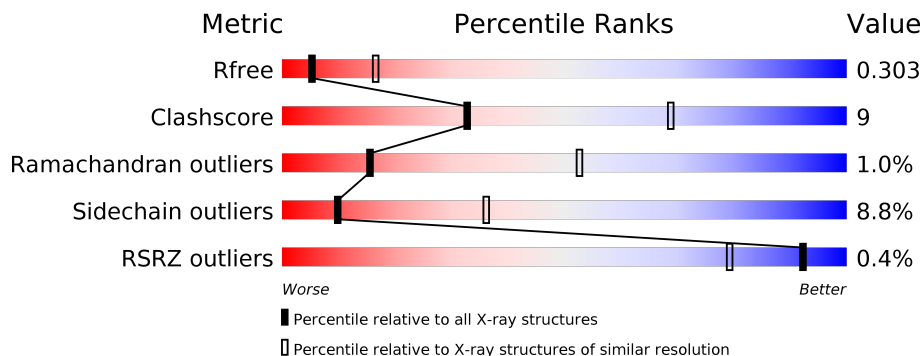
1 Overall quality at a glance

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}	130704	2092 (3.00-3.00)
Clashscore	141614	2416 (3.00-3.00)
Ramachandran outliers	138981	2333 (3.00-3.00)
Sidechain outliers	138945	2336 (3.00-3.00)
RSRZ outliers	127900	1990 (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 ≥ 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 $\leq 5\%$. 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	166	
1	C	166	
1	E	166	
1	G	166	
2	B	108	
2	D	108	

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	108	 <p>%</p> <p>59% 19% 20%</p>
2	H	108	 <p>54% 11% 35%</p>

2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 5914 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 E3 ubiquitin-protein ligase RNF31.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	129	859	545	151	157	3	3	0	0	0
1	C	130	876	547	154	169	3	3	0	0	0
1	E	143	1005	633	182	184	3	3	0	0	0
1	G	140	970	611	173	180	3	3	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	474	GLY	-	expression tag	UNP Q96EP0
A	475	PRO	-	expression tag	UNP Q96EP0
A	476	GLY	-	expression tag	UNP Q96EP0
A	477	SER	-	expression tag	UNP Q96EP0
A	478	GLU	-	expression tag	UNP Q96EP0
A	479	PHE	-	expression tag	UNP Q96EP0
C	474	GLY	-	expression tag	UNP Q96EP0
C	475	PRO	-	expression tag	UNP Q96EP0
C	476	GLY	-	expression tag	UNP Q96EP0
C	477	SER	-	expression tag	UNP Q96EP0
C	478	GLU	-	expression tag	UNP Q96EP0
C	479	PHE	-	expression tag	UNP Q96EP0
E	474	GLY	-	expression tag	UNP Q96EP0
E	475	PRO	-	expression tag	UNP Q96EP0
E	476	GLY	-	expression tag	UNP Q96EP0
E	477	SER	-	expression tag	UNP Q96EP0
E	478	GLU	-	expression tag	UNP Q96EP0
E	479	PHE	-	expression tag	UNP Q96EP0
G	474	GLY	-	expression tag	UNP Q96EP0
G	475	PRO	-	expression tag	UNP Q96EP0
G	476	GLY	-	expression tag	UNP Q96EP0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
G	477	SER	-	expression tag	UNP Q96EP0
G	478	GLU	-	expression tag	UNP Q96EP0
G	479	PHE	-	expression tag	UNP Q96EP0

- Molecule 2 is a protein called Sharpin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	82	Total	C	N	O	S	0	0	0
			584	377	101	103	3			
2	D	84	Total	C	N	O	S	0	0	0
			590	382	102	103	3			
2	F	86	Total	C	N	O	S	0	0	0
			604	384	105	113	2			
2	H	70	Total	C	N	O	S	0	0	0
			426	271	80	73	2			

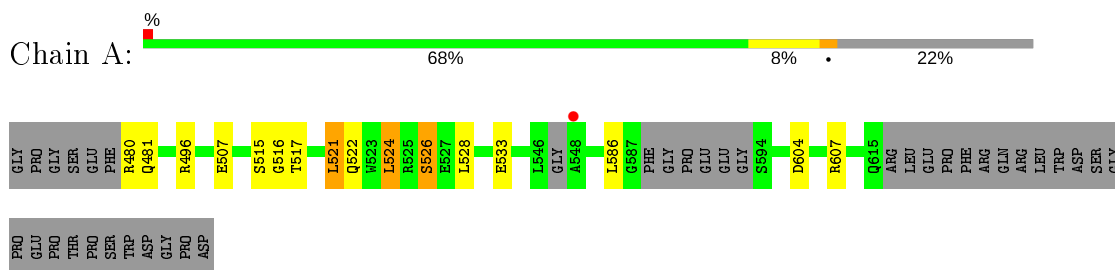
There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	202	GLY	-	expression tag	UNP Q9H0F6
B	203	PRO	-	expression tag	UNP Q9H0F6
B	204	GLY	-	expression tag	UNP Q9H0F6
B	205	SER	-	expression tag	UNP Q9H0F6
D	202	GLY	-	expression tag	UNP Q9H0F6
D	203	PRO	-	expression tag	UNP Q9H0F6
D	204	GLY	-	expression tag	UNP Q9H0F6
D	205	SER	-	expression tag	UNP Q9H0F6
F	202	GLY	-	expression tag	UNP Q9H0F6
F	203	PRO	-	expression tag	UNP Q9H0F6
F	204	GLY	-	expression tag	UNP Q9H0F6
F	205	SER	-	expression tag	UNP Q9H0F6
H	202	GLY	-	expression tag	UNP Q9H0F6
H	203	PRO	-	expression tag	UNP Q9H0F6
H	204	GLY	-	expression tag	UNP Q9H0F6
H	205	SER	-	expression tag	UNP Q9H0F6

