



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 15, 2016 – 01:11 PM EDT

PDB ID : 5L3T  
Title : Structure of the *Saccharomyces cerevisiae* TREX-2 complex  
Authors : Stewart, M.; Aibara, S.  
Deposited on : 2016-05-24  
Resolution : 4.93 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-20027939  
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-20027939

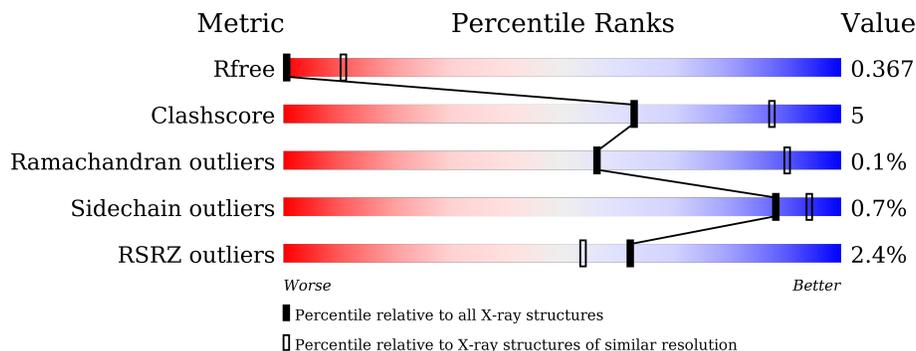
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 4.93 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	1119 (6.22-3.60)
Clashscore	102246	1018 (6.10-3.66)
Ramachandran outliers	100387	1158 (6.22-3.60)
Sidechain outliers	100360	1136 (6.22-3.60)
RSRZ outliers	91569	1122 (6.22-3.60)

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	1301	
2	B	455	
3	C	89	
4	D	24	
4	E	24	
5	F	12	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 6831 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 Nuclear mRNA export protein SAC3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	294	2438	1564	420	441	13	0	0	0

- Molecule 2 is a protein called Nuclear mRNA export protein THP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	455	3714	2392	645	659	18	0	0	0

- Molecule 3 is a protein called 26S proteasome complex subunit SEM1.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	48	419	258	64	97	0	0	0

- Molecule 4 is a protein called Sac3polyAla.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	D	17	85	51	17	17	0	0	0
4	E	24	120	72	24	24	0	0	0

- Molecule 5 is a protein called Sac3polyAla.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
5	F	11	55	33	11	11	0	0	0







- Molecule 5: Sac3polyAla

Chain F:   
83% 8% 8%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	78.17Å 84.09Å 165.48Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	47.08 – 4.93 47.08 – 4.93	Depositor EDS
% Data completeness (in resolution range)	57.1 (47.08-4.93) 46.4 (47.08-4.93)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	0.91 (at 4.85Å)	Xtrriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, $R_{free}$	0.338 , 0.370 0.333 , 0.367	Depositor DCC
$R_{free}$ test set	173 reflections (5.74%)	DCC
Wilson B-factor (Å <sup>2</sup> )	68.5	Xtrriage
Anisotropy	0.208	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 13.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.65	EDS
Total number of atoms	6831	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	80.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.08% 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 [i](#)

### 5.1 Standard geometry [i](#)

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	0/2492	0.43	1/3368 (0.0%)
2	B	0.26	0/3804	0.40	0/5168
3	C	0.24	0/426	0.49	1/575 (0.2%)
All	All	0.25	0/6722	0.42	2/9111 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	350	ASP	CB-CG-OD2	5.21	122.99	118.30
3	C	41	ASP	CB-CG-OD2	5.21	122.99	118.30

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	2438	0	2449	38	2
2	B	3714	0	3754	31	1
3	C	419	0	355	4	0
4	D	85	0	19	0	0
4	E	120	0	26	1	0
5	F	55	0	13	1	0
All	All	6831	0	6616	71	2

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 5.

The worst 5 of 71 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:276:LEU:HB3	1:A:277:PRO:HD3	1.73	0.70
3:C:38:PHE:HE1	3:C:40:ILE:HD13	1.57	0.69
1:A:266:ASP:OD1	1:A:310:ARG:NH2	2.28	0.67
1:A:256:ARG:HG2	5:F:1:UNK:N	2.10	0.66
1:A:326:ASN:O	1:A:329:PHE:CZ	2.48	0.66

All (2) 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:288:ARG:NH2	2:B:213:GLU:OE2[4_545]	2.11	0.09
1:A:271:ASN:OD1	1:A:486:ARG:NH2[3_654]	2.16	0.04

## 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	292/1301 (22%)	290 (99%)	1 (0%)	1 (0%)	46	83
2	B	453/455 (100%)	444 (98%)	9 (2%)	0	100	100
3	C	44/89 (49%)	44 (100%)	0	0	100	100
All	All	789/1845 (43%)	778 (99%)	10 (1%)	1 (0%)	56	90

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	276	LEU

### 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	274/1212 (23%)	272 (99%)	2 (1%)	88	94
2	B	419/419 (100%)	416 (99%)	3 (1%)	88	94
3	C	47/81 (58%)	47 (100%)	0	100	100
All	All	740/1712 (43%)	735 (99%)	5 (1%)	88	94

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	404	ARG
1	A	445	PHE
2	B	323	ARG
2	B	370	ARG
2	B	402	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	139	HIS

### 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 carbohydrates 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	294/1301 (22%)	0.18	4 (1%) 78 70	30, 76, 101, 134	0
2	B	455/455 (100%)	0.23	8 (1%) 71 63	44, 82, 117, 144	0
3	C	48/89 (53%)	0.86	7 (14%) 3 5	61, 93, 114, 123	0
4	D	0/24	-	-	-	-
4	E	0/24	-	-	-	-
5	F	0/12	-	-	-	-
All	All	797/1905 (41%)	0.25	19 (2%) 62 54	30, 80, 114, 144	0

The worst 5 of 19 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	66	ASP	3.4
2	B	13	HIS	3.2
1	A	468	LYS	3.2
3	C	30	GLU	3.1
3	C	75	ASN	3.0

### 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

There are no such residues in this entry.