



Full wwPDB X-ray Structure Validation Report ⓘ

May 22, 2020 – 12:25 am BST

PDB ID : 5H2W
Title : Crystal structure of the karyopherin Kap60p bound to the SUMO protease Ulp1p (150-340)
Authors : Hirano, H.; Matsuura, Y.
Deposited on : 2016-10-18
Resolution : 2.50 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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
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

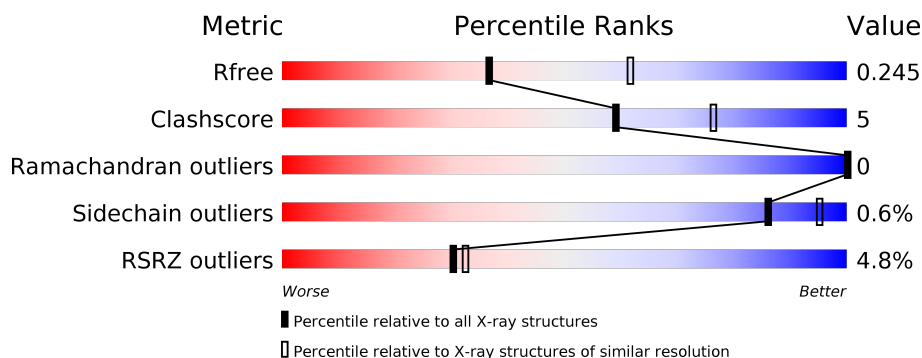
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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	4661 (2.50-2.50)
Clashscore	141614	5346 (2.50-2.50)
Ramachandran outliers	138981	5231 (2.50-2.50)
Sidechain outliers	138945	5233 (2.50-2.50)
RSRZ outliers	127900	4559 (2.50-2.50)

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	423	<div> <div>4%</div> <div> <div></div> <div>86%</div> <div>12%</div> <div></div> </div> <div></div> </div>
1	C	423	<div> <div>4%</div> <div> <div></div> <div>88%</div> <div>10%</div> <div></div> </div> <div></div> </div>
2	B	191	<div> <div>0%</div> <div> <div></div> <div>91%</div> <div></div> </div> <div>7%</div> </div>
2	D	191	<div> <div>2%</div> <div> <div></div> <div>90%</div> <div></div> </div> <div>7%</div> </div>

2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 6752 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 Importin subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	416	Total	C	N	O	S	0	0	0
			3171	2015	534	606	16			
1	C	414	Total	C	N	O	S	0	0	0
			3171	2011	529	615	16			

- Molecule 2 is a protein called Ubiquitin-like-specific protease 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	17	Total	C	N	O	0	0	0
			155	99	36	20			
2	D	19	Total	C	N	O	0	0	0
			170	107	38	25			


- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	30	Total	O	0	0
			30	30		
3	B	4	Total	O	0	0
			4	4		
3	C	48	Total	O	0	0
			48	48		
3	D	3	Total	O	0	0
			3	3		

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 ($\text{RSRZ} > 2$). 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.

- Chain A:
-
- Sequence logo for Chain A. The y-axis represents information content in bits (0.00 to 0.15). The x-axis shows positions 1 to 300. A color scale at the top indicates conservation levels: 4% (red), 86% (green), and 12% (yellow). Amino acids are labeled below the bars, and red dots above indicate specific mutations.
- Key mutations (red dots) are observed at positions: 1, 18, 32, 48, 51, 52, 59, 103, 119, 122, 132, 133, 138, 141, 146, 168, 187, 192, 209, 213, 227, 228, 231, 244, 245, 246, 247, 279, 298, 299, 300, 310, 311, 312, 313, 314, 315, 316, 317, 318, 319, 320, 321, 322, 323, 324, 325, 326, 327, 328, 329, 330, 331, 332, 333, 334, 335, 336, 337, 338, 339, 340, 341, 342, 343, 344, 345, 346, 347, 348, 349, 350, 351, 352, 353, 354, 355, 356, 357, 358, 359, 360, 361, 362, 363, 364, 365, 366, 367, 368, 369, 370, 371, 372, 373, 374, 375, 376, 377, 378, 379, 380, 381, 382, 383, 384, 385, 386, 387, 388, 389, 390, 391, 392, 393, 394, 395, 396, 397, 398, 399, 400.

