



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 10:22 PM GMT

PDB ID : 5CYK  
Title : Structure of Ytm1 bound to the C-terminal domain of Erb1-R486E  
Authors : Wegrecki, M.; Bravo, J.  
Deposited on : 2015-07-30  
Resolution : 3.00 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  
<http://wwpdb.org/validation/2016/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 : unknown  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026982  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : rb-20026982

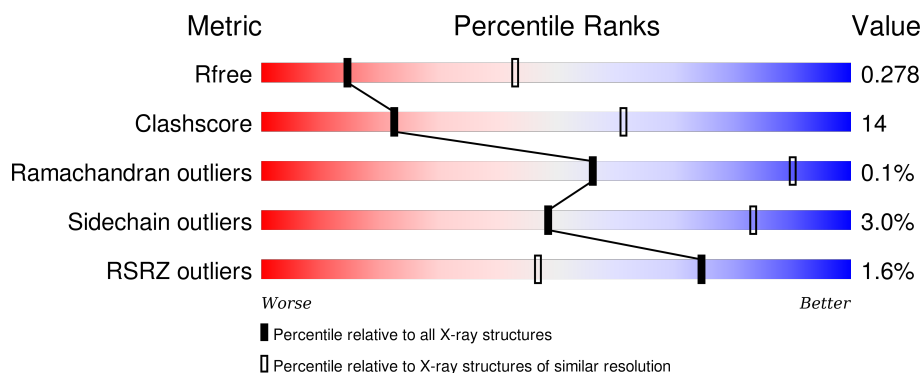
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 3.00 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))
$R_{free}$	91344	1578 (3.00-3.00)
Clashscore	102246	1912 (3.00-3.00)
Ramachandran outliers	100387	1853 (3.00-3.00)
Sidechain outliers	100360	1856 (3.00-3.00)
RSRZ outliers	91569	1592 (3.00-3.00)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. 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	495	<div> <div> <div></div> <div>61%</div> <div>28%</div> <div>•</div> <div>8%</div> </div> </div>
2	B	389	<div> <div> <div>2%</div> <div>67%</div> <div>21%</div> <div>•</div> <div>10%</div> </div> </div>

## 2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 6182 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 Ribosome biogenesis protein YTM1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	454	Total	C	N	O	S	0	0	0
			3424	2146	599	673	6			

- Molecule 2 is a protein called Ribosome biogenesis protein ERB1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	350	Total	C	N	O	S	0	0	0
			2753	1755	502	490	6			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	413	MET	-	initiating methionine	UNP G0SCK6
B	414	ALA	-	expression tag	UNP G0SCK6
B	415	HIS	-	expression tag	UNP G0SCK6
B	416	HIS	-	expression tag	UNP G0SCK6
B	417	HIS	-	expression tag	UNP G0SCK6
B	418	HIS	-	expression tag	UNP G0SCK6
B	419	HIS	-	expression tag	UNP G0SCK6
B	420	HIS	-	expression tag	UNP G0SCK6
B	421	SER	-	expression tag	UNP G0SCK6
B	422	ALA	-	expression tag	UNP G0SCK6
B	423	ALA	-	expression tag	UNP G0SCK6
B	424	LEU	-	expression tag	UNP G0SCK6
B	425	GLU	-	expression tag	UNP G0SCK6
B	426	VAL	-	expression tag	UNP G0SCK6
B	427	LEU	-	expression tag	UNP G0SCK6
B	428	PHE	-	expression tag	UNP G0SCK6
B	429	GLN	-	expression tag	UNP G0SCK6
B	430	GLY	-	expression tag	UNP G0SCK6
B	431	PRO	-	expression tag	UNP G0SCK6
B	432	GLY	-	expression tag	UNP G0SCK6

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	486	GLU	ARG	engineered mutation	UNP G0SCK6

- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	3	Total Cl 3 3	0	0
3	A	2	Total Cl 2 2	0	0



