



Full wwPDB X-ray Structure Validation Report ⓘ

Aug 29, 2020 – 06:03 PM BST

PDB ID : 6QU1
Title : Crystal structure of the KAP1 RBCC domain in complex with the SMAR-CAD1 CUE1 domain at 3.7 angstrom resolution.
Authors : Newman, J.A.; Aitkenhead, H.; Gavard, A.; Lim, M.; Williams, H.L.; Svejstrup, J.Q.; von Delft, F.; Arrowsmith, C.H.; Edwards, A.; Bountra, C.; Gileadi, O.
Deposited on : 2019-02-26
Resolution : 3.70 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 : 2.13
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) : 2.13

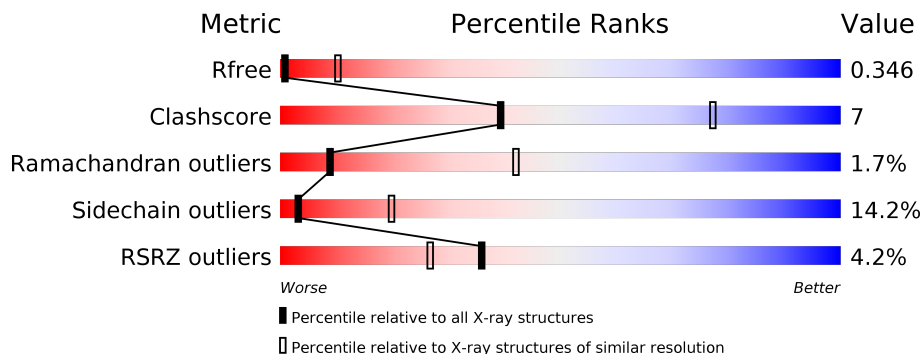
1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)

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	322	
2	D	48	

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 2444 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 Transcription intermediary factor 1-beta, Transcription intermediary factor 1-beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	262	2076	1298	383	376	19	0	0	0

- Molecule 2 is a protein called SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	D	48	364	230	56	76	2	0	0	0

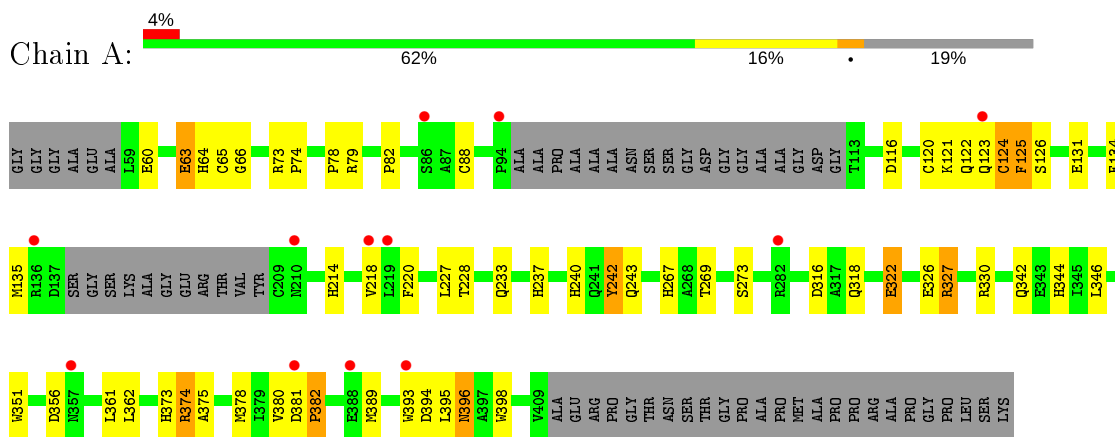
- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Zn		
3	A	4	4	4	0	0

