



wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 09:07 PM GMT

PDB ID : 1NKC
Title : A BACILLUS DNA POLYMERASE I PRODUCT COMPLEX BOUND TO A GUANINE-THYMINE MISMATCH AFTER FIVE ROUNDS OF PRIMER EXTENSION, FOLLOWING INCORPORATION OF DCTP, DGTP, DTTP, AND DATP.
Authors : Johnson, S.J.; Beese, L.S.
Deposited on : 2003-01-02
Resolution : 1.80 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

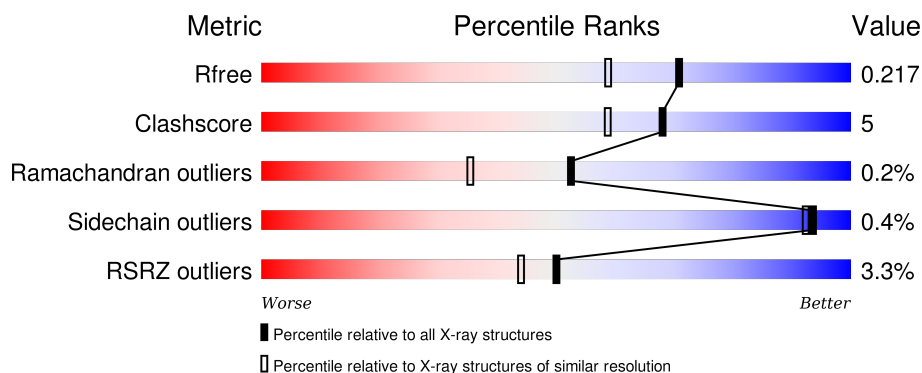
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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}	91344	4533 (1.80-1.80)
Clashscore	102246	5383 (1.80-1.80)
Ramachandran outliers	100387	5320 (1.80-1.80)
Sidechain outliers	100360	5319 (1.80-1.80)
RSRZ outliers	91569	4547 (1.80-1.80)

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	B	15	 67% 33% 7%
2	C	15	 87% 13% 3%
3	A	580	 88% 12% 3%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	SUC	A	901	-	-	-	X

2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 5912 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 DNA PRIMER STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	B	15	Total	C	N	O	P	0	0	0
			305	145	59	87	14			

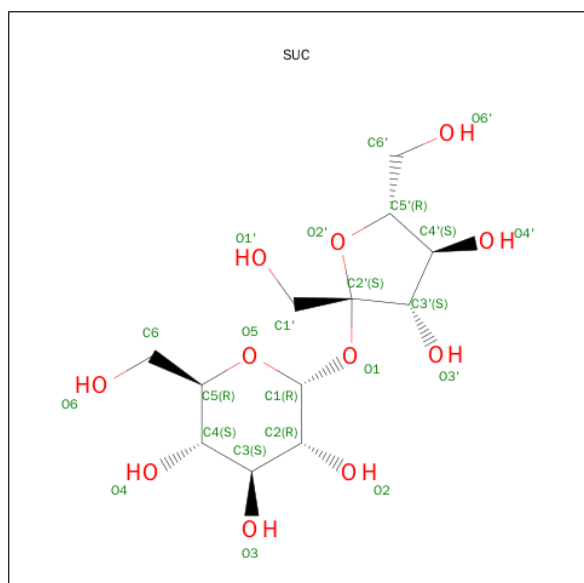
- Molecule 2 is a DNA chain called DNA TEMPLATE STRAND.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	15	Total	C	N	O	P	0	0	0
			305	146	55	90	14			

- Molecule 3 is a protein called DNA POLYMERASE I.

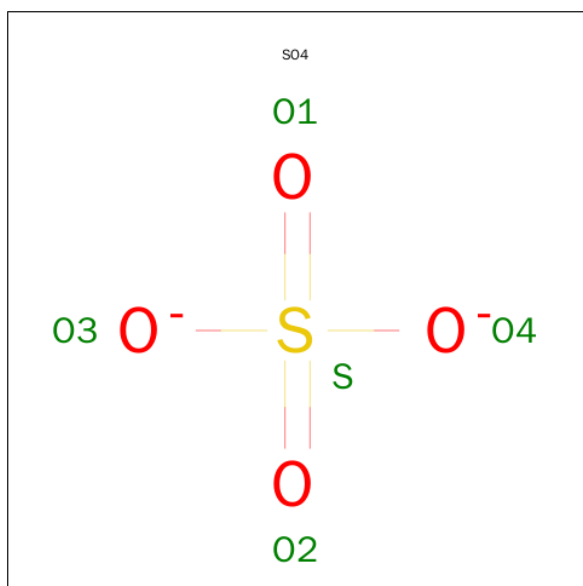
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	580	Total	C	N	O	S	0	0	0
			4650	2956	807	870	17			

- Molecule 4 is SUGAR (SUCROSE) (three-letter code: SUC) (formula: C₁₂H₂₂O₁₁).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			23	12	11		

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		
5	A	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Mg	0	0
			1	1		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	532	Total	O	0	0
			532	532		
7	B	32	Total	O	0	0
			32	32		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	C	49	Total	O	0	0
			49	49		

- Molecule 1: DNA PRIMER STRAND

A diagram of a protein structure with four residues highlighted: G27 (green), T35 (yellow), G36 (yellow), and C41 (red). The residues are connected by lines, indicating a sequence or structural relationship.

