Full wwPDB X-ray Structure Validation Report ⓘ

May 26, 2020 – 03:15 am BST

PDB ID	:	5Y5A
Title	:	Crystal structure of Est1 and Cdc13
Authors	:	Chen, H.; Xue, J.; Wu, J.; Lei, M.
Deposited on	:	2017-08-08
Resolution	:	2.21 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 ①) were used in the production of this report:

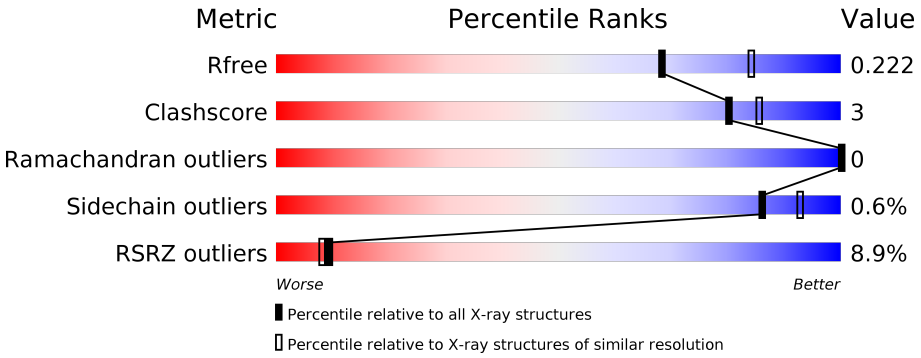
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.21 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 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 <=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	567	8% 84% 7% 9%
2	B	26	15% 65% 35%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4442 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 KLLA0F20702p.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	514	Total	C	N	O	S	0	0	0
			4233	2722	726	757	28			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	THR	deletion	UNP Q6CJ80
A	?	-	THR	deletion	UNP Q6CJ80
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	ALA	deletion	UNP Q6CJ80
A	?	-	ALA	deletion	UNP Q6CJ80
A	?	-	THR	deletion	UNP Q6CJ80
A	?	-	ILE	deletion	UNP Q6CJ80
A	?	-	MET	deletion	UNP Q6CJ80
A	?	-	PRO	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	ASP	deletion	UNP Q6CJ80
A	?	-	LYS	deletion	UNP Q6CJ80
A	?	-	THR	deletion	UNP Q6CJ80
A	?	-	LEU	deletion	UNP Q6CJ80
A	?	-	MET	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	GLY	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	GLY	deletion	UNP Q6CJ80
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	ALA	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	ILE	deletion	UNP Q6CJ80

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80
A	?	-	SER	deletion	UNP Q6CJ80
A	?	-	GLY	deletion	UNP Q6CJ80
A	?	-	ASN	deletion	UNP Q6CJ80

- Molecule 2 is a protein called KLLA0F20922p.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	17	Total	C	N	O	0	0	0
			136	88	21	27			

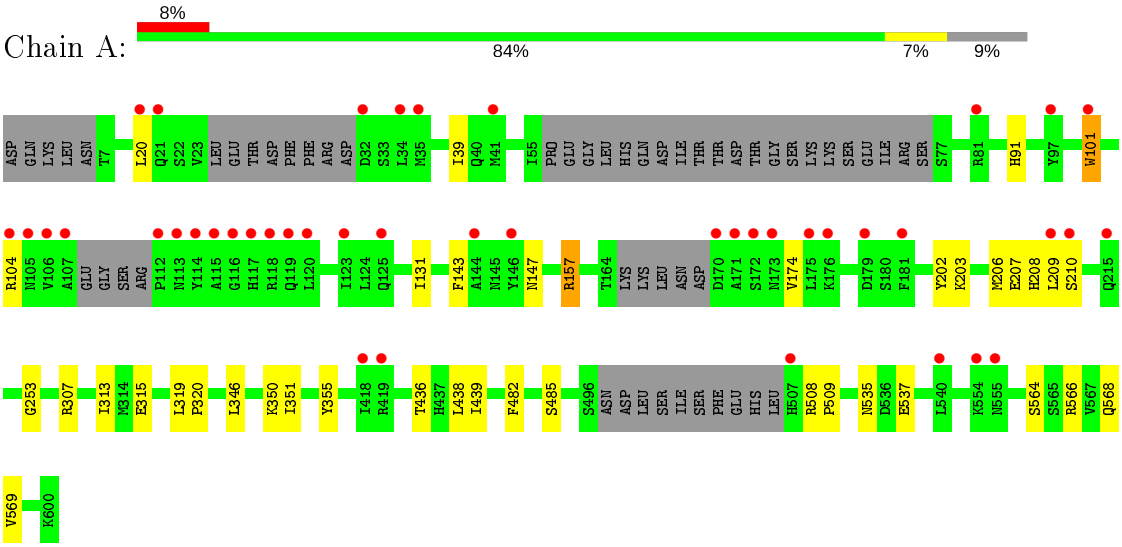
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	72	Total	O	0	0
			72	72		
3	B	1	Total	O	0	0
			1	1		

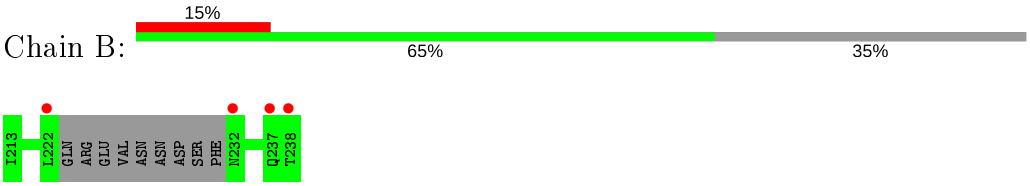
3 Residue-property plots (i)

