



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:09 PM GMT

PDB ID : 1NOZ
Title : T4 DNA POLYMERASE FRAGMENT (RESIDUES 1-388) AT 110K
Authors : Wang, J.; Yu, P.; Lin, T.C.; Konigsberg, W.H.; Steitz, T.A.
Deposited on : 1996-02-16
Resolution : 2.20 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
Mogul : 1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026688
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
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) : trunk26865

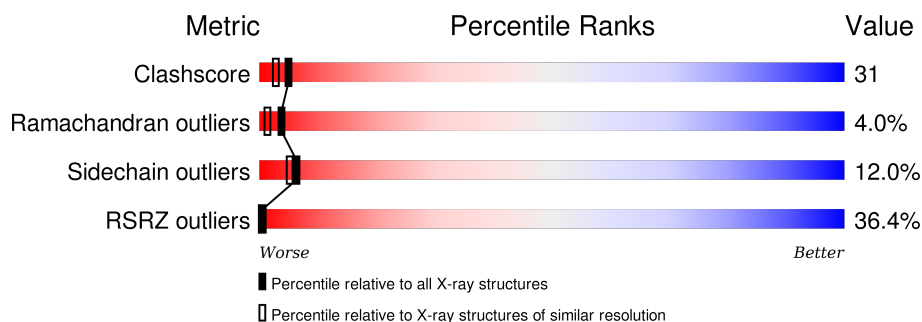
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 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))
Clashscore	102246	4477 (2.20-2.20)
Ramachandran outliers	100387	4404 (2.20-2.20)
Sidechain outliers	100360	4405 (2.20-2.20)
RSRZ outliers	91569	3781 (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. 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	388	
1	B	388	

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 5725 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 DNA POLYMERASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	346	Total	C	N	O	S	0	0	0
			2840	1822	466	533	19			
1	B	346	Total	C	N	O	S	0	0	0
			2840	1822	466	533	19			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ASP	LYS	CONFLICT	UNP P04415
A	250	LEU	ILE	CONFLICT	UNP P04415
B	2	ASP	LYS	CONFLICT	UNP P04415
B	250	LEU	ILE	CONFLICT	UNP P04415

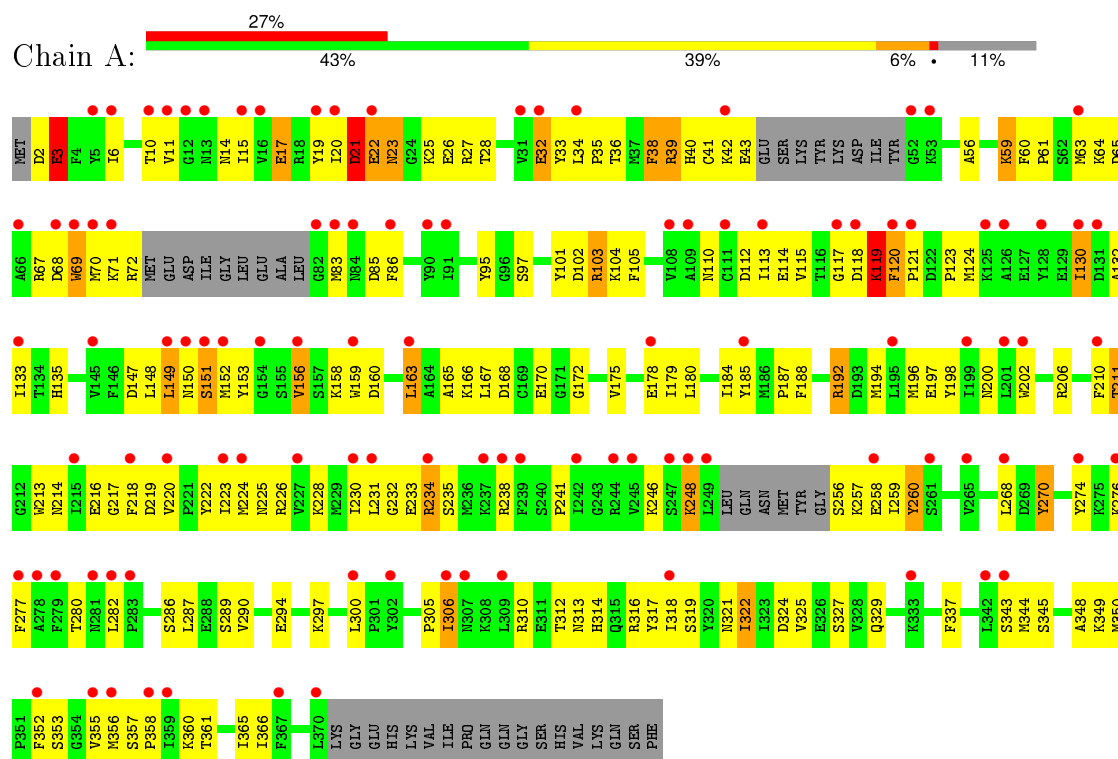
- Molecule 2 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	34	Total	O	0	0
			34	34		
2	B	11	Total	O	0	0
			11	11		