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	170.75Å 170.75Å 155.66Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	85.38 – 3.00 85.38 – 3.00	Depositor EDS
% Data completeness (in resolution range)	98.1 (85.38-3.00) 98.6 (85.38-3.00)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.46 (at 3.01Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, $R_{free}$	0.202 , 0.262 0.219 , 0.278	Depositor DCC
$R_{free}$ test set	1346 reflections (5.25%)	DCC
Wilson B-factor (Å <sup>2</sup> )	86.9	Xtriage
Anisotropy	0.006	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 62.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 26967 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	6182	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	89.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.85% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: CL

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	0.34	1/3497 (0.0%)	0.69	7/4759 (0.1%)
2	B	0.31	0/2827	0.60	2/3843 (0.1%)
All	All	0.33	1/6324 (0.0%)	0.66	9/8602 (0.1%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	37	VAL	CA-CB	6.10	1.67	1.54

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	754	LEU	CA-CB-CG	7.51	132.56	115.30
1	A	87	LEU	CA-CB-CG	7.00	131.40	115.30
1	A	85	LEU	CA-CB-CG	6.58	130.43	115.30
1	A	35	LEU	CA-CB-CG	6.05	129.21	115.30
2	B	773	LYS	CD-CE-NZ	-5.82	98.32	111.70

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	3424	0	3381	106	1
2	B	2753	0	2759	66	1
3	A	2	0	0	2	0
3	B	3	0	0	0	0
All	All	6182	0	6140	169	1

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 14.

The worst 5 of 169 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)
2:B:588:ARG:HG3	2:B:613:SER:HB2	1.57	0.86
2:B:478:ARG:HB2	2:B:480:TRP:HE1	1.41	0.85
2:B:634:GLY:HA2	2:B:654:ARG:HH21	1.46	0.81
1:A:287:ARG:NH2	3:A:502:CL:CL	2.55	0.75
1:A:355:SER:OG	1:A:373:ARG:NH2	2.19	0.75

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:257:ALA:O	2:B:645:ARG:NH1[8_555]	2.13	0.07

## 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	444/495 (90%)	416 (94%)	27 (6%)	1 (0%)	52	88
2	B	346/389 (89%)	325 (94%)	21 (6%)	0	100	100
All	All	790/884 (89%)	741 (94%)	48 (6%)	1 (0%)	56	90



All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	473	ILE

### 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	376/410 (92%)	364 (97%)	12 (3%)	46	82
2	B	296/325 (91%)	288 (97%)	8 (3%)	52	85
All	All	672/735 (91%)	652 (97%)	20 (3%)	48	83

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	355	SER
1	A	420	SER
2	B	654	ARG
1	A	293	TRP
1	A	354	THR

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

Mol	Chain	Res	Type
2	B	611	GLN
2	B	753	GLN

### 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 [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 5 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, 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	454/495 (91%)	-0.01	6 (1%)	79 53	46, 79, 151, 199	2 (0%)
2	B	350/389 (89%)	0.02	7 (2%)	68 39	50, 85, 154, 203	3 (0%)
All	All	804/884 (90%)	-0.00	13 (1%)	74 47	46, 82, 152, 203	5 (0%)

The worst 5 of 13 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	14	GLN	5.1
2	B	632	LEU	4.3
1	A	17	VAL	3.7
2	B	613	SER	3.5
1	A	86	SER	3.2

### 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](#)

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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95<sup>th</sup> 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	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	CL	B	903	1/1	0.95	0.09	-	87,87,87,87	0
3	CL	B	902	1/1	0.97	0.32	-	75,75,75,75	0
3	CL	B	901	1/1	0.86	0.18	-	93,93,93,93	0
3	CL	A	501	1/1	0.82	0.22	-	83,83,83,83	0
3	CL	A	502	1/1	0.87	0.23	-	82,82,82,82	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.