- Chain C:
-
- 88% 10% 4%

- Chain B: 

ASP
ASP
LEU
VAL
PHE

● Molecule 2: Ubiquitin-like-specific protease 1



SER
SER
ASP
THR
R154
D159
T160
S161
T162
W163
A164
L165
R171
I172
GLU
SER
SER
GLU
GLY
VAL
GLY
THR
THR
PRO
SER
SER
THR
SER
PRO
PRO
ILE
SER
SER
SER
LEU
ALA
ALA
GLN
GLN
LYS
SER
ASN
CYS
ASP
SER
SER
ASP
ASN
SER
ILE
THR
PHE
SER
SER
ARG
ASP
PRO
PHE
GLY
TRP
ASN
LYS
TRP
LYS
THR
SER

ALA
ILE
GLY
SER
ASN
SER
GLU
ASN
ASN
THR
THR
SER
ASP
GLN
LYS
ASN
TYR
ASP
ARG
ARG
ARG
GLN
TYR
GLY
THR
ALA
PHE
PHE
ILE
ARG
GLY
LYS
LYS
VAL
LYS
GLN
ASN
ILE
ASN
ASN
THR
LYS
LYS
LEU
VAL
SER
SER
ARG
ALA
GLN
PHE
SER
GLU
GLU
VAL
THR
TYR
THR
LEU
ARG
GLN
ILE
PHE
ASN
THR
GLY

GLU
TYR
LYS
VAL
PRO
LYS
ILE
LEU
LYS
GLU
GLU
ARG
GLU
ARG
GLN
LEU
LYS
LEU
MET
ASP
MET
ASP
LYS
GLY
GLU
LYS
ASP
THR
GLY
LEU
LYS
LYS
SER
ILE
ILE
ASP
LEU
THR
GLU
LYS
ILE
ILE
THR
ILE
LEU
ILE
GLU
ASN
ASN
LYS
ASN
ARG
LEU
GLN
THR
ARG
ASN
GLU
ASN
ASP
ASP
GLY

ASP
LEU
VAL
PHE

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	52.61Å 63.54Å 80.91Å 106.09° 107.72° 90.50°	Depositor
Resolution (Å)	27.03 – 2.50 27.03 – 2.50	Depositor EDS
% Data completeness (in resolution range)	90.5 (27.03-2.50) 90.5 (27.03-2.50)	Depositor EDS
R_{merge}	0.05	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.14 (at 2.50Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.213 , 0.247 0.215 , 0.245	Depositor DCC
R_{free} test set	1544 reflections (5.17%)	wwPDB-VP
Wilson B-factor (Å ²)	29.8	Xtriage
Anisotropy	0.633	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.38 , 35.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.028 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.93	EDS
Total number of atoms	6752	wwPDB-VP
Average B, all atoms (Å ²)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.71% 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	0.22	0/3222	0.43	0/4392
1	C	0.22	0/3221	0.41	0/4389
2	B	0.37	0/158	0.45	0/207
2	D	0.32	0/174	0.58	0/231
All	All	0.23	0/6775	0.43	0/9219

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	D	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
2	D	160	THR	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	3171	0	3198	33	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	3171	0	3191	27	0
2	B	155	0	167	7	0
2	D	170	0	179	5	0
3	A	30	0	0	0	0
3	B	4	0	0	0	0
3	C	48	0	0	1	0
3	D	3	0	0	0	0
All	All	6752	0	6735	63	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 5.