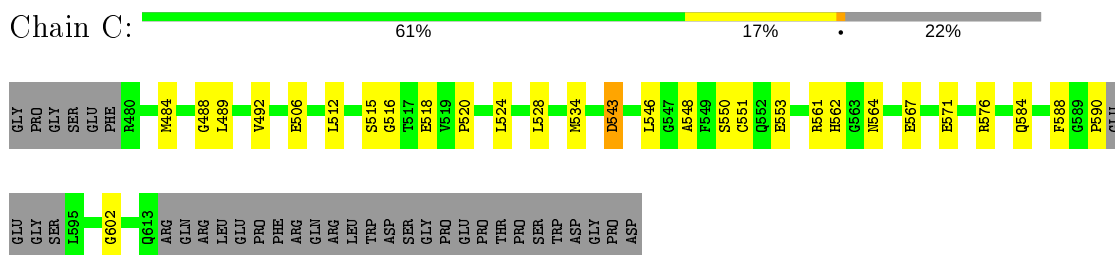
3 Residue-property plots [i](#)

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.

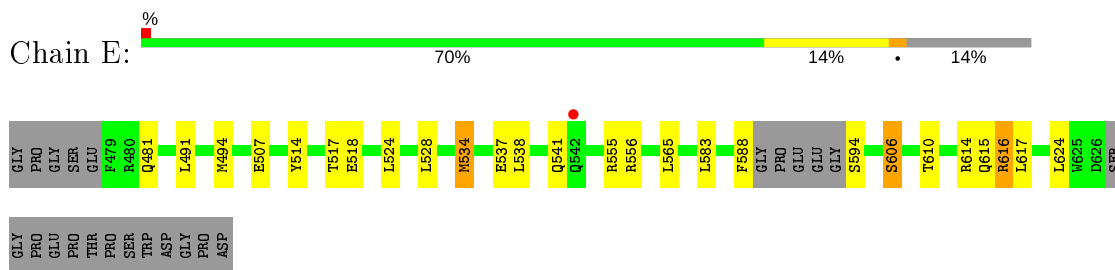
- Molecule 1: E3 ubiquitin-protein ligase RNF31



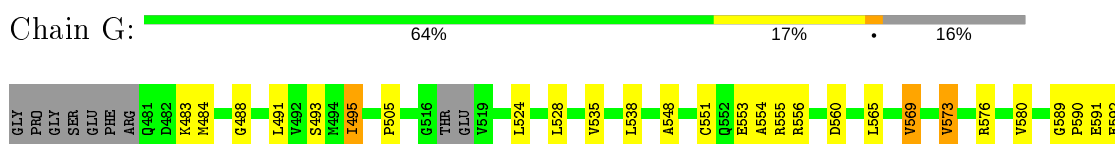
- Molecule 1: E3 ubiquitin-protein ligase RNF31

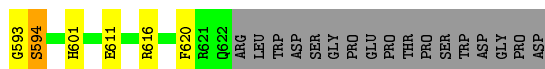


- Molecule 1: E3 ubiquitin-protein ligase RNF31

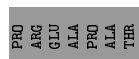


- Molecule 1: E3 ubiquitin-protein ligase RNF31

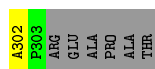
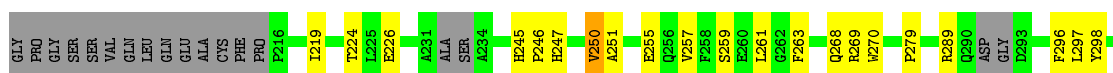