K760	Y587	A297
R779	Y588	L303
R784	E589	V319
E788	G590	G334
T794	L591	P348
P795	L592	G406
A813	K593	V407
E817	N607	D408
A822	L610	M416
V828	S617	E426
E831	P627	A433
A836	L628	K434
L844	P628	A453
V848	L629	L457
P849	P640	E458
R859	H664	R459
K876	R677	P460
	L678	R466
	L679	L485
	F690	E489
	G691	Q510
	V692	V514
	S693	S530
	D696	P531
	N700	K532
	H701	Q533
	Q704	L534
	T716	G535
	S717	V536
	D718	K549
	Y719	R550
	Q723	K551
	N726	T552
	T727	G553
	S728	A565
	R729	P566
	K730	I574
	E731	Q579
	E734	O584
	E737	
	R738	
	E741	

4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	87.80 Å 93.49 Å 105.47 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	26.76 – 1.80 26.76 – 1.80	Depositor EDS
% Data completeness (in resolution range)	75.1 (26.76-1.80) 76.5 (26.76-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.04	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.50 (at 1.80 Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, R_{free}	0.186 , 0.218 0.186 , 0.217	Depositor DCC
R_{free} test set	3090 reflections (5.08%)	DCC
Wilson B-factor (Å ²)	17.5	Xtriage
Anisotropy	0.056	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.40 , 53.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 61954 reflections	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5912	wwPDB-VP
Average B, all atoms (Å ²)	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.11% 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](#)

Bond lengths and bond angles in the following residue types are not validated in this section: MG, SUC, SO4

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	0.30	0/342	0.77	1/526 (0.2%)
2	C	0.35	0/341	0.70	0/525
3	A	0.30	0/4734	0.56	0/6398
All	All	0.30	0/5417	0.59	1/7449 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	12	DC	C1'-O4'-C4'	-5.64	104.46	110.10

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	B	305	0	169	5	0
2	C	305	0	171	2	0
3	A	4650	0	4698	43	0
4	A	23	0	22	1	0
5	A	15	0	0	0	0
6	A	1	0	0	0	0
7	A	532	0	0	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	B	32	0	0	0	0
7	C	49	0	0	0	0
All	All	5912	0	5060	49	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 5.

The worst 5 of 49 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:738:ARG:HH11	3:A:738:ARG:HA	1.56	0.71
3:A:408:ASP:HB2	4:A:901:SUC:H1'1	1.75	0.69
1:B:6:DT:H2''	1:B:7:DC:C5'	2.25	0.66
3:A:737:GLU:O	3:A:741:GLU:HG3	1.96	0.66
1:B:6:DT:H2''	1:B:7:DC:H5'	1.78	0.65

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	578/580 (100%)	564 (98%)	13 (2%)	1 (0%)	52 35

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	A	628	ILE

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	495/496 (100%)	493 (100%)	2 (0%)	93	92

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	303	LEU
3	A	844	LEU

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

Mol	Chain	Res	Type
3	A	782	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 carbohydrates in this entry.

5.6 Ligand geometry [i](#)

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link

column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# $ Z > 2$	Counts	RMSZ	# $ Z > 2$
4	SUC	A	901	-	24,24,24	1.55	3 (12%)	36,36,36	0.93	2 (5%)
5	SO4	A	910	-	4,4,4	0.22	0	6,6,6	0.07	0
5	SO4	A	911	-	4,4,4	0.21	0	6,6,6	0.09	0
5	SO4	A	912	-	4,4,4	0.18	0	6,6,6	0.11	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the chemical component dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	SUC	A	901	-	-	0/12/51/51	0/2/2/2
5	SO4	A	910	-	-	0/0/0/0	0/0/0/0
5	SO4	A	911	-	-	0/0/0/0	0/0/0/0
5	SO4	A	912	-	-	0/0/0/0	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	901	SUC	O5-C1	2.43	1.48	1.41
4	A	901	SUC	C4-C5	2.44	1.58	1.53
4	A	901	SUC	C3-C2	5.62	1.67	1.52

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	901	SUC	C1-C2-C3	-2.09	105.85	109.97
4	A	901	SUC	C1-O5-C5	2.10	117.82	113.75

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	901	SUC	1	0

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	B	15/15 (100%)	0.11	0 100 100	13, 32, 40, 43	0
2	C	15/15 (100%)	0.05	1 (6%) 21 17	12, 23, 45, 51	0
3	A	580/580 (100%)	0.06	19 (3%) 50 44	7, 16, 34, 52	0
All	All	610/610 (100%)	0.06	20 (3%) 50 44	7, 16, 36, 52	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	A	719	TYR	7.7
3	A	552	THR	4.9
3	A	551	LYS	4.4
3	A	729	ARG	3.7
3	A	297	ALA	3.5

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](#)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors

of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
4	SUC	A	901	23/23	0.80	0.17	2.47	22,28,30,31	0
5	SO4	A	912	5/5	0.96	0.12	0.16	38,38,39,40	0
5	SO4	A	911	5/5	0.81	0.19	-	79,80,80,80	0
6	MG	A	920	1/1	0.99	0.05	-	25,25,25,25	0
5	SO4	A	910	5/5	0.94	0.17	-	58,58,58,58	0

6.5 Other polymers [i](#)

There are no such residues in this entry.