wwPDB X-ray Structure Validation Summary Report (i)

Oct 17, 2021 – 06:26 AM EDT

PDB ID	:	1KU7
Title	:	Crystal Structure of Thermus aquatics RNA Polymerase SigmaA Subunit Region 4 Bound to-35 Element DNA
Authors	:	Campbell, E.A.; Muzzin, O.; Chlenov, M.; Sun, J.L.; Olson, C.A.; Weinman, O.; Trester-Zedlitz, M.L.; Darst, S.A.
Deposited on	:	2002-01-21
Resolution	:	2.40 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 following versions of software and data (see references (i)) were used in the production of this report:

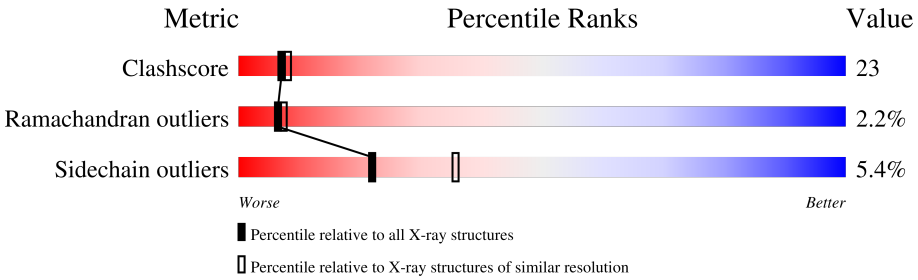
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.23.2

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 2.40 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))
Clashscore	141614	4398 (2.40-2.40)
Ramachandran outliers	138981	4318 (2.40-2.40)
Sidechain outliers	138945	4319 (2.40-2.40)

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%

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	B	11	27% 55% 9% 9%
2	C	11	36% 55% 9%
3	A	73	74% 23% .
3	D	73	60% 27% 5% 7%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 1645 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 DNA chain called 5'-D(*CP*CP*TP*TP*GP*AP*CP*AP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	11	Total	C	N	O	P	0	0	0
			222	107	43	62	10			

- Molecule 2 is a DNA chain called 5'-D(*CP*CP*TP*TP*TP*GP*TP*CP*AP*AP*G)-3'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	11	Total	C	N	O	P	0	0	0
			220	107	37	66	10			

- Molecule 3 is a protein called sigma factor sigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	73	Total	C	N	O	S	0	0	0
			611	380	117	112	2			
3	D	68	Total	C	N	O	S	0	0	0
			525	326	105	92	2			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	386	MET	LEU	engineered mutation	UNP Q9EZJ8
D	386	MET	LEU	engineered mutation	UNP Q9EZJ8

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	B	11	Total	O	0	0
			11	11		
4	C	20	Total	O	0	0
			20	20		
4	A	30	Total	O	0	0
			30	30		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	D	6	Total O 6 6	0	0

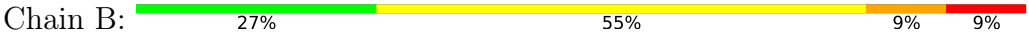
3 Residue-property plots (i)

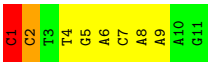
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry. Residues are color-coded 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. 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.

Note EDS was not executed.

• Molecule 1: 5'-D(*CP*CP*TP*TP*GP*AP*CP*AP*AP*AP*G)-3'

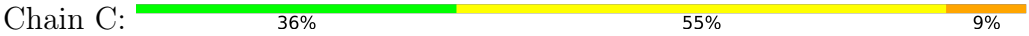
Chain B:

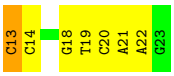




• Molecule 2: 5'-D(*CP*CP*TP*TP*TP*GP*TP*CP*AP*AP*G)-3'

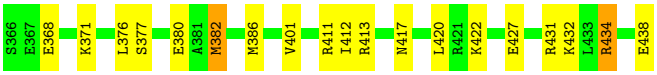
Chain C:





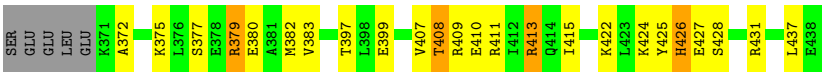
• Molecule 3: sigma factor sigA

Chain A:



• Molecule 3: sigma factor sigA

Chain D:



4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section is therefore incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	36.18Å 62.65Å 94.81Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	25.00 – 2.40	Depositor
% Data completeness (in resolution range)	97.8 (25.00-2.40)	Depositor
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.255 , 0.293	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	1645	wwPDB-VP
Average B, all atoms (Å^2)	45.0	wwPDB-VP

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	B	0.53	0/249	0.99	1/382 (0.3%)
2	C	0.47	0/245	0.91	0/376
3	A	0.55	1/616 (0.2%)	0.62	0/814
3	D	0.54	1/528 (0.2%)	0.56	0/700
All	All	0.53	2/1638 (0.1%)	0.73	1/2272 (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	B	0	2
2	C	0	1
All	All	0	3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	382	MET	CG-SD	6.38	1.97	1.81
3	A	382	MET	CG-SD	5.61	1.95	1.81

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	1	DC	N1-C1'-C2'	-6.19	100.84	112.60

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	1	DC	Sidechain

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Group
1	B	2	DC	Sidechain
2	C	13	DC	Sidechain

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	B	222	0	125	12	0
2	C	220	0	127	7	0
3	A	611	0	635	22	0
3	D	525	0	523	28	0
4	A	30	0	0	6	0
4	B	11	0	0	0	0
4	C	20	0	0	0	0
4	D	6	0	0	0	0
All	All	1645	0	1410	68	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 23.

The worst 5 of 68 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)
3:A:431:ARG:HG3	3:A:434:ARG:HG3	1.32	1.10
3:D:410:GLU:O	3:D:413:ARG:HD3	1.72	0.88
3:A:371:LYS:HD2	3:A:432:LYS:HZ1	1.40	0.85
1:B:1:DC:H2"	1:B:2:DC:C5'	2.09	0.82
3:D:408:THR:HG23	3:D:410:GLU:H	1.44	0.81

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
3	A	71/73 (97%)	68 (96%)	3 (4%)	0	100 100
3	D	66/73 (90%)	59 (89%)	4 (6%)	3 (4%)	2 2
All	All	137/146 (94%)	127 (93%)	7 (5%)	3 (2%)	6 7

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	D	426	HIS
3	D	437	LEU
3	D	372	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.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
3	A	64/65 (98%)	61 (95%)	3 (5%)	26 42
3	D	48/65 (74%)	45 (94%)	3 (6%)	18 28
All	All	112/130 (86%)	106 (95%)	6 (5%)	22 36

5 of 6 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	379	ARG
3	D	408	THR
3	D	413	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	A	420	LEU
3	A	413	ARG

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

Mol	Chain	Res	Type
3	D	417	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 (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section is therefore empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section is therefore empty.

6.3 Carbohydrates

EDS was not executed - this section is therefore empty.

6.4 Ligands

EDS was not executed - this section is therefore empty.

6.5 Other polymers

EDS was not executed - this section is therefore empty.