



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 08:20 PM EDT

PDB ID : 5Y58  
Title : Crystal structure of Ku70/80 and TLC1  
Authors : Chen, H.; Xue, J.; Wu, J.; Lei, M.  
Deposited on : 2017-08-08  
Resolution : 2.80 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](mailto: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 types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.37.1  
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) : 2.37.1

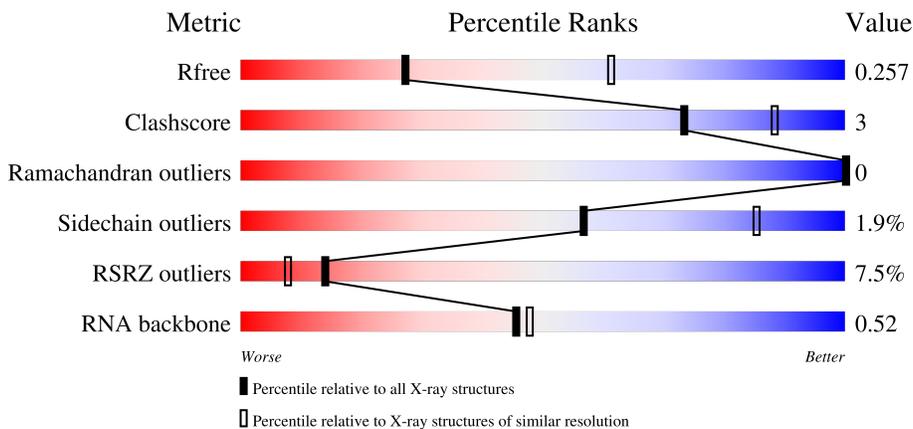
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

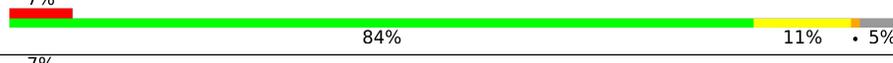
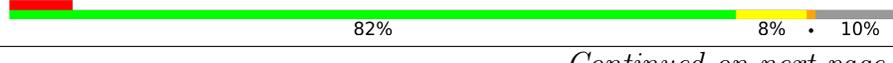
The reported resolution of this entry is 2.80 Å.

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.



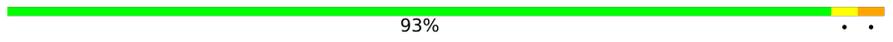
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)
RNA backbone	3102	1227 (3.10-2.50)

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 of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 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	575	 7% 83% 11% • 5%
1	C	575	 7% 84% 10% • 5%
1	E	575	 7% 84% 11% • 5%
2	B	628	 7% 82% 8% • 10%

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	628	 <p>9% 79% 9% 11%</p>
2	F	628	 <p>15% 81% 8% 10%</p>
3	X	30	 <p>93% . .</p>
3	Y	30	 <p>87% 10% .</p>
3	Z	30	 <p>87% 10% .</p>

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 29320 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 ATP-dependent DNA helicase II subunit 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	548	Total	C	N	O	S	0	0	0
			4549	2926	743	863	17			
1	C	548	Total	C	N	O	S	0	0	0
			4549	2926	743	863	17			
1	E	548	Total	C	N	O	S	0	0	0
			4549	2926	743	863	17			

- Molecule 2 is a protein called ATP-dependent DNA helicase II subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	568	Total	C	N	O	S	0	0	0
			4509	2849	750	887	23			
2	D	560	Total	C	N	O	S	0	0	0
			4442	2812	737	870	23			
2	F	565	Total	C	N	O	S	0	0	0
			4493	2841	747	882	23			

- Molecule 3 is a RNA chain called TLC1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	X	30	Total	C	N	O	P	0	0	0
			641	287	116	208	30			
3	Y	30	Total	C	N	O	P	0	0	0
			641	287	116	208	30			
3	Z	30	Total	C	N	O	P	0	0	0
			641	287	116	208	30			