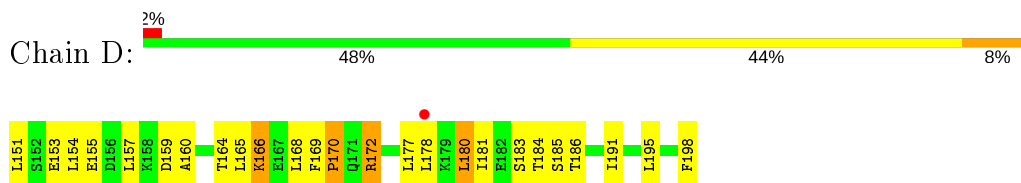
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Transcription intermediary factor 1-beta, Transcription intermediary factor 1-beta



- Molecule 2: SWI/SNF-related matrix-associated actin-dependent regulator of chromatin subfamily A containing DEAD/H box 1



4 Data and refinement statistics

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α , β , γ	64.50Å 64.50Å 287.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	29.43 – 3.70 29.43 – 3.70	Depositor EDS
% Data completeness (in resolution range)	98.7 (29.43-3.70) 98.9 (29.43-3.70)	Depositor EDS
R_{merge}	0.23	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.11 (at 3.75Å)	Xtrriage
Refinement program	PHENIX (1.13_2998: ???)	Depositor
R, R_{free}	0.300 , 0.346 0.299 , 0.346	Depositor DCC
R_{free} test set	557 reflections (7.92%)	wwPDB-VP
Wilson B-factor (Å ²)	107.5	Xtrriage
Anisotropy	0.598	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 64.7	EDS
L-test for twinning ²	$\langle L \rangle = 0.44$, $\langle L^2 \rangle = 0.27$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	2444	wwPDB-VP
Average B, all atoms (Å ²)	119.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.41% 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

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section:
ZN

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	0/2111	0.46	0/2850
2	D	0.24	0/366	0.42	0/494
All	All	0.25	0/2477	0.46	0/3344

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

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	2076	0	2042	23	0
2	D	364	0	363	14	0
3	A	4	0	0	0	0
All	All	2444	0	2405	36	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 7.

All (36) 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:342:GLN:HE21	1:A:380:VAL:HG11	1.48	0.76
1:A:78:PRO:HA	1:A:88:CYS:HA	1.74	0.69
1:A:74:PRO:HB3	1:A:135:MET:HG2	1.77	0.66
1:A:82:PRO:HG2	1:A:124:CYS:HB2	1.80	0.63
1:A:65:CYS:SG	1:A:66:GLY:N	2.75	0.59
1:A:79:ARG:HA	1:A:131:GLU:HA	1.85	0.58
1:A:228:THR:OG1	1:A:233:GLN:NE2	2.37	0.57
1:A:375:ALA:HA	1:A:378:MET:HB3	1.88	0.56
1:A:125:PHE:HD1	1:A:126:SER:H	1.55	0.55
1:A:380:VAL:C	1:A:382:PRO:HD3	2.27	0.54
1:A:374:ARG:O	1:A:378:MET:N	2.40	0.54
1:A:356:ASP:OD1	1:A:356:ASP:N	2.44	0.51
1:A:396:ASN:OD1	1:A:396:ASN:N	2.45	0.50
2:D:151:LEU:HA	2:D:154:LEU:HD23	1.93	0.50
1:A:380:VAL:O	1:A:382:PRO:HD3	2.13	0.49
2:D:172:ARG:H	2:D:172:ARG:HG2	1.37	0.48
2:D:165:LEU:HD21	2:D:191:ILE:HG22	1.97	0.47
2:D:177:LEU:HA	2:D:180:LEU:HD23	1.97	0.47
1:A:220:PHE:HD1	1:A:227:LEU:HD12	1.81	0.46
1:A:63:GLU:H	1:A:63:GLU:HG2	1.51	0.44
1:A:269:THR:O	1:A:273:SER:N	2.48	0.43
2:D:184:THR:O	2:D:185:SER:OG	2.29	0.43
1:A:116:ASP:HA	1:A:122:GLN:HA	2.00	0.43
2:D:166:LYS:O	2:D:170:PRO:HB3	2.18	0.43
1:A:318:GLN:O	1:A:322:GLU:HG2	2.19	0.43
2:D:178:LEU:HD23	2:D:181:ILE:HD11	2.01	0.43
2:D:180:LEU:HD11	2:D:191:ILE:HG23	2.00	0.43
2:D:153:GLU:O	2:D:157:LEU:HG	2.20	0.42
2:D:160:ALA:O	2:D:164:THR:HG23	2.20	0.42
1:A:237:HIS:HB3	1:A:240:HIS:HB2	2.01	0.41
1:A:233:GLN:HG3	1:A:242:TYR:CE1	2.55	0.41
2:D:165:LEU:HD22	2:D:195:LEU:HD13	2.03	0.41
2:D:186:THR:HA	2:D:191:ILE:HD11	2.02	0.41
1:A:121:LYS:CB	2:D:170:PRO:HB2	2.50	0.41
2:D:168:LEU:O	2:D:170:PRO:HD3	2.20	0.41
1:A:327:ARG:HB2	1:A:327:ARG:NH1	2.36	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	256/322 (80%)	238 (93%)	14 (6%)	4 (2%)	9	43
2	D	46/48 (96%)	44 (96%)	1 (2%)	1 (2%)	6	37
All	All	302/370 (82%)	282 (93%)	15 (5%)	5 (2%)	9	42

