WORLDWIDE wwPDB PROTEIN DATA BANK

Full wwPDB X-ray Structure Validation Report ⓘ

Feb 13, 2017 – 02:43 am GMT

PDB ID : 5IIM
Title : Crystal structure of the pre-catalytic ternary extension complex of DNA polymerase lambda with an 8-oxo-dG:dA base-pair
Authors : Burak, M.J.; Guja, K.E.; Garcia-Diaz, M.
Deposited on : 2016-03-01
Resolution : 1.94 Å(reported)

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

MolProbity : 4.02b-467
Mogul : 1.7.2 (RC1), CSD as538be (2017)
Xtriage (Phenix) : 1.9-1692
EDS : trunk28620
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
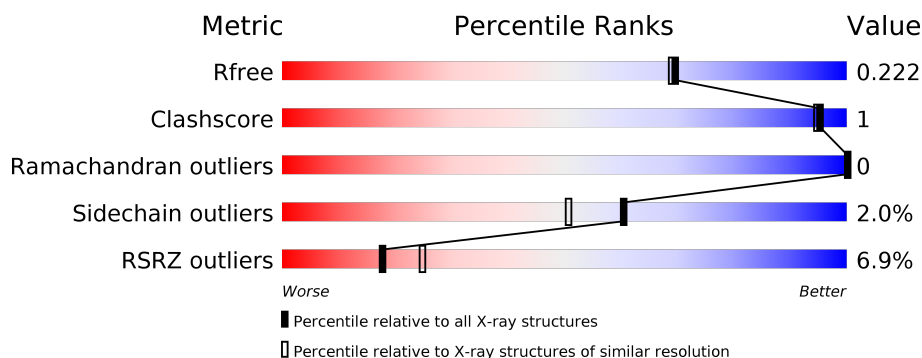Validation Pipeline (wwPDB-VP) : recalc28949

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 1.94 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	Percentile Ranks	Value
Rfree		0.222
Clashscore		1
Ramachandran outliers		0
Sidechain outliers		2.0%
RSRZ outliers		6.9%

Worse · Better

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	100719	3233 (1.96-1.92)
Clashscore	112137	3430 (1.96-1.92)
Ramachandran outliers	110173	3395 (1.96-1.92)
Sidechain outliers	110143	3395 (1.96-1.92)
RSRZ outliers	101464	3250 (1.96-1.92)

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 <=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	334	7% / 93% · · ·
2	D	4	75% / 25%
3	P	6	100%
4	T	11	91% / 9%

2 Entry composition ⓘ

There are 7 unique types of molecules in this entry. The entry contains 3301 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 lambda.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	326	Total	C	N	O	S	0	2	0
			2493	1567	451	463	12			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	D	4	Total	C	N	O	P	0	0	0
			83	38	16	25	4			

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

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	6	Total	C	N	O	P	0	0	0
			121	59	25	32	5			

- Molecule 4 is a DNA chain called DNA (5'-D(*CP*GP*GP*CP*AP*(8OG)P*TP*AP*CP*TP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	T	11	Total	C	N	O	P	0	0	0
			225	107	43	65	10			

- Molecule 5 is 2'-DEOXYURIDINE 5'-ALPHA,BETA-IMIDO-TRIPHOSPHATE (three-letter code: DUP) (formula: $C_9H_{16}N_3O_{13}P_3$).
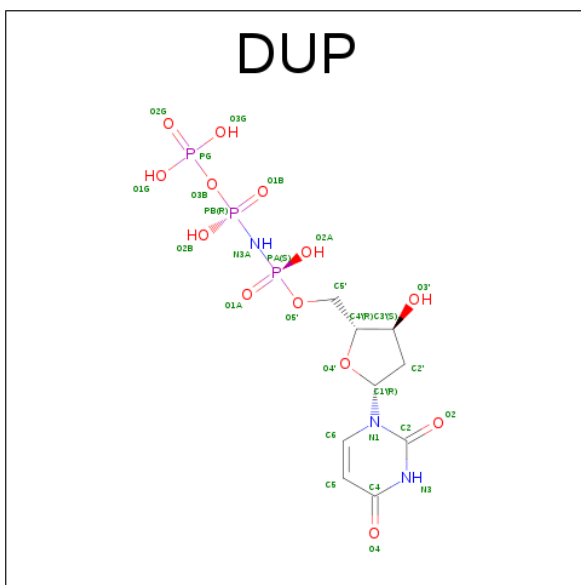
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
5	A	1	Total	C	N	O	P	0	0
			28	9	3	13	3		