All (63) 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:A:244:ARG:NH2	2:B:165:LEU:O	2.09	0.84
1:C:244:ARG:NH2	2:D:165:LEU:O	2.11	0.80
1:A:247:LYS:HB2	2:B:162:THR:HG21	1.65	0.78
1:A:95:GLN:HG3	1:A:103:GLU:HG2	1.64	0.77
1:A:192:GLN:HE21	2:B:171:ARG:HH11	1.32	0.76
1:C:95:GLN:NE2	1:C:103:GLU:OE1	2.25	0.69
1:C:463:LYS:HZ3	1:C:471:ASN:N	1.93	0.67
1:C:310:HIS:HD2	1:C:312:SER:H	1.42	0.66
1:A:310:HIS:HD2	1:A:312:SER:H	1.43	0.66
1:A:310:HIS:CD2	1:A:312:SER:H	2.20	0.60
1:A:460:GLU:O	1:A:463:LYS:HB2	2.02	0.58
1:C:310:HIS:CD2	1:C:312:SER:H	2.20	0.58
1:C:414:GLY:HA3	1:C:421:ILE:HG12	1.86	0.57
1:A:299:ARG:HH11	1:A:299:ARG:HG3	1.70	0.56
1:A:95:GLN:HG2	1:A:107:ALA:HB2	1.86	0.56
1:A:106:SER:OG	1:A:110:LYS:NZ	2.39	0.55
1:A:209:ASP:O	1:A:213:GLN:HG2	2.07	0.55
1:C:392:LEU:HD22	1:C:407:ILE:HD12	1.89	0.54
1:A:365:ILE:HG21	1:A:388:LEU:HD11	1.92	0.52
1:A:464:GLU:OE1	1:A:464:GLU:N	2.40	0.51
1:A:192:GLN:HE21	2:B:171:ARG:NH1	2.05	0.50
1:A:392:LEU:HD22	1:A:407:ILE:HD12	1.93	0.50
2:B:163:TRP:CE3	2:B:163:TRP:HA	2.46	0.50
1:A:354:LYS:HE2	1:A:357:ILE:HG13	1.93	0.50
1:A:481:GLY:O	1:A:485:LYS:HG2	2.12	0.49
1:C:279:TRP:CD2	1:C:318:PRO:HB3	2.46	0.49

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:282:SER:HB2	1:C:321:ARG:HG3	1.95	0.48
1:C:247:LYS:HA	1:C:248:PRO:HA	1.66	0.48
1:C:471:ASN:HB3	1:C:474:ALA:HB3	1.96	0.47
1:C:138:PHE:HB3	1:C:147:LEU:HG	1.97	0.46
1:C:493:GLU:H	1:C:493:GLU:CD	2.18	0.46
1:A:98:SER:O	1:A:134:ARG:NH2	2.48	0.46
1:C:365:ILE:HG21	1:C:388:LEU:HD11	1.98	0.46
1:C:493:GLU:OE1	1:C:493:GLU:N	2.46	0.46
1:A:187:VAL:HB	1:A:228:LYS:HD2	1.98	0.46
1:A:137:GLU:O	1:A:140:ARG:HG3	2.16	0.45
1:A:478:GLU:OE1	1:A:479:LYS:HD2	2.17	0.45
1:A:122:PRO:HB2	1:A:125:VAL:HG22	1.98	0.44
2:B:163:TRP:HE3	2:B:163:TRP:HA	1.82	0.44
1:A:246:LYS:HB2	2:B:162:THR:HG23	2.00	0.44
1:A:228:LYS:HD3	1:A:231:LEU:HB2	1.99	0.44
1:C:298:VAL:HG23	1:C:300:ILE:HG13	2.00	0.44
1:C:359:LYS:HD3	1:C:398:LYS:HG3	2.00	0.44
1:A:122:PRO:HB2	1:A:125:VAL:CG2	2.48	0.44
1:A:279:TRP:CD2	1:A:318:PRO:HB3	2.53	0.44
1:C:226:SER:OG	1:C:228:LYS:HB2	2.18	0.43
1:A:463:LYS:HE3	1:A:471:ASN:N	2.33	0.43
1:C:192:GLN:HE21	2:D:171:ARG:NH1	2.16	0.43
2:D:162:THR:O	2:D:163:TRP:CD1	2.72	0.43
1:C:358:LYS:NZ	3:C:611:HOH:O	2.52	0.42
1:A:496:LYS:O	1:A:500:LYS:HB2	2.20	0.42
1:A:417:ARG:HA	1:A:418:PRO:HD2	1.88	0.42
1:A:138:PHE:HB3	1:A:147:LEU:HG	2.01	0.42
1:A:450:ASP:OD1	1:A:500:LYS:HE2	2.20	0.42
1:A:298:VAL:HG23	1:A:300:ILE:HG13	2.01	0.41
1:C:434:CYS:SG	1:C:482:GLY:HA3	2.61	0.41
1:C:397:TYR:CE2	1:C:401:LYS:HD2	2.56	0.41
1:C:431:LYS:HB3	1:C:432:PRO:HD3	2.03	0.41
1:C:452:LEU:HD13	1:C:486:ILE:HD11	2.03	0.41
1:C:192:GLN:NE2	2:D:171:ARG:HD3	2.37	0.40
1:A:127:ILE:HD13	1:A:168:VAL:HG21	2.02	0.40
1:C:354:LYS:HA	1:C:354:LYS:HD2	1.82	0.40
1:C:247:LYS:CB	2:D:162:THR:HG21	2.51	0.40

