



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 1, 2021 – 12:07 PM JST

PDB ID : 7E34
Title : Crystal structure of SUN1-Speedy A-CDK2
Authors : Chen, Y.; Huang, C.; Wu, J.; Lei, M.
Deposited on : 2021-02-08
Resolution : 3.19 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.2
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) : 2.23.2

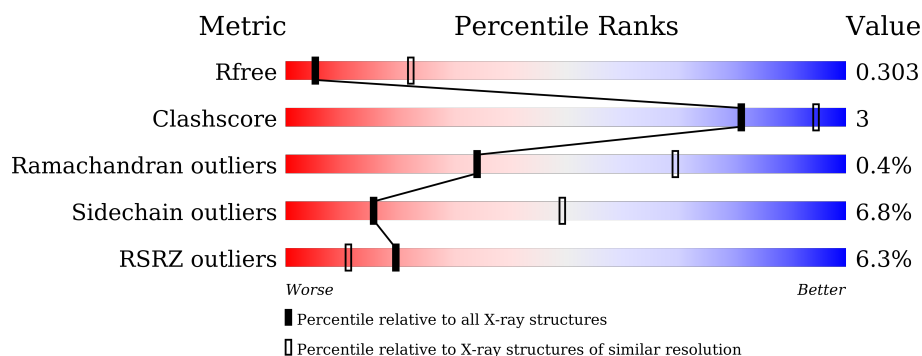
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.19 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	1133 (3.20-3.20)
Clashscore	141614	1253 (3.20-3.20)
Ramachandran outliers	138981	1234 (3.20-3.20)
Sidechain outliers	138945	1233 (3.20-3.20)
RSRZ outliers	127900	1095 (3.20-3.20)

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 of 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	298	<div> <div>4%</div> <div>87%</div> <div>13%</div> </div>
2	B	153	<div> <div>6%</div> <div>78%</div> <div>14%</div> <div>• 7%</div> </div>
3	C	46	<div> <div>15%</div> <div>35%</div> <div>• •</div> <div>59%</div> </div>

2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 3815 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 Cyclin-dependent kinase 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	298	Total	C	N	O	S	0	0	0
			2395	1557	408	423	7			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	SER	-	expression tag	UNP P24941

- Molecule 2 is a protein called Speedy protein A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	142	Total	C	N	O	S	0	0	0
			1216	793	204	209	10			

- Molecule 3 is a protein called SUN domain-containing protein 1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	19	Total	C	N	O	0	0	0
			166	104	27	35			

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total	Zn	0	0
			1	1		

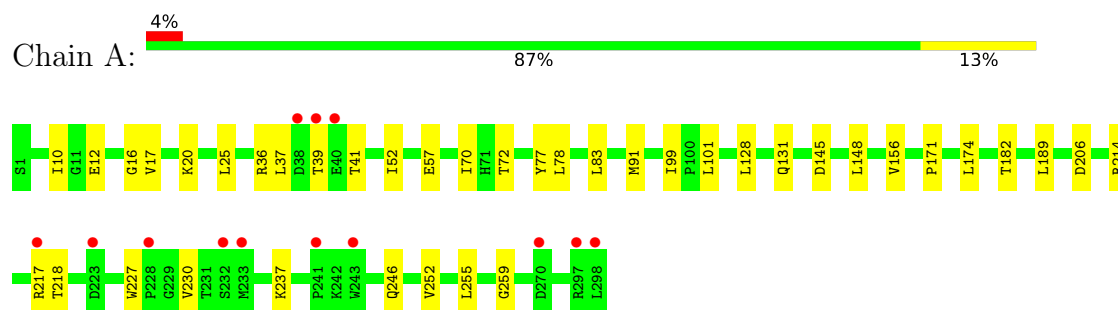
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	25	Total	O	0	0
			25	25		
6	B	3	Total	O	0	0
			3	3		
6	C	3	Total	O	0	0
			3	3		