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	2	Total	Na	0	0
			2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	245	Total	O	0	0
			245	245		
7	D	11	Total	O	0	0
			11	11		
7	P	32	Total	O	0	0
			32	32		
7	T	61	Total	O	0	0
			61	61		

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 the various outlier classes 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 lambda
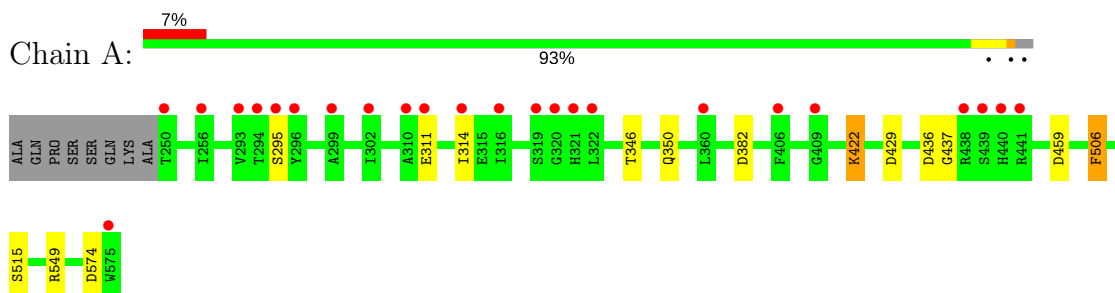
Chain A:

• Molecule 2: DNA (5'-D(P*GP*CP*CP*G)-3')
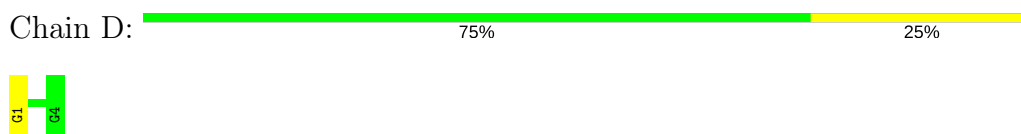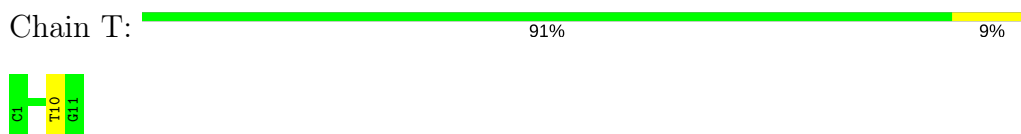
Chain D:

• Molecule 3: DNA (5'-D(*CP*AP*GP*TP*AP*A)-3')

Chain P:

There are no outlier residues recorded for this chain.

• Molecule 4: DNA (5'-D(*CP*GP*GP*CP*AP*(8OG)P*TP*AP*CP*TP*G)-3')

Chain T:

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	56.27Å 63.70Å 140.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	43.95 – 1.94 43.95 – 1.94	Depositor EDS
% Data completeness (in resolution range)	94.4 (43.95-1.94) 94.6 (43.95-1.94)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.87 (at 1.94Å)	Xtriage
Refinement program	PHENIX (1.10.1_2155: ???)	Depositor
R, R_{free}	0.183 , 0.216 0.190 , 0.222	Depositor DCC
R_{free} test set	3596 reflections (9.97%)	DCC
Wilson B-factor (Å^2)	26.3	Xtriage
Anisotropy	0.298	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.38 , 53.7	EDS
L-test for twinning[2]	$<|L|> = 0.48$, $<L^2> = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	3301	wwPDB-VP
Average B, all atoms (Å^2)	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.63% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $<|L|>$, $<L^2>$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality ⓘ