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: KLLA0F20702p



- Molecule 2: KLLA0F20922p



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	108.61Å 41.73Å 147.06Å 90.00° 97.45° 90.00°	Depositor
Resolution (Å)	40.86 – 2.21 40.86 – 2.21	Depositor EDS
% Data completeness (in resolution range)	91.5 (40.86-2.21) 91.6 (40.86-2.21)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 2.20Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.176 , 0.225 0.184 , 0.222	Depositor DCC
R_{free} test set	1535 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	45.9	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 53.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	4442	wwPDB-VP
Average B, all atoms (Å^2)	65.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.46% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ 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

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	0.32	0/4322	0.44	0/5808
2	B	0.27	0/137	0.46	0/184
All	All	0.32	0/4459	0.44	0/5992

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	A	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	A	157	ARG	Sidechain

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	4233	0	4280	24	0
2	B	136	0	129	0	0
3	A	72	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
All	All	4442	0	4409	24	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 (24) 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:203:LYS:HE2	1:A:207:GLU:OE2	1.95	0.67
1:A:20:LEU:HD21	1:A:39:ILE:HD13	1.79	0.65
1:A:203:LYS:HE2	1:A:207:GLU:CD	2.20	0.62
1:A:143:PHE:O	1:A:157:ARG:NH1	2.39	0.55
1:A:91:HIS:CE1	1:A:131:ILE:HG23	2.42	0.54
1:A:147:ASN:HB3	1:A:174:VAL:HA	1.94	0.50
1:A:313:ILE:HD12	1:A:350:LYS:HE3	1.95	0.49
1:A:209:LEU:HD12	1:A:210:SER:N	2.31	0.45
1:A:509:PRO:HG2	1:A:569:VAL:HA	1.96	0.45
1:A:202:TYR:O	1:A:206:MET:HG2	2.17	0.45
1:A:253:GLY:HA3	1:A:315:GLU:HG2	1.99	0.45
1:A:351:ILE:O	1:A:355:TYR:HB3	2.18	0.44
1:A:319:LEU:HB2	1:A:320:PRO:HD3	1.99	0.44
1:A:209:LEU:HG	1:A:210:SER:H	1.83	0.43
1:A:20:LEU:HD22	1:A:39:ILE:HD11	2.01	0.43
1:A:101:TRP:O	1:A:104:ARG:HB2	2.19	0.42
1:A:439:ILE:HG21	1:A:485:SER:HB3	2.02	0.42
1:A:346:LEU:HA	1:A:346:LEU:HD12	1.91	0.42
1:A:436:THR:HG22	1:A:482:PHE:HA	2.02	0.41
1:A:566:ARG:HA	1:A:569:VAL:HG12	2.02	0.41
1:A:438:LEU:HD23	1:A:438:LEU:HA	1.91	0.40
1:A:535:ASN:OD1	1:A:537:GLU:HB2	2.21	0.40
1:A:508:ARG:HA	1:A:509:PRO:HD3	1.84	0.40
1:A:564:SER:O	1:A:568:GLN:HG3	2.22	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	502/567 (88%)	498 (99%)	4 (1%)	0	100 100
2	B	13/26 (50%)	12 (92%)	1 (8%)	0	100 100
All	All	515/593 (87%)	510 (99%)	5 (1%)	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	480/530 (91%)	477 (99%)	3 (1%)	86 93
2	B	16/25 (64%)	16 (100%)	0	100 100
All	All	496/555 (89%)	493 (99%)	3 (1%)	86 93

All (3) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	101	TRP
1	A	208	HIS
1	A	307	ARG

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

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

There are no ligands in this entry.

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, 95^{th} 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	514/567 (90%)	0.44	43 (8%)	11	9	36, 58, 115, 168	0
2	B	17/26 (65%)	0.93	4 (23%)	0	0	53, 80, 143, 146	0
All	All	531/593 (89%)	0.46	47 (8%)	9	8	36, 59, 117, 168	0

All (47) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	114	TYR	8.0
1	A	171	ALA	7.5
1	A	115	ALA	7.4
1	A	173	ASN	6.7
1	A	170	ASP	6.2
1	A	106	VAL	6.1
1	A	113	ASN	5.1
1	A	105	ASN	5.1
2	B	232	ASN	4.2
1	A	104	ARG	4.0
1	A	35	MET	4.0
1	A	112	PRO	4.0
1	A	507	HIS	4.0
1	A	118	ARG	3.9
1	A	210	SER	3.8
1	A	179	ASP	3.7
2	B	237	GLN	3.4
1	A	97	TYR	3.3
1	A	32	ASP	3.3
1	A	21	GLN	3.2
1	A	119	GLN	3.1
2	B	238	THR	3.1
1	A	41	MET	3.1
1	A	116	GLY	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	117	HIS	3.0
1	A	81	ARG	2.7
1	A	540	LEU	2.7
1	A	418	ILE	2.7
1	A	125	GLN	2.7
1	A	123	ILE	2.7
1	A	554	LYS	2.6
1	A	120	LEU	2.6
1	A	20	LEU	2.6
1	A	34	LEU	2.6
2	B	222	LEU	2.6
1	A	419	ARG	2.6
1	A	107	ALA	2.5
1	A	101	TRP	2.4
1	A	176	LYS	2.2
1	A	181	PHE	2.2
1	A	209	LEU	2.2
1	A	172	SER	2.2
1	A	144	ALA	2.1
1	A	146	TYR	2.1
1	A	215	GLN	2.1
1	A	555	ASN	2.1
1	A	175	LEU	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.