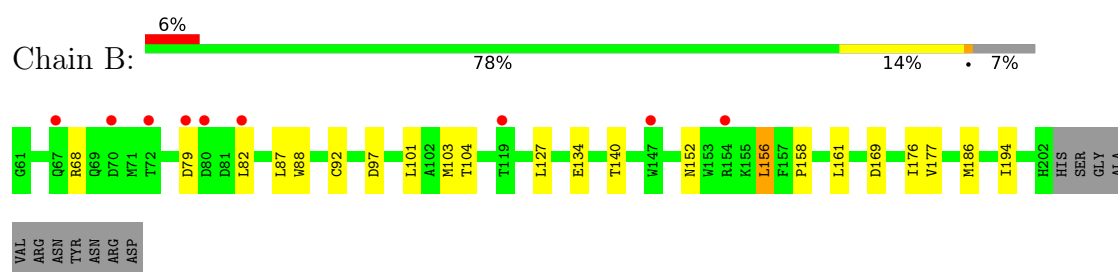
3 Residue-property plots [i](#)

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.

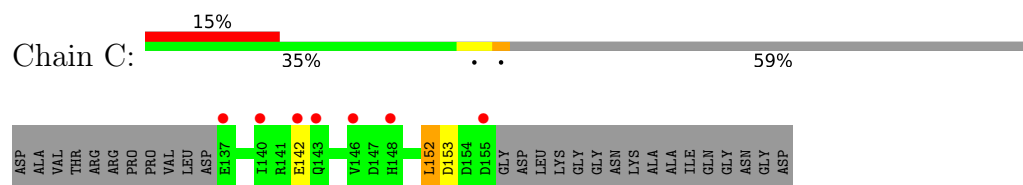
- Molecule 1: Cyclin-dependent kinase 2



- Molecule 2: Speedy protein A



- Molecule 3: SUN domain-containing protein 1



4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α , β , γ	59.34Å 183.53Å 103.83Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.59 – 3.19 49.60 – 3.19	Depositor EDS
% Data completeness (in resolution range)	92.5 (42.59-3.19) 92.5 (49.60-3.19)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.13 (at 3.19Å)	Xtriage
Refinement program	BUSTER 2.10.3	Depositor
R, R_{free}	0.252 , 0.300 0.257 , 0.303	Depositor DCC
R_{free} test set	473 reflections (5.20%)	wwPDB-VP
Wilson B-factor (Å ²)	63.9	Xtriage
Anisotropy	0.127	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 75.4	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.87	EDS
Total number of atoms	3815	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.47% 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: GOL, 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	0.40	0/2457	0.60	0/3336
2	B	0.41	0/1250	0.58	0/1689
3	C	0.35	0/171	0.50	0/233
All	All	0.40	0/3878	0.59	0/5258

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	2395	0	2446	13	0
2	B	1216	0	1191	8	0
3	C	166	0	137	2	0
4	A	6	0	8	0	0
5	B	1	0	0	0	0
6	A	25	0	0	0	0
6	B	3	0	0	0	0
6	C	3	0	0	0	0
All	All	3815	0	3782	20	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 3.

All (20) 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:52:ILE:HD11	1:A:78:LEU:HD11	1.84	0.59
2:B:176:ILE:HG21	3:C:152:LEU:HD23	1.86	0.56
1:A:12:GLU:HG3	1:A:17:VAL:HG22	1.88	0.55
1:A:252:VAL:HG21	1:A:255:LEU:HD12	1.89	0.55
1:A:70:ILE:HB	1:A:77:TYR:HB2	1.90	0.52
2:B:101:LEU:HD13	2:B:194:ILE:HB	1.95	0.49
1:A:171:PRO:HA	1:A:174:LEU:HB2	1.95	0.48
2:B:68:ARG:HB2	3:C:142:GLU:HA	1.96	0.48
1:A:227:TRP:HB3	1:A:230:VAL:HB	1.97	0.47
1:A:255:LEU:HD22	1:A:259:GLY:HA3	1.97	0.46
1:A:10:ILE:HD11	1:A:20:LYS:HB2	1.99	0.44
1:A:128:LEU:HD13	1:A:189:LEU:HG	1.99	0.44
2:B:152:ASN:HB3	2:B:156:LEU:HD12	1.99	0.44
1:A:91:MET:HG2	1:A:99:ILE:HD11	2.00	0.43
2:B:101:LEU:HA	2:B:104:THR:HG22	2.01	0.43
2:B:158:PRO:HA	2:B:161:LEU:HD12	2.00	0.43
2:B:103:MET:HG2	2:B:177:VAL:HB	2.02	0.42
1:A:16:GLY:HA2	1:A:36:ARG:HG3	2.03	0.41
1:A:218:THR:HA	1:A:246:GLN:HE21	1.85	0.41
1:A:41:THR:HB	2:B:161:LEU:HD13	2.04	0.40