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	412/423 (97%)	403 (98%)	9 (2%)	0	100	100
1	C	410/423 (97%)	404 (98%)	6 (2%)	0	100	100
2	B	13/191 (7%)	13 (100%)	0	0	100	100
2	D	17/191 (9%)	15 (88%)	2 (12%)	0	100	100
All	All	852/1228 (69%)	835 (98%)	17 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	344/364 (94%)	343 (100%)	1 (0%)	92	97
1	C	347/364 (95%)	346 (100%)	1 (0%)	92	97
2	B	16/178 (9%)	15 (94%)	1 (6%)	18	34
2	D	18/178 (10%)	17 (94%)	1 (6%)	21	40
All	All	725/1084 (67%)	721 (99%)	4 (1%)	86	95

All (4) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	491	GLN
2	B	163	TRP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	290	GLU
2	D	154	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	192	GLN
1	A	310	HIS
1	A	494	ASN
1	C	95	GLN
1	C	142	ASN
1	C	192	GLN
1	C	213	GLN
1	C	310	HIS

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 ⓘ

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, 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	416/423 (98%)	0.36	19 (4%) 32 34	27, 39, 60, 74	0
1	C	414/423 (97%)	0.31	17 (4%) 37 40	25, 36, 60, 80	0
2	B	17/191 (8%)	0.38	2 (11%) 4 4	31, 39, 77, 84	0
2	D	19/191 (9%)	0.84	4 (21%) 1 0	32, 46, 83, 84	0
All	All	866/1228 (70%)	0.34	42 (4%) 30 32	25, 37, 62, 84	0

All (42) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	118	GLU	3.9
2	D	160	THR	3.8
1	A	365	ILE	3.5
1	A	464	GLU	3.3
1	A	332	LEU	3.2
1	A	227	ASN	3.2
1	A	348	LEU	3.2
1	C	493	GLU	3.2
1	A	502	TYR	3.1
2	B	162	THR	3.0
1	C	142	ASN	3.0
1	C	90	PRO	3.0
2	D	162	THR	2.9
1	C	415	LEU	2.8
1	A	353	PRO	2.8
1	C	246	LYS	2.8
1	A	415	LEU	2.7
1	A	441	ASP	2.7
1	A	119	HIS	2.6
1	C	365	ILE	2.6
2	B	163	TRP	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	495	ASP	2.6
1	C	494	ASN	2.6
1	C	498	TYR	2.6
1	C	227	ASN	2.5
2	D	159	ASP	2.5
1	A	383	ASN	2.4
1	A	498	TYR	2.4
1	A	89	LEU	2.4
1	A	140	ARG	2.4
1	C	416	GLN	2.4
2	D	163	TRP	2.4
1	A	493	GLU	2.3
1	C	89	LEU	2.3
1	C	141	GLU	2.3
1	C	463	LYS	2.2
1	A	460	GLU	2.2
1	A	351	SER	2.2
1	A	381	ASP	2.1
1	C	403	ALA	2.1
1	A	388	LEU	2.1
1	C	193	ALA	2.0

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.