



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 14, 2017 – 08:00 pm GMT

PDB ID : 3PSI
Title : Crystal Structure of the Spt6 core domain from *Saccharomyces cerevisiae*,
Form Spt6(239-1451)
Authors : Close, D.; Hill, C.P.; Johnson, S.J.
Deposited on : 2010-12-01
Resolution : 3.30 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

<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
Xtriage (Phenix)	:	1.9-1692
EDS	:	trunk28620
Percentile statistics	:	20161228.v01 (using entries in the PDB archive December 28th 2016)
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)	:	recalc28949

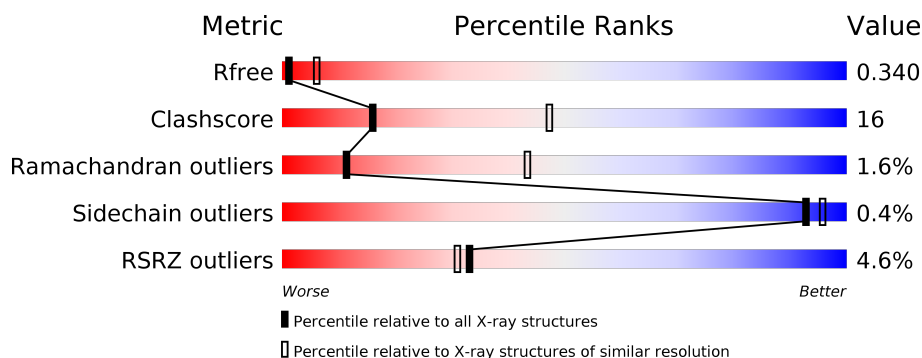
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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}	100719	1034 (3.36-3.24)
Clashscore	112137	1100 (3.36-3.24)
Ramachandran outliers	110173	1081 (3.36-3.24)
Sidechain outliers	110143	1080 (3.36-3.24)
RSRZ outliers	101464	1039 (3.36-3.24)

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. 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	1219	<div> <div>3%</div> <div> <div></div> <div>48%</div> <div>20%</div> <div>•</div> <div>31%</div> </div> </div>

2 Entry composition

There is only 1 type of molecule in this entry. The entry contains 13677 atoms, of which 6800 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	843	Total	C	H	N	O	S	0	0	0
			13677	4370	6800	1164	1326	17			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLY	-	EXPRESSION TAG	UNP P23615
A	234	ILE	-	EXPRESSION TAG	UNP P23615
A	235	ASP	-	EXPRESSION TAG	UNP P23615
A	236	PRO	-	EXPRESSION TAG	UNP P23615
A	237	PHE	-	EXPRESSION TAG	UNP P23615
A	238	THR	-	EXPRESSION TAG	UNP P23615

ASN	ASN	VAL
ALA	ILE	ILE
ASN	GLN	ARG
SER	ASN	GLN
LYS	LYS	SER
LEU	VAL	SER
TYR	ARG	ARG
THR	LEU	GLY
TRP	LEU	ASP
ASN	ASN	ASP
VAL	GLU	HIS
LYS	MET	LEU
LEU	THR	VAL
THR	SER	ILE
ASN	GLU	THR
THR	LYS	TRP
GLY	PHE	LYS
TYR	LYS	LEU
PHE	SER	ASP
LEU	GLY	LYS
VAL	THR	ASP
ASN	LYS	LEU
TYR	LYS	PHE
ASN	ASP	GLN
TYR	VAL	HIS
PRO	VAL	ILE
SER	LYS	ASP
VAL	PHE	ILE
ILE	ILE	GLN
GLN	GLU	GLU
LEU	ASP	LEU
CYS	TYR	GLU
ASN	SER	LYS
GLY	ARG	GLU
PHE	VAL	ASN
LYS	ASN	PRO
THR	PRO	LEU
LEU	ASN	ALA
LEU	LYS	LEU
LYS	SER	GLY
SER	VAL	LYS
ASN	TYR	VAL
SER	TYR	VAL
SER	PHE	LEU
LYS	SER	ILE
ASN	LEU	VAL
ARG	ASP	ASP
MET	HIS	ASN
ASN	LYS	GLN
ASN	ASP	LYS
TYR	ASN	TYR
ARG	PRO	ASN
	GLY	ASP
	TRP	LEU
	PHE	ASP
	TYR	GLN
	LEU	ILE
	MET	ILE
	PHE	VAL
	LYS	GLU
	ILE	TYR

4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, α , β , γ	118.75Å 118.75Å 214.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.67 – 3.30 45.67 – 3.29	Depositor EDS
% Data completeness (in resolution range)	99.4 (45.67-3.30) 99.4 (45.67-3.29)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.44 (at 3.32Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.6.4_486)	Depositor
R, R_{free}	0.265 , 0.308 0.247 , 0.340	Depositor DCC
R_{free} test set	2827 reflections (10.53%)	DCC
Wilson B-factor (Å ²)	111.6	Xtriage
Anisotropy	0.000	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 136.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.017 for -h,-k,l	Xtriage
F_o, F_c correlation	0.91	EDS
Total number of atoms	13677	wwPDB-VP
Average B, all atoms (Å ²)	170.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.75% 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	0.62	0/7015	0.77	2/9487 (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	1020	ARG	NE-CZ-NH1	-7.92	116.34	120.30
1	A	1101	LEU	CB-CG-CD1	-5.88	101.00	111.00

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	6877	6800	6777	221	0
All	All	6877	6800	6777	221	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 16.

The worst 5 of 221 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:980:ILE:HD12	1:A:981:THR:N	1.84	0.92

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:601:LEU:HD11	1:A:658:HIS:HB2	1.61	0.83
1:A:980:ILE:HD12	1:A:981:THR:H	1.41	0.81
1:A:641:PHE:CE1	1:A:713:LEU:HD23	2.21	0.75
1:A:653:LEU:HD12	1:A:653:LEU:C	2.07	0.74

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	831/1219 (68%)	760 (92%)	58 (7%)	13 (2%)	11 43

5 of 13 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	1117	GLY
1	A	742	VAL
1	A	551	SER
1	A	628	ASN
1	A	762	ALA

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.

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
-----	-------	----------	-----------	----------	-------------

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	A	762/1108 (69%)	759 (100%)	3 (0%)	93 96

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

Mol	Chain	Res	Type
1	A	689	PHE
1	A	726	ARG
1	A	1204	LEU

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	909	ASN
1	A	1147	HIS
1	A	982	HIS
1	A	429	ASN
1	A	1074	ASN

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 [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	843/1219 (69%)	0.27	39 (4%)	33 31	74, 150, 305, 453	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	534	ILE	7.3
1	A	638	PRO	5.4
1	A	519	ILE	5.3
1	A	679	THR	4.8
1	A	569	LEU	4.3

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.