5.1 Standard geometry ⓘ

Bond lengths and bond angles in the following residue types are not validated in this section: DUP, NA, 8OG

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	0.34	0/2550	0.54	0/3452
2	D	1.27	1/92 (1.1%)	0.83	0/138
3	P	0.90	0/136	0.98	0/208
4	T	0.91	0/225	1.08	1/343 (0.3%)
All	All	0.50	1/3003 (0.0%)	0.64	1/4141 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1	DG	OP3-P	-10.53	1.48	1.61

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	T	10	DT	O4'-C1'-N1	-5.23	104.34	108.00

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-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 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	2493	0	2429	8	0
2	D	83	0	45	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	P	121	0	69	0	0
4	T	225	0	125	0	0
5	A	28	0	12	0	0
6	A	2	0	0	0	0
7	A	245	0	0	4	1
7	D	11	0	0	0	0
7	P	32	0	0	0	0
7	T	61	0	0	0	1
All	All	3301	0	2680	8	1

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 1.

All (8) 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:A:382:ASP:OD2	7:A:701:HOH:O	2.06	0.73
1:A:459:ASP:OD2	7:A:702:HOH:O	2.12	0.68
1:A:436:ASP:OD1	1:A:437:GLY:N	2.33	0.61
1:A:346[B]:THR:HG23	7:A:868:HOH:O	2.11	0.50
1:A:350:GLN:NE2	7:A:704:HOH:O	2.32	0.48
1:A:311:GLU:HA	1:A:314:ILE:HG22	1.97	0.46
1:A:422:LYS:NZ	1:A:574:ASP:OD2	2.51	0.43
1:A:429:ASP:HB3	1:A:506:PHE:CZ	2.55	0.42

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
7:A:741:HOH:O	7:T:121:HOH:O[4_555]	2.02	0.18

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	326/334 (98%)	321 (98%)	5 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	255/280 (91%)	250 (98%)	5 (2%)	60	51

All (5) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	295	SER
1	A	422	LYS
1	A	506	PHE
1	A	515	SER
1	A	549	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

1 non-standard protein/DNA/RNA residue is modelled in this entry.

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	8OG	T	6	3,4	17,25,26	5.77	12 (70%)	22,37,40	1.87	5 (22%)

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	8OG	T	6	3,4	-	0/3/21/22	0/3/3/3

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	T	6	8OG	C3'-C4'	-7.14	1.33	1.53
4	T	6	8OG	O4'-C1'	-4.85	1.31	1.42
4	T	6	8OG	O5'-C5'	-3.60	1.39	1.44
4	T	6	8OG	C2-N3	2.32	1.47	1.35
4	T	6	8OG	O3'-C3'	2.71	1.49	1.43
4	T	6	8OG	C2-N2	4.92	1.44	1.34
4	T	6	8OG	C6-C5	5.00	1.50	1.41
4	T	6	8OG	C4-N3	6.03	1.45	1.35
4	T	6	8OG	C6-N1	6.36	1.44	1.33
4	T	6	8OG	C2-N1	7.06	1.48	1.35
4	T	6	8OG	O4'-C4'	7.25	1.61	1.45
4	T	6	8OG	C8-N7	15.21	1.50	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	T	6	8OG	N3-C2-N1	-4.48	120.92	127.46
4	T	6	8OG	C5-C6-N1	-3.50	118.50	123.48
4	T	6	8OG	C4'-O4'-C1'	-3.45	100.98	109.42
4	T	6	8OG	C6-N1-C2	2.58	119.77	116.06
4	T	6	8OG	C2-N3-C4	2.97	118.62	115.16

