



wwPDB X-ray Structure Validation Summary Report ⓘ

May 16, 2020 – 04:52 am BST

PDB ID : 3PSF
Title : Crystal Structure of the Spt6 core domain from *Saccharomyces cerevisiae*,
Form Spt6(236-1259)
Authors : Close, D.; Hill, C.P.
Deposited on : 2010-12-01
Resolution : 2.59 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

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
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

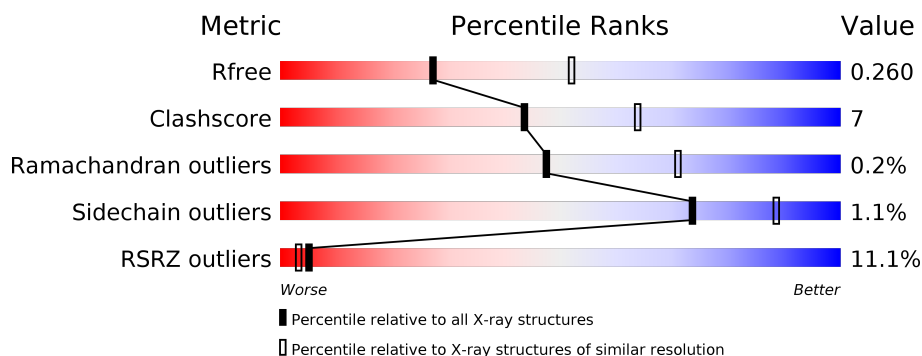
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.59 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	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 ≥ 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	1030	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 13444 atoms, of which 6655 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 Transcription elongation factor SPT6.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	826	Total	C	H	N	O	S	0	0	0
			13396	4272	6655	1134	1317	18			

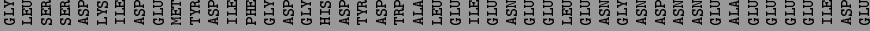
There are 5 discrepancies between the modelled and reference sequences:

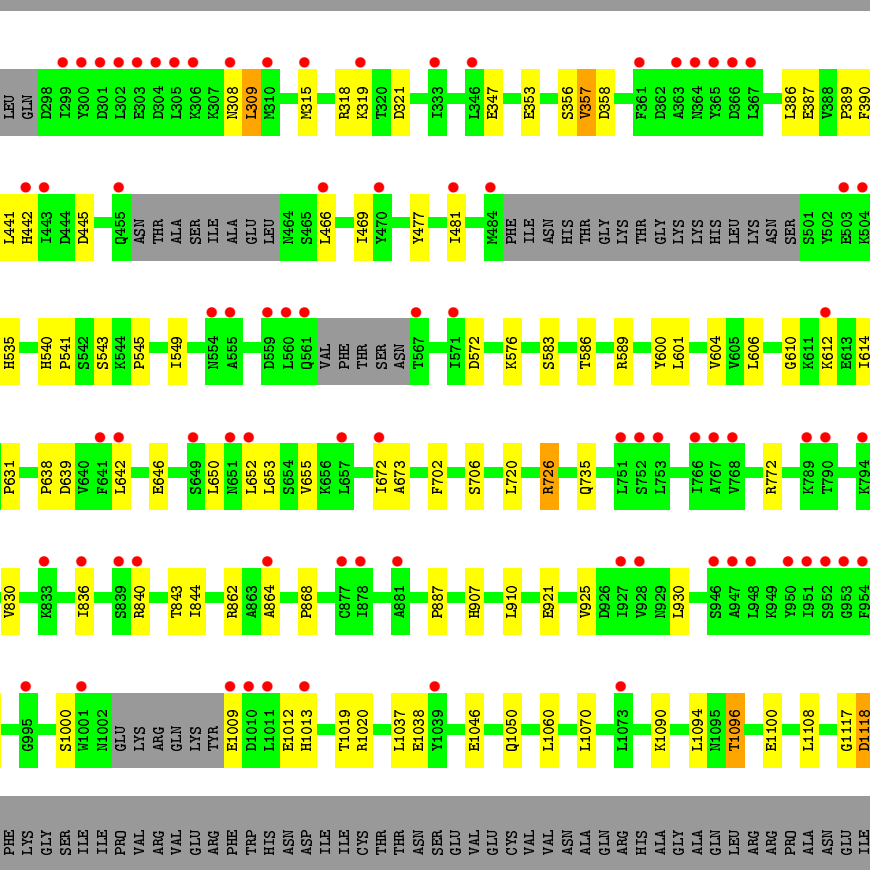
Chain	Residue	Modelled	Actual	Comment	Reference
A	230	GLY	-	EXPRESSION TAG	UNP P23615
A	231	ILE	-	EXPRESSION TAG	UNP P23615
A	232	ASP	-	EXPRESSION TAG	UNP P23615
A	233	PRO	-	EXPRESSION TAG	UNP P23615
A	234	PHE	-	EXPRESSION TAG	UNP P23615

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	48	Total	O	0	0
			48	48		

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.