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	60	Total	0	0
			60		

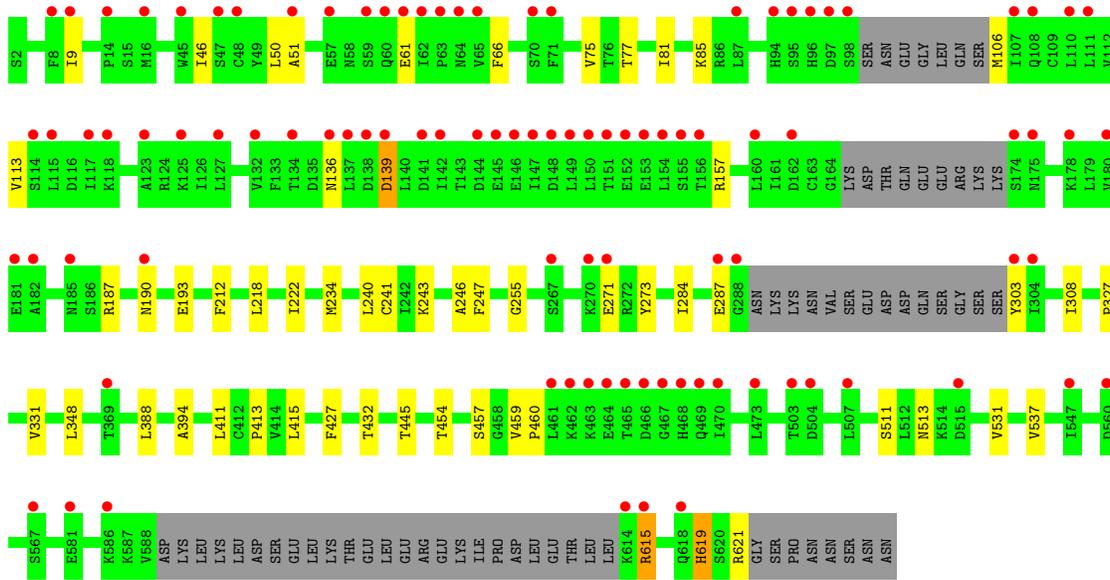
*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
4	B	66	Total O 66 66	0	0
4	X	14	Total O 14 14	0	0
4	C	19	Total O 19 19	0	0
4	D	48	Total O 48 48	0	0
4	Y	13	Total O 13 13	0	0
4	E	34	Total O 34 34	0	0
4	F	45	Total O 45 45	0	0
4	Z	7	Total O 7 7	0	0



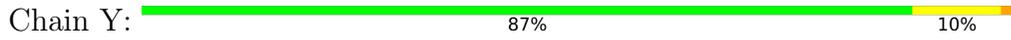




• Molecule 3: TLC1



• Molecule 3: TLC1



• Molecule 3: TLC1



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	116.30Å 115.16Å 115.85Å 77.43° 78.48° 63.73°	Depositor
Resolution (Å)	44.11 – 2.80 48.33 – 2.78	Depositor EDS
% Data completeness (in resolution range)	91.8 (44.11-2.80) 90.6 (48.33-2.78)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.70 (at 2.77Å)	Xtrriage
Refinement program	PHENIX 1.9_1692	Depositor
R, $R_{free}$	0.215 , 0.258 0.218 , 0.257	Depositor DCC
$R_{free}$ test set	5915 reflections (4.97%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	43.4	Xtrriage
Anisotropy	0.016	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 55.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.022 for -k,-h,-l	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	29320	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	68.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.29% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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

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	0.24	0/4647	0.37	0/6253
1	C	0.26	0/4647	0.38	0/6253
1	E	0.25	0/4647	0.37	0/6253
2	B	0.23	0/4583	0.41	0/6203
2	D	0.24	0/4514	0.41	0/6109
2	F	0.25	0/4567	0.42	0/6182
3	X	0.24	0/717	0.72	0/1115
3	Y	0.23	0/717	0.74	0/1115
3	Z	0.29	0/717	0.72	0/1115
All	All	0.25	0/29756	0.43	0/40598

