



wwPDB X-ray Structure Validation Summary Report i

Mar 1, 2014 – 04:35 AM GMT

PDB ID : 3K5O
Title : Crystal structure of E.coli Pol II
Authors : Yang, W.; Wang, F.
Deposited on : 2009-10-07
Resolution : 2.20 Å(reported)

This is a wwPDB 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/ValidationPDFNotes.html>

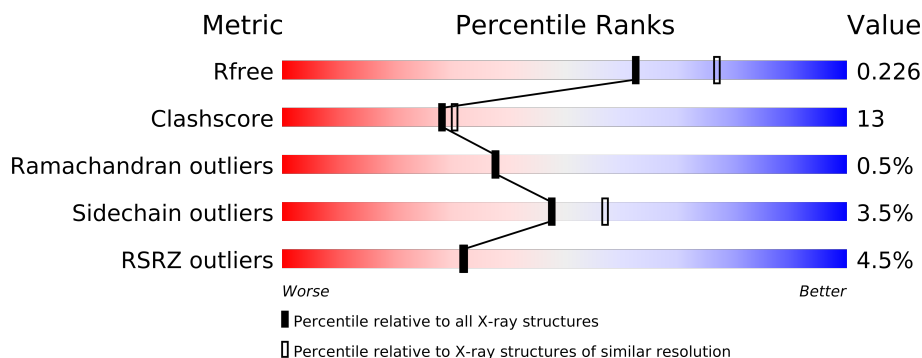
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

1 Overall quality at a glance

The reported resolution of this entry is 2.20 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}	66092	2938 (2.20-2.20)
Clashscore	79885	3751 (2.20-2.20)
Ramachandran outliers	78287	3681 (2.20-2.20)
Sidechain outliers	78261	3682 (2.20-2.20)
RSRZ outliers	66119	2939 (2.20-2.20)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	786	
1	B	786	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 12674 atoms, of which 0 are hydrogen and 0 are deuterium.

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 DNA polymerase II.

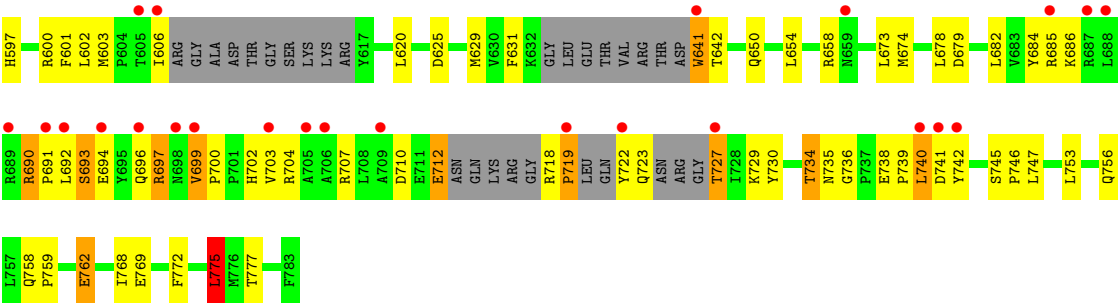
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	758	Total	C	N	O	S	0	0	0
			6137	3918	1087	1109	23			
1	B	744	Total	C	N	O	S	0	0	0
			6015	3847	1056	1089	23			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	EXPRESSION TAG	UNP P21189
A	-1	PRO	-	EXPRESSION TAG	UNP P21189
A	0	HIS	-	EXPRESSION TAG	UNP P21189
A	335	ASN	ASP	ENGINEERED	UNP P21189
B	-2	GLY	-	EXPRESSION TAG	UNP P21189
B	-1	PRO	-	EXPRESSION TAG	UNP P21189
B	0	HIS	-	EXPRESSION TAG	UNP P21189
B	335	ASN	ASP	ENGINEERED	UNP P21189

- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	250	Total	O	0	0
			250	250		
2	B	272	Total	O	0	0
			272	272		



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	92.85Å 116.82Å 163.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	50.00 – 2.20 47.52 – 2.20	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.20) 91.1 (47.52-2.20)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.78 (at 2.20Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, R_{free}	0.218 , 0.223 0.221 , 0.226	Depositor DCC
R_{free} test set	2029 reflections (2.45%)	DCC
Wilson B-factor (Å ²)	31.2	Xtriage
Anisotropy	0.163	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 32.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	2 of 89919 reflections (0.002%)	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12674	wwPDB-VP
Average B, all atoms (Å ²)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 73.70 % of the origin peak, indicating pseudo translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 1.7531e-06. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

¹Intensities estimated from amplitudes.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

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	0.40	0/6299	0.69	0/8546
1	B	0.40	0/6175	0.71	4/8376 (0.0%)
All	All	0.40	0/12474	0.70	4/16922 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	734	THR	N-CA-C	-6.39	93.75	111.00
1	B	775	LEU	CA-CB-CG	5.90	128.87	115.30
1	B	408	ARG	N-CA-C	-5.28	96.75	111.00
1	B	736	GLY	N-CA-C	5.14	125.95	113.10

There are no chirality outliers.

There are no planarity outliers.

5.2 Close contacts ⓘ

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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	6137	0	5992	114	0
1	B	6015	0	5849	191	0
2	A	250	0	0	3	0
2	B	272	0	0	8	0
All	All	12674	0	11841	302	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

The worst 5 of 302 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:522:ARG:NH1	1:A:525:GLN:HE22	1.49	1.10
1:A:522:ARG:HH11	1:A:525:GLN:NE2	1.56	1.02
1:B:674:MET:HE3	1:B:753:LEU:HD22	1.40	1.01
1:A:311:TRP:CZ2	1:B:696:GLN:HB2	2.00	0.95
1:B:258:HIS:HD2	1:B:266:PHE:HD2	1.10	0.95

There are no symmetry-related clashes.

5.3 Torsion angles

5.3.1 Protein backbone ⓘ

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	748/786 (95%)	733 (98%)	12 (2%)	3 (0%)	43	45
1	B	730/786 (93%)	709 (97%)	17 (2%)	4 (0%)	38	38
All	All	1478/1572 (94%)	1442 (98%)	29 (2%)	7 (0%)	38	38

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	693	SER
1	A	698	ASN
1	B	122	SER
1	B	699	VAL
1	A	122	SER

5.3.2 Protein sidechains ⓘ

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	642/672 (96%)	625 (97%)	17 (3%)	59	70
1	B	629/672 (94%)	602 (96%)	27 (4%)	40	47
All	All	1271/1344 (95%)	1227 (96%)	44 (4%)	48	57

5 of 44 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	175	GLN
1	B	303	GLU
1	B	747	LEU
1	B	268	GLN
1	B	289	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	557	HIS
1	A	735	ASN
1	B	597	HIS
1	A	571	GLN
1	A	584	GLN

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data ⓘ

6.1 Protein, DNA and RNA chains ⓘ

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	758/786 (96%)	0.03	34 (4%) 32 32	18, 35, 65, 85	0
1	B	744/786 (94%)	0.07	33 (4%) 33 33	17, 36, 67, 81	0
All	All	1502/1572 (95%)	0.05	67 (4%) 32 32	17, 36, 66, 85	0

The worst 5 of 67 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	709	ALA	5.8
1	A	398	ALA	5.4
1	A	636	THR	5.3
1	A	697	ARG	4.6
1	A	698	ASN	4.6

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4 Ligands ⓘ

There are no ligands in this entry.

6.5 Other polymers ⓘ

There are no such residues in this entry.