- Chain A: 



Protein	Sequence
PRO	E1127
ALA	S1128
LYS	GLU
VAL	L974
ILE	P817
THR	P818
PHE	P819
PHE	P820
LYS	P821
LYS	P822
GLY	P823
ALA	P824
ASN	P825
ILE	P826
ILE	P827
PRO	P828
ALA	P829
VAL	P830
VAL	P831
GLU	P832
LYS	P833
GLN	P834
ARG	P835
GLU	P836
GLU	P837
GLU	P838
GLU	P839
GLU	P840
GLU	P841
GLU	P842
GLU	P843
GLU	P844
GLU	P845
GLU	P846
GLU	P847
GLU	P848
GLU	P849
GLU	P850
GLU	P851
GLU	P852
GLU	P853
GLU	P854
GLU	P855
GLU	P856
GLU	P857
GLU	P858
GLU	P859
GLU	P860
GLU	P861
GLU	P862
GLU	P863
GLU	P864
GLU	P865
GLU	P866
GLU	P867
GLU	P868
GLU	P869
GLU	P870
GLU	P871
GLU	P872
GLU	P873
GLU	P874
GLU	P875
GLU	P876
GLU	P877
GLU	P878
GLU	P879
GLU	P880
GLU	P881
GLU	P882
GLU	P883
GLU	P884
GLU	P885
GLU	P886
GLU	P887
GLU	P888
GLU	P889
GLU	P890
GLU	P891
GLU	P892
GLU	P893
GLU	P894
GLU	P895
GLU	P896
GLU	P897
GLU	P898
GLU	P899
GLU	P900
GLU	P901
GLU	P902
GLU	P903
GLU	P904
GLU	P905
GLU	P906
GLU	P907
GLU	P908
GLU	P909
GLU	P910
GLU	P911
GLU	P912
GLU	P913
GLU	P914
GLU	P915
GLU	P916
GLU	P917
GLU	P918
GLU	P919
GLU	P920
GLU	P921
GLU	P922
GLU	P923
GLU	P924
GLU	P925
GLU	P926
GLU	P927
GLU	P928
GLU	P929
GLU	P930
GLU	P931
GLU	P932
GLU	P933
GLU	P934
GLU	P935
GLU	P936
GLU	P937
GLU	P938
GLU	P939
GLU	P940
GLU	P941
GLU	P942
GLU	P943
GLU	P944
GLU	P945
GLU	P946
GLU	P947
GLU	P948
GLU	P949
GLU	P950
GLU	P951
GLU	P952
GLU	P953
GLU	P954
GLU	P955
GLU	P956
GLU	P957
GLU	P958
GLU	P959
GLU	P960
GLU	P961
GLU	P962
GLU	P963
GLU	P964
GLU	P965
GLU	P966
GLU	P967
GLU	P968
GLU	P969
GLU	P970
GLU	P971
GLU	P972
GLU	P973
GLU	P974
GLU	P975
GLU	P976
GLU	P977
GLU	P978
GLU	P979
GLU	P980
GLU	P981
GLU	P982
GLU	P983
GLU	P984
GLU	P985
GLU	P986
GLU	P987
GLU	P988
GLU	P989
GLU	P990
GLU	P991
GLU	P992
GLU	P993
GLU	P994
GLU	P995
GLU	P996
GLU	P997
GLU	P998
GLU	P999
GLU	P1000
GLU	P1001
GLU	P1002
GLU	P1003
GLU	P1004
GLU	P1005

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	115.05Å 116.18Å 117.35Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	32.32 – 2.59 32.32 – 2.59	Depositor EDS
% Data completeness (in resolution range)	96.5 (32.32-2.59) 96.0 (32.32-2.59)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.21 (at 2.57Å)	Xtriage
Refinement program	PHENIX 1.6.4_486	Depositor
R, R_{free}	0.224 , 0.265 0.216 , 0.260	Depositor DCC
R_{free} test set	1978 reflections (3.99%)	wwPDB-VP
Wilson B-factor (Å ²)	65.1	Xtriage
Anisotropy	0.581	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 56.3	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	0.020 for -h,l,k 0.017 for -l,-k,-h 0.016 for k,h,-l 0.007 for k,l,h 0.007 for l,h,k	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	13444	wwPDB-VP
Average B, all atoms (Å ²)	103.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.28% 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 [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	0.33	0/6869	0.48	0/9278

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 [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	6741	6655	6635	89	0
2	A	48	0	0	2	0
All	All	6789	6655	6635	89	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 7.

The worst 5 of 89 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:626:ALA:HB1	1:A:629:ARG:HE	1.60	0.65
1:A:532:HIS:NE2	1:A:534:ILE:HG22	2.15	0.61
1:A:357:VAL:HG12	1:A:358:ASP:N	2.16	0.60
1:A:412:ASP:OD1	1:A:726:ARG:NH1	2.35	0.59
1:A:817:ASN:HB2	1:A:818:PRO:CD	2.33	0.58

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	814/1030 (79%)	772 (95%)	40 (5%)	2 (0%)	47 71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	616	LYS
1	A	357	VAL

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	748/929 (80%)	740 (99%)	8 (1%)	73 88

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	813	ILE
1	A	1118	ASP
1	A	1013	HIS
1	A	726	ARG
1	A	1000	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	540	HIS
1	A	1013	HIS
1	A	575	GLN
1	A	535	HIS
1	A	885	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, 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	826/1030 (80%)	0.73	92 (11%) 5 3	57, 89, 148, 201	0

The worst 5 of 92 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1011	LEU	6.0
1	A	560	LEU	5.8
1	A	504	LYS	5.7
1	A	481	ILE	5.7
1	A	649	SER	5.5

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.