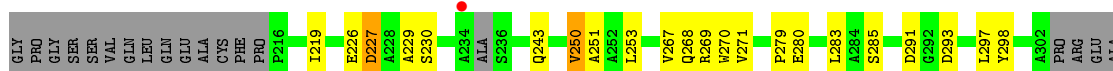
- Molecule 2: Sharpin



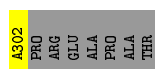
- Molecule 2: Sharpin



- Molecule 2: Sharpin



- Molecule 2: Sharpin



4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, α , β , γ	101.35Å 101.35Å 146.98Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	33.90 – 3.00 33.89 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.8 (33.90-3.00) 99.5 (33.89-3.00)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.30 (at 3.00Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.261 , 0.302 0.262 , 0.303	Depositor DCC
R_{free} test set	1707 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å ²)	83.5	Xtriage
Anisotropy	0.375	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 47.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.46$, $\langle L^2 \rangle = 0.29$	Xtriage
Estimated twinning fraction	0.047 for -h,-k,l 0.248 for h,-h-k,-l 0.049 for -k,-h,-l	Xtriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	5914	wwPDB-VP
Average B, all atoms (Å ²)	91.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.10% 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, 0.333 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	0.25	0/868	0.43	0/1159
1	C	0.26	0/886	0.45	0/1193
1	E	0.27	0/1020	0.47	0/1373
1	G	0.27	0/984	0.43	0/1320
2	B	0.30	0/598	0.54	0/816
2	D	0.27	0/605	0.49	0/826
2	F	0.26	0/616	0.50	0/841
2	H	0.27	0/434	0.51	0/583
All	All	0.27	0/6011	0.47	0/8111

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	E	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	E	616	ARG	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	859	0	713	10	0
1	C	876	0	724	19	0
1	E	1005	0	861	11	0
1	G	970	0	843	18	0
2	B	584	0	542	11	0
2	D	590	0	550	15	0
2	F	604	0	553	14	0
2	H	426	0	330	8	0
All	All	5914	0	5116	96	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 9.

The worst 5 of 96 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:506:GLU:OE2	2:D:269:ARG:NH1	1.95	0.99
1:A:496:ARG:NH1	2:B:227:ASP:O	1.95	0.98
1:G:553:GLU:OE1	1:G:576:ARG:NH1	2.09	0.86
1:A:507:GLU:HB3	1:A:524:LEU:HD11	1.74	0.69
1:C:518:GLU:HA	1:C:520:PRO:HD3	1.77	0.67

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	123/166 (74%)	118 (96%)	3 (2%)	2 (2%)	9 40
1	C	126/166 (76%)	119 (94%)	7 (6%)	0	100 100
1	E	139/166 (84%)	132 (95%)	6 (4%)	1 (1%)	22 60

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	G	136/166 (82%)	129 (95%)	6 (4%)	1 (1%)	22	60
2	B	78/108 (72%)	75 (96%)	3 (4%)	0	100	100
2	D	78/108 (72%)	71 (91%)	6 (8%)	1 (1%)	12	45
2	F	82/108 (76%)	76 (93%)	6 (7%)	0	100	100
2	H	62/108 (57%)	53 (86%)	6 (10%)	3 (5%)	2	13
All	All	824/1096 (75%)	773 (94%)	43 (5%)	8 (1%)	15	53

5 of 8 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	586	LEU
1	G	483	LYS
2	D	289	ARG
2	H	285	SER
1	A	516	GLY

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	60/134 (45%)	57 (95%)	3 (5%)	24	60
1	C	61/134 (46%)	56 (92%)	5 (8%)	11	39
1	E	80/134 (60%)	71 (89%)	9 (11%)	6	24
1	G	78/134 (58%)	67 (86%)	11 (14%)	3	16
2	B	51/84 (61%)	47 (92%)	4 (8%)	12	42
2	D	54/84 (64%)	53 (98%)	1 (2%)	57	84
2	F	53/84 (63%)	48 (91%)	5 (9%)	8	32
2	H	28/84 (33%)	25 (89%)	3 (11%)	6	26
All	All	465/872 (53%)	424 (91%)	41 (9%)	10	36

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	G	493	SER
1	G	555	ARG
2	F	285	SER
1	G	495	ILE
1	G	538	LEU

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

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	126/166 (75%)	-0.40	1 (0%) 86 65	55, 97, 140, 177	0
1	C	127/166 (76%)	-0.47	0 100 100	57, 96, 142, 164	0
1	E	140/166 (84%)	-0.38	1 (0%) 87 69	52, 81, 113, 148	0
1	G	137/166 (82%)	-0.32	0 100 100	67, 96, 125, 144	0
2	B	82/108 (75%)	-0.37	0 100 100	51, 73, 104, 161	0
2	D	84/108 (77%)	-0.33	0 100 100	59, 79, 135, 155	0
2	F	86/108 (79%)	-0.25	1 (1%) 79 54	53, 86, 119, 151	0
2	H	70/108 (64%)	-0.37	0 100 100	74, 102, 133, 143	0
All	All	852/1096 (77%)	-0.37	3 (0%) 92 79	51, 88, 133, 177	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	234	ALA	2.4
1	A	548	ALA	2.2
1	E	542	GLN	2.1

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

There are no ligands in this entry.

6.5 Other polymers

There are no such residues in this entry.