There are no symmetry-related clashes.

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	296/298 (99%)	267 (90%)	27 (9%)	2 (1%)	22 61
2	B	140/153 (92%)	138 (99%)	2 (1%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	17/46 (37%)	16 (94%)	1 (6%)	0	100	100
All	All	453/497 (91%)	421 (93%)	30 (7%)	2 (0%)	34	69

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	182	THR
1	A	156	VAL

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	263/263 (100%)	249 (95%)	14 (5%)	22	58
2	B	130/139 (94%)	118 (91%)	12 (9%)	9	33
3	C	18/37 (49%)	16 (89%)	2 (11%)	6	25
All	All	411/439 (94%)	383 (93%)	28 (7%)	16	49

All (28) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	25	LEU
1	A	37	LEU
1	A	39	THR
1	A	57	GLU
1	A	72	THR
1	A	83	LEU
1	A	101	LEU
1	A	131	GLN
1	A	145	ASP
1	A	148	LEU
1	A	206	ASP
1	A	214	ARG
1	A	217	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	237	LYS
2	B	79	ASP
2	B	82	LEU
2	B	87	LEU
2	B	88	TRP
2	B	92	CYS
2	B	97	ASP
2	B	127	LEU
2	B	134	GLU
2	B	140	THR
2	B	156	LEU
2	B	169	ASP
2	B	186	MET
3	C	152	LEU
3	C	153	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	62	ASN
1	A	211	GLN
1	A	246	GLN
2	B	118	HIS
2	B	202	HIS

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 monosaccharides in this entry.

5.6 Ligand geometry

Of 2 ligands modelled in this entry, 1 is monoatomic - leaving 1 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z > 2$	Counts	RMSZ	$\# Z > 2$
4	GOL	A	301	-	5,5,5	0.05	0	5,5,5	0.22	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	GOL	A	301	-	-	0/4/4/4	-

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

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, 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	298/298 (100%)	0.25	13 (4%) 34 21	38, 74, 156, 169	0
2	B	142/153 (92%)	0.45	9 (6%) 20 11	47, 74, 100, 124	0
3	C	19/46 (41%)	1.46	7 (36%) 0 0	72, 91, 111, 122	0
All	All	459/497 (92%)	0.37	29 (6%) 20 11	38, 74, 142, 169	0

All (29) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	39	THR	6.8
3	C	148	HIS	5.3
1	A	223	ASP	5.2
1	A	40	GLU	4.7
1	A	233	MET	4.4
3	C	140	ILE	4.1
1	A	228	PRO	4.0
3	C	137	GLU	3.8
2	B	154	ARG	3.7
1	A	298	LEU	3.6
1	A	38	ASP	3.2
1	A	232	SER	3.1
1	A	243	TRP	3.0
2	B	82	LEU	2.9
2	B	80	ASP	2.9
2	B	119	THR	2.9
1	A	297	ARG	2.9
2	B	72	THR	2.7
2	B	79	ASP	2.5
2	B	70	ASP	2.4
3	C	143	GLN	2.4
1	A	217	ARG	2.3
2	B	147	TRP	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	B	67	GLN	2.2
1	A	241	PRO	2.2
3	C	142	GLU	2.1
3	C	146	VAL	2.1
3	C	155	ASP	2.1
1	A	270	ASP	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
4	GOL	A	301	6/6	0.82	0.16	82,82,82,82	0
5	ZN	B	301	1/1	0.89	0.23	40,40,40,40	1

6.5 Other polymers [i](#)

There are no such residues in this entry.