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 3 ligands modelled in this entry, 2 are monoatomic - leaving 1 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	#$|Z| > 2$
5	DUP	A	601	6	29,29,29	3.84	14 (48%)	37,45,45	2.24	8 (21%)

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
5	DUP	A	601	6	-	0/18/34/34	0/2/2/2

All (14) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	DUP	C6-N1	-9.46	1.35	1.47
5	A	601	DUP	C6-C5	-8.28	1.37	1.52
5	A	601	DUP	O4'-C4'	-6.03	1.31	1.45
5	A	601	DUP	C2'-C1'	-5.59	1.37	1.52
5	A	601	DUP	C5-C4	-5.02	1.38	1.50
5	A	601	DUP	O3'-C3'	-2.50	1.37	1.43
5	A	601	DUP	PA-O2A	-2.25	1.50	1.56
5	A	601	DUP	PA-O5'	2.68	1.64	1.57
5	A	601	DUP	PB-O3B	2.77	1.62	1.59

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	601	DUP	PA-N3A	3.73	1.73	1.63
5	A	601	DUP	PA-O1A	3.73	1.50	1.46
5	A	601	DUP	PB-N3A	3.81	1.73	1.63
5	A	601	DUP	PB-O1B	4.50	1.51	1.46
5	A	601	DUP	O4'-C1'	8.29	1.61	1.42

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	601	DUP	C4-N3-C2	-5.69	120.94	125.81
5	A	601	DUP	C2'-C1'-N1	-3.88	110.87	115.61
5	A	601	DUP	O2-C2-N1	-2.42	120.09	123.12
5	A	601	DUP	PG-O3B-PB	-2.21	124.57	132.38
5	A	601	DUP	O1A-PA-N3A	2.40	115.37	111.79
5	A	601	DUP	N3-C2-N1	2.77	119.49	116.73
5	A	601	DUP	C5-C4-N3	2.84	119.54	116.72
5	A	601	DUP	C5-C6-N1	9.20	120.27	110.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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(Å²)	Q<0.9
1	A	326/334 (97%)	0.58	24 (7%)	15	22	13, 30, 58, 79	0
2	D	4/4 (100%)	-0.07	0	100	100	29, 35, 35, 36	0
3	P	6/6 (100%)	-0.09	0	100	100	15, 17, 19, 21	0
4	T	10/11 (90%)	-0.13	0	100	100	17, 20, 27, 32	0
All	All	346/355 (97%)	0.54	24 (6%)	18	25	13, 29, 58, 79	0

All (24) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	440	HIS	5.6
1	A	320	GLY	4.2
1	A	256	ILE	4.0
1	A	314	ILE	3.9
1	A	250	THR	3.6
1	A	406	PHE	3.6
1	A	294	THR	3.5
1	A	319	SER	3.3
1	A	316	ILE	3.1
1	A	322	LEU	2.9
1	A	295	SER	2.9
1	A	299	ALA	2.7
1	A	293	VAL	2.7
1	A	441	ARG	2.7
1	A	296	TYR	2.6
1	A	310	ALA	2.6
1	A	409	GLY	2.6
1	A	438	ARG	2.4
1	A	575	TRP	2.4
1	A	439	SER	2.3
1	A	360	LEU	2.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	302	ILE	2.1
1	A	311	GLU	2.0
1	A	321	HIS	2.0

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

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(Å^2)	Q<0.9
4	8OG	T	6	23/24	0.97	0.12	-	14,17,20,21	0

6.3　Carbohydrates ⓘ

There are no carbohydrates in this entry.

6.4　Ligands ⓘ

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(Å^2)	Q<0.9
5	DUP	A	601	28/28	0.97	0.12	0.37	13,17,21,22	0
6	NA	A	603	1/1	0.99	0.11	-1.49	16,16,16,16	0
6	NA	A	602	1/1	0.97	0.12	-	17,17,17,17	0

6.5　Other polymers ⓘ

There are no such residues in this entry.