wwPDB X-ray Structure Validation Summary Report (i)

Jun 16, 2024 – 05:11 PM EDT

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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (1)) were used in the production of this report:

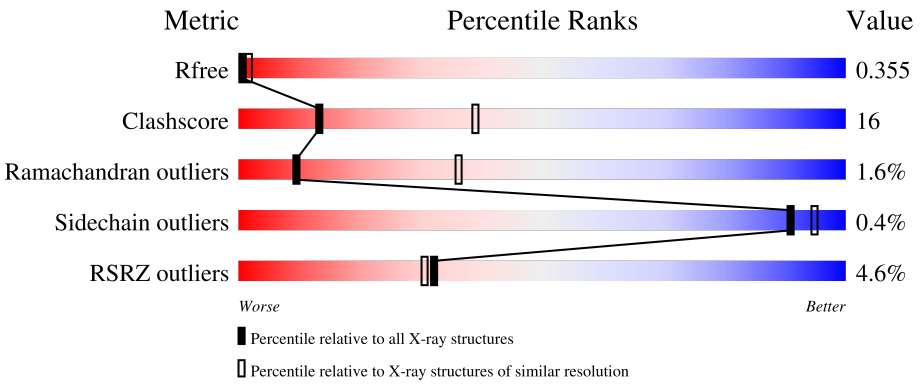
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

1 Overall quality at a glance (i)

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}	130704	1149 (3.34-3.26)
Clashscore	141614	1205 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 >=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 <=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	3% 48% 20% • 31%

2 Entry composition (i)

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
			Total	C	H	N	O	S			
1	A	843	13677	4370	6800	1164	1326	17	0	0	0

There are 6 discrepancies between the modelled and reference sequences:

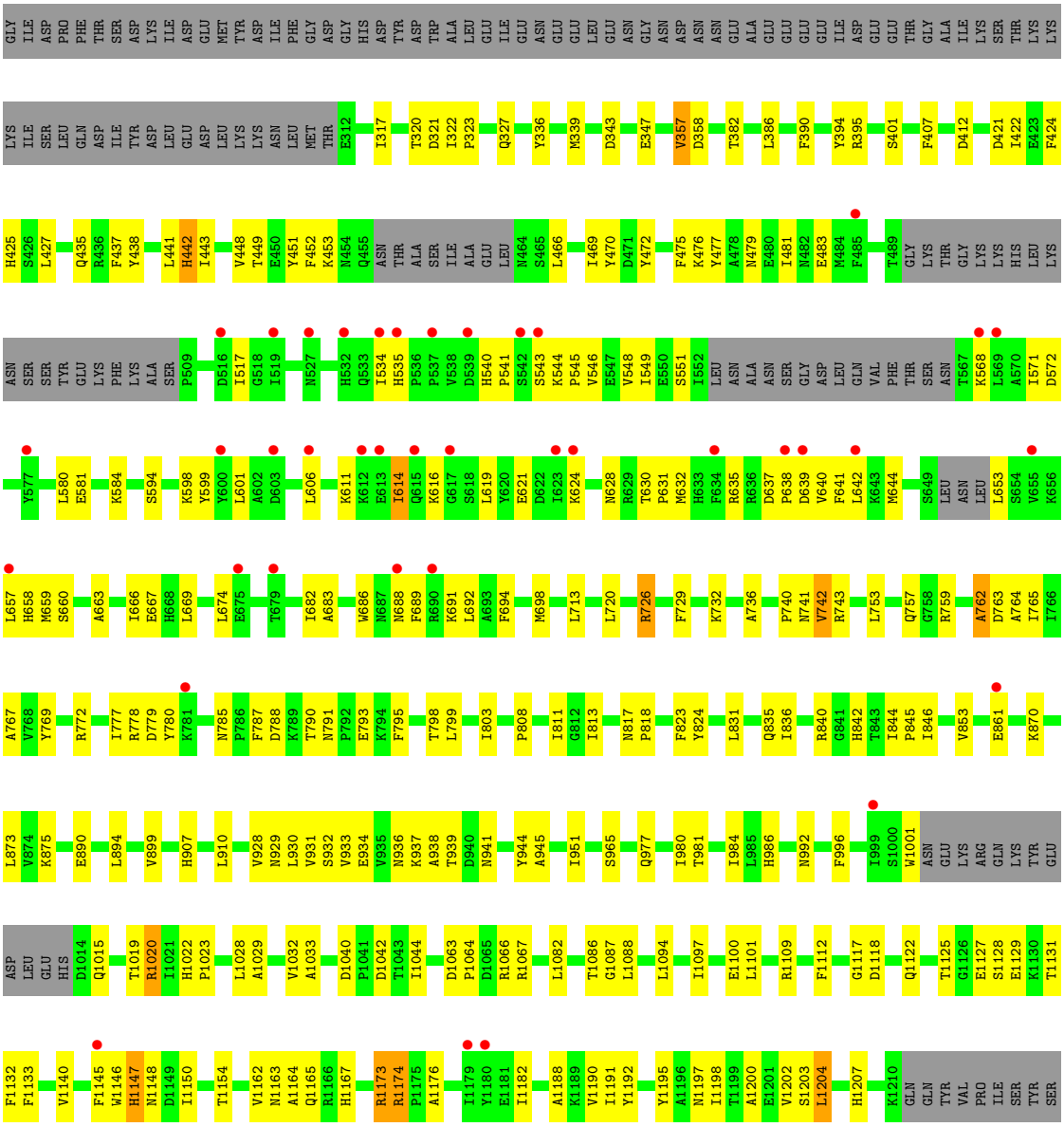
Chain	Residue	Modelled	Actual	Comment	Reference
A	233	GLY	-	EXPRESSION TAG	UNP P23615
A	234	ILE	-	EXPRESSION TAG	UNP P23615
A	235	ASP	-	EXPRESSION TAG	UNP P23615
A	236	PRO	-	EXPRESSION TAG	UNP P23615
A	237	PHE	-	EXPRESSION TAG	UNP P23615
A	238	THR	-	EXPRESSION TAG	UNP P23615

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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.

- Molecule 1: Transcription elongation factor SPT6

Chain A:



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

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

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

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

4 Data and refinement statistics (i)

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
$< I/\sigma(I) >$ [1]	3.44 (at 3.32Å)	Xtriage
Refinement program	PHENIX 1.6.4_486	Depositor
R, R_{free}	0.265 , 0.308 0.259 , 0.355	Depositor DCC
R_{free} test set	2836 reflections (10.50%)	wwPDB-VP
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]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	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.*

[1]Intensities estimated from amplitudes.

[2]Theoretical values of $< |L| >$, $< L^2 >$ 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(^{o})	Ideal(^{o})
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%)	9 36

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
1	A	762/1108 (69%)	759 (100%)	3 (0%)	91 95

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

Sometimes 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	982	HIS
1	A	1074	ASN
1	A	1147	HIS
1	A	429	ASN
1	A	425	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 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, 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%) 32 30	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	5.3
1	A	519	ILE	5.3
1	A	679	THR	4.7
1	A	569	LEU	4.4

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 (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.