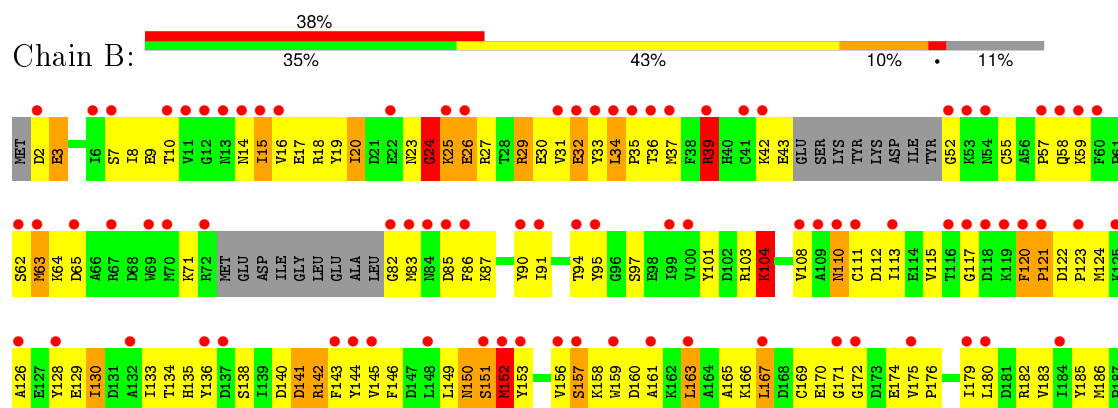
3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors 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: DNA POLYMERASE