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	B	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	B	258	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	4549	0	4523	37	0
1	C	4549	0	4523	37	0
1	E	4549	0	4523	36	0
2	B	4509	0	4542	27	0
2	D	4442	0	4488	38	0
2	F	4493	0	4529	33	0
3	X	641	0	323	1	0
3	Y	641	0	323	2	0
3	Z	641	0	323	2	0
4	A	60	0	0	0	0
4	B	66	0	0	0	0
4	C	19	0	0	0	0
4	D	48	0	0	0	0
4	E	34	0	0	0	0
4	F	45	0	0	0	0
4	X	14	0	0	0	0
4	Y	13	0	0	0	0
4	Z	7	0	0	0	0
All	All	29320	0	28097	188	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.

The worst 5 of 188 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:E:30:GLU:HG3	1:E:165:ASN:HB3	1.67	0.74
2:B:212:PHE:HB3	2:B:246:ALA:HB3	1.73	0.70
1:C:506:LYS:HE3	2:D:333:GLN:O	1.91	0.70
1:E:526:GLU:OE2	1:E:534:ARG:HD3	1.91	0.69
2:F:615:ARG:HH21	2:F:619:HIS:HB2	1.57	0.69

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	544/575 (95%)	529 (97%)	15 (3%)	0	100	100
1	C	544/575 (95%)	526 (97%)	18 (3%)	0	100	100
1	E	544/575 (95%)	527 (97%)	17 (3%)	0	100	100
2	B	558/628 (89%)	539 (97%)	19 (3%)	0	100	100
2	D	548/628 (87%)	531 (97%)	17 (3%)	0	100	100
2	F	555/628 (88%)	536 (97%)	19 (3%)	0	100	100
All	All	3293/3609 (91%)	3188 (97%)	105 (3%)	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	510/536 (95%)	500 (98%)	10 (2%)	55	84
1	C	510/536 (95%)	501 (98%)	9 (2%)	59	86
1	E	510/536 (95%)	500 (98%)	10 (2%)	55	84
2	B	527/585 (90%)	517 (98%)	10 (2%)	57	85
2	D	519/585 (89%)	508 (98%)	11 (2%)	53	84
2	F	525/585 (90%)	517 (98%)	8 (2%)	65	89
All	All	3101/3363 (92%)	3043 (98%)	58 (2%)	57	85

5 of 58 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	433	SER
2	F	457	SER
2	D	445	THR
2	F	445	THR
1	E	477	MET

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	229	ASN
2	F	619	HIS

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
3	X	29/30 (96%)	2 (6%)	0
3	Y	29/30 (96%)	2 (6%)	0
3	Z	29/30 (96%)	2 (6%)	0
All	All	87/90 (96%)	6 (6%)	0

5 of 6 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
3	X	301	U
3	X	302	A
3	Y	301	U
3	Y	302	A
3	Z	301	U

There are no RNA pucker outliers to report.

## 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](#)

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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	548/575 (95%)	-0.16	7 (1%) 77 72	16, 51, 99, 142	0
1	C	548/575 (95%)	-0.27	8 (1%) 73 68	17, 46, 94, 150	0
1	E	548/575 (95%)	0.45	39 (7%) 16 9	29, 71, 130, 181	0
2	B	568/628 (90%)	0.22	47 (8%) 11 6	16, 64, 143, 204	0
2	D	560/628 (89%)	0.29	58 (10%) 6 3	18, 64, 149, 189	0
2	F	565/628 (89%)	0.85	97 (17%) 1 1	22, 82, 163, 200	0
3	X	30/30 (100%)	-0.21	0 100 100	32, 55, 109, 130	0
3	Y	30/30 (100%)	-0.13	0 100 100	34, 57, 100, 122	0
3	Z	30/30 (100%)	-0.14	0 100 100	35, 55, 116, 126	0
All	All	3427/3699 (92%)	0.22	256 (7%) 14 8	16, 60, 140, 204	0

The worst 5 of 256 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	F	145	GLU	8.7
2	F	60	GLN	8.5
2	B	150	LEU	8.3
2	B	97	ASP	7.6
2	F	59	SER	7.4

### 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

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.