



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 3, 2021 – 08:06 am BST

PDB ID : 7OGG
Title : Nse5/6 complex
Authors : Basquin, J.; Taschner, M.; Gruber, S.
Deposited on : 2021-05-06
Resolution : 3.29 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
Mogul : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.23.1
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) : 2.23.1

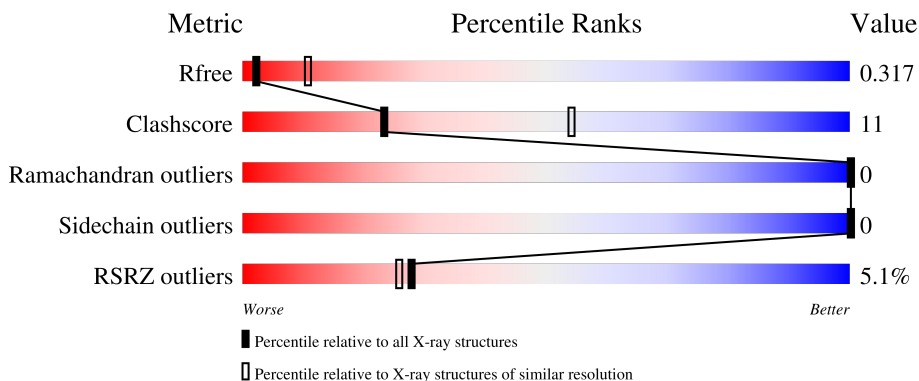
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.29 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 $\leq 5\%$. 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	R	607	 3% 60% 15% 22%
2	Q	308	 4% 52% 16% 28%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 4888 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 Non-structural maintenance of chromosome element 5,Non-structural maintenance of chromosome element 5,Non-structural maintenance of chromosome element 5.

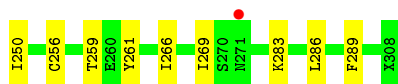
Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	R	458	3271	2098	552	608	8	5	0	0	0

- Molecule 2 is a protein called DNA repair protein KRE29,DNA repair protein KRE29,DNA repair protein KRE29.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	Q	207	1617	1061	259	288	5	4	0	1	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
Q	1	GLY	-	expression tag	UNP P40026
Q	2	PRO	-	expression tag	UNP P40026
Q	3	ALA	-	expression tag	UNP P40026
Q	4	ALA	-	expression tag	UNP P40026
Q	5	MSE	-	expression tag	UNP P40026
Q	294	GLY	-	insertion	UNP P40026



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	99.12Å 147.37Å 74.44Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.85 – 3.29 19.85 – 3.29	Depositor EDS
% Data completeness (in resolution range)	99.0 (19.85-3.29) 91.0 (19.85-3.29)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.25 (at 3.29Å)	Xtrriage
Refinement program	PHENIX 1.18.2_3874, PHENIX 1.18.2_3874	Depositor
R, R_{free}	0.297 , 0.317 0.297 , 0.317	Depositor DCC
R_{free} test set	1690 reflections (10.01%)	wwPDB-VP
Wilson B-factor (Å ²)	102.1	Xtrriage
Anisotropy	0.733	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.25 , 95.0	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.88	EDS
Total number of atoms	4888	wwPDB-VP
Average B, all atoms (Å ²)	133.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.21% 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	R	0.26	0/3070	0.50	1/4183 (0.0%)
2	Q	0.29	0/1583	0.46	0/2133
All	All	0.27	0/4653	0.49	1/6316 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	R	0	1

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
1	R	288	LEU	CA-CB-CG	5.61	128.20	115.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	R	264	ALA	Peptide

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	R	3271	0	2839	68	0
2	Q	1617	0	1467	43	0
All	All	4888	0	4306	104	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 11.

The worst 5 of 104 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:Q:192:LEU:HD12	2:Q:243:THR:HG23	1.63	0.81
1:R:44:TYR:HA	1:R:47:VAL:HG12	1.64	0.79
2:Q:145:LEU:O	2:Q:190:ARG:NH2	2.16	0.78
1:R:134:GLN:NE2	1:R:185:TYR:OH	2.19	0.75
1:R:12:VAL:HA	1:R:22:VAL:HG12	1.73	0.69

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	R	397/607 (65%)	393 (99%)	4 (1%)	0	100	100
2	Q	192/308 (62%)	191 (100%)	1 (0%)	0	100	100
All	All	589/915 (64%)	584 (99%)	5 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	R	296/504 (59%)	296 (100%)	0	100	100
2	Q	158/270 (58%)	158 (100%)	0	100	100
All	All	454/774 (59%)	454 (100%)	0	100	100

There are no protein residues with a non-rotameric sidechain to report.

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

Mol	Chain	Res	Type
1	R	183	ASN

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

The following chains have linkage breaks:

Mol	Chain	Number of breaks
1	R	2
2	Q	1

All chain breaks are listed below:

Model	Chain	Residue-1	Atom-1	Residue-2	Atom-2	Distance (Å)
1	Q	294:GLY	C	295:UNK	N	47.71
1	R	599:UNK	C	600:UNK	N	43.57
1	R	573:UNK	C	574:UNK	N	31.35

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	R	401/607 (66%)	-0.02	18 (4%) 33 32	74, 129, 200, 309	1 (0%)
2	Q	189/308 (61%)	0.17	12 (6%) 20 20	76, 120, 198, 264	0
All	All	590/915 (64%)	0.04	30 (5%) 28 26	74, 127, 203, 309	1 (0%)

The worst 5 of 30 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	Q	106	ASN	6.2
2	Q	128	ASP	5.4
1	R	268	SER	4.8
1	R	314	TYR	4.7
2	Q	271	ASN	4.2

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 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.