• Molecule 1: DNA POLYMERASE



GLN	F188	•
ASN	D189	•
MET	N190	•
TYR	E191	
GLY	R192	
S256	D193	
K257	M194	•
E258	L195	•
I259	M196	•
Y260	E197	
S261	Y198	•
I262	I199	
D263	N200	
G264	L201	•
V265	W202	•
S266	E203	•
Q269	Q204	•
A330	K205	•
K333	R206	•
I338	P207	
D339	A208	
L340	I209	•
V341	F210	
L342	T211	
S343	G212	
M344	W213	
Y346	M214	
F352	I215	
V355	F218	•
M356	D219	
S357	Y222	•
K360	I223	
T361	R226	
W362	V227	
I366	K228	
F367	M229	•
N368	I230	•
S369	G232	
L370	L231	
LYS	E233	
GLY	R234	•
GLU	S235	
HIS	M236	
LYS	K237	
VAL	R238	•
ILE	F239	•
PRO	S240	
GLN	P241	
GLN	I242	•
GLY	G243	•
SER	N244	
HIS	V245	•
VAL	K246	•
LYS	S247	•
	K248	•
	L249	•
	LEU	•
GLN	GLN	
SER	ASN	
PHE	MET	
	TYR	
	GLY	
	S256	
	K257	•
	E258	
	I259	
	Y260	•
	S261	
	I262	•
	D263	
	G264	
	V265	•
	S266	•
	Q269	•
	A330	
	K333	
	I338	•
	D339	
	L340	
	V341	•
	L342	
	S343	
	M344	
	Y346	
	F352	•
	V355	
	M356	
	S357	
	K360	•
	T361	
	W362	
	I366	
	F367	•
	N368	•
	S369	
	L370	•
	LYS	
	GLY	
	GLU	
	HIS	
	LYS	
	VAL	
	ILE	
	PRO	
	GLN	
	GLN	
	GLY	
	SER	
	HIS	
	VAL	
	LYS	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, α , β , γ	113.62Å 109.31Å 68.58Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	6.00 – 2.20 58.09 – 2.72	Depositor EDS
% Data completeness (in resolution range)	(Not available) (6.00-2.20) 82.2 (58.09-2.72)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.45 (at 2.73Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.222 , (Not available) (Not available) , (Not available)	Depositor DCC
R_{free} test set	NotAvailable	DCC
Wilson B-factor (Å ²)	44.9	Xtriage
Anisotropy	0.487	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 92.8	EDS
Estimated twinning fraction	0.017 for k,h,-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 20521 reflections	Xtriage
F_o, F_c correlation	0.83	EDS
Total number of atoms	5725	wwPDB-VP
Average B, all atoms (Å ²)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.11% 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, 0.375 respectively for untwinned datasets, and 0.333, 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	0.65	0/2904	0.85	2/3911 (0.1%)
1	B	0.59	0/2904	0.89	3/3911 (0.1%)
All	All	0.62	0/5808	0.87	5/7822 (0.1%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	24	GLY	N-CA-C	6.41	129.13	113.10
1	A	32	GLU	N-CA-C	-6.15	94.39	111.00
1	B	280	THR	N-CA-C	6.13	127.56	111.00
1	B	29	ARG	NE-CZ-NH1	5.98	123.29	120.30
1	A	38	PHE	N-CA-C	5.73	126.46	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	2840	0	2783	151	0
1	B	2840	0	2783	210	0
2	A	34	0	0	2	0
2	B	11	0	0	1	0
All	All	5725	0	5566	350	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 31.

The worst 5 of 350 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:B:169:CYS:SG	1:B:170:GLU:HG3	2.05	0.97
1:A:114:GLU:HB2	1:A:132:ALA:HB3	1.53	0.88
1:B:179:ILE:HG22	1:B:322:ILE:HD11	1.62	0.82
1:A:67:ARG:O	1:A:71:LYS:HG2	1.81	0.81
1:B:10:THR:HG23	1:B:14:ASN:O	1.81	0.79

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	338/388 (87%)	299 (88%)	27 (8%)	12 (4%)	4	2
1	B	338/388 (87%)	294 (87%)	29 (9%)	15 (4%)	3	1
All	All	676/776 (87%)	593 (88%)	56 (8%)	27 (4%)	4	1

5 of 27 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	120	PHE
1	A	150	ASN
1	A	257	LYS
1	B	3	GLU
1	B	141	ASP

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	312/350 (89%)	280 (90%)	32 (10%)	9	8
1	B	312/350 (89%)	269 (86%)	43 (14%)	4	3
All	All	624/700 (89%)	549 (88%)	75 (12%)	6	5

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	20	ILE
1	B	58	GLN
1	B	313	ASN
1	B	23	ASN
1	B	34	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 16 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	313	ASN
1	A	321	ASN
1	B	204	GLN
1	A	225	ASN
1	B	292	GLN

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 [i](#)

There are no carbohydrates 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	346/388 (89%)	1.73	105 (30%) 1 0	20, 39, 61, 72	0
1	B	346/388 (89%)	2.14	147 (42%) 0 0	25, 54, 87, 101	0
All	All	692/776 (89%)	1.93	252 (36%) 0 0	20, 46, 78, 101	0

The worst 5 of 252 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	249	LEU	11.7
1	B	82	GLY	8.4
1	B	298	GLY	7.8
1	B	120	PHE	7.4
1	B	245	VAL	7.4

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 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.