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	381	ASP
1	A	382	PRO
1	A	120	CYS
2	D	170	PRO
1	A	218	VAL

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	227/273 (83%)	197 (87%)	30 (13%)	4	22
2	D	40/42 (95%)	32 (80%)	8 (20%)	1	8
All	All	267/315 (85%)	229 (86%)	38 (14%)	3	20

All (38) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	60	GLU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	63	GLU
1	A	64	HIS
1	A	73	ARG
1	A	123	GLN
1	A	124	CYS
1	A	125	PHE
1	A	134	PHE
1	A	214	HIS
1	A	242	TYR
1	A	243	GLN
1	A	267	HIS
1	A	316	ASP
1	A	322	GLU
1	A	326	GLU
1	A	327	ARG
1	A	330	ARG
1	A	344	HIS
1	A	346	LEU
1	A	351	TRP
1	A	361	LEU
1	A	362	LEU
1	A	373	HIS
1	A	374	ARG
1	A	389	MET
1	A	393	TRP
1	A	394	ASP
1	A	395	LEU
1	A	396	ASN
1	A	398	TRP
2	D	155	GLU
2	D	159	ASP
2	D	166	LYS
2	D	169	PHE
2	D	172	ARG
2	D	180	LEU
2	D	183	SER
2	D	198	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (4) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	233	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	271	GLN
1	A	342	GLN
2	D	171	GLN

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	262/322 (81%)	0.03	12 (4%) 32 24	80, 119, 150, 189	0
2	D	48/48 (100%)	0.02	1 (2%) 63 52	90, 110, 143, 155	0
All	All	310/370 (83%)	0.03	13 (4%) 36 27	80, 117, 150, 189	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	86	SER	4.3
1	A	357	ASN	3.6
1	A	393	TRP	3.5
1	A	94	PRO	3.5
1	A	123	GLN	2.9
1	A	210	ASN	2.8
2	D	178	LEU	2.7
1	A	388	GLU	2.4
1	A	219	LEU	2.2
1	A	381	ASP	2.1
1	A	282	ARG	2.1
1	A	218	VAL	2.0
1	A	136	ARG	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 monosaccharides in this entry.

6.4 Ligands [i](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q < 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
3	ZN	A	1001	1/1	0.93	0.07	125,125,125,125	0
3	ZN	A	1003	1/1	0.96	0.10	126,126,126,126	0
3	ZN	A	1002	1/1	0.97	0.10	90,90,90,90	0
3	ZN	A	1004	1/1	0.97	0.10	133,133,133,133	0

6.5 Other polymers [i](#)

There are no such residues in this entry.