Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 04:36 PM GMT

PDB ID	:	4FJJ
Title	:	RB69 DNA polymerase ternary complex with dTTP/dC
Authors	:	Xia, S.; Wang, J.; Konigsberg, W.H.
Deposited on	:	2012-06-11
Resolution	:	1.99 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

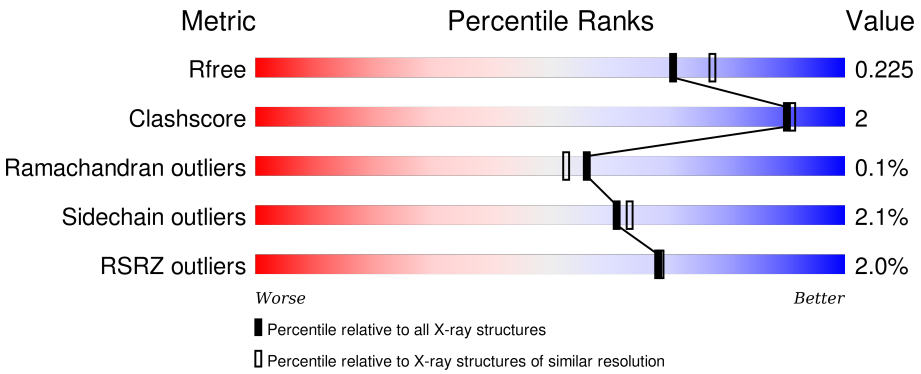
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.99 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	6249 (2.00-2.00)
Clashscore	102246	7340 (2.00-2.00)
Ramachandran outliers	100387	7248 (2.00-2.00)
Sidechain outliers	100360	7247 (2.00-2.00)
RSRZ outliers	91569	6262 (2.00-2.00)

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	903	2% 93% 7%
2	T	18	89% 11%
3	P	13	62% 31% 8%

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 8615 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
			Total	C	N	O	S			
1	A	901	7374	4733	1233	1375	33	0	5	0

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	222	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
A	327	ALA	ASP	ENGINEERED MUTATION	UNP Q38087
A	415	ALA	LEU	ENGINEERED MUTATION	UNP Q38087
A	561	ALA	LEU	ENGINEERED MUTATION	UNP Q38087
A	565	GLY	SER	ENGINEERED MUTATION	UNP Q38087
A	567	ALA	TYR	ENGINEERED MUTATION	UNP Q38087

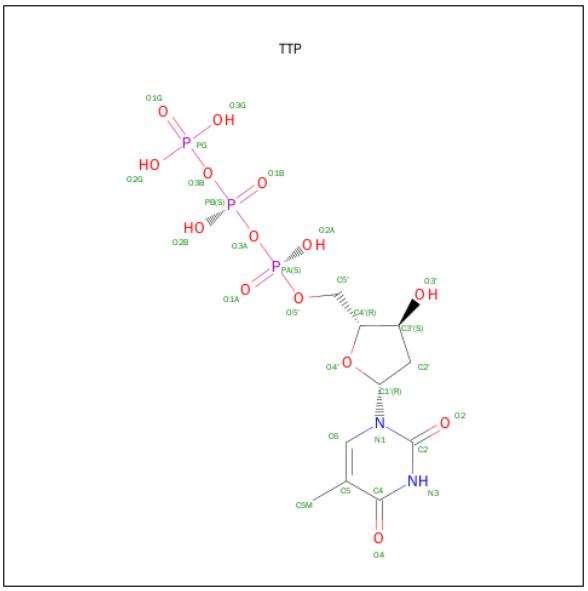
- Molecule 2 is a DNA chain called DNA template.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
2	T	18	366	174	69	106	17	0	0	0

- Molecule 3 is a DNA chain called DNA primer.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
3	P	13	262	126	48	76	12	0	0	0

- Molecule 4 is THYMIDINE-5'-TRIPHOSPHATE (three-letter code: TTP) (formula: $C_{10}H_{17}N_2O_{14}P_3$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
4	A	1	29	10	2	14	3	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
5	A	5	5	5	0	0

- Molecule 6 is water.

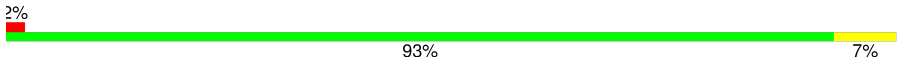
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
6	A	521	525	525	0	4
			Total	O		
6	T	36	36	36	0	0
			Total	O		
6	P	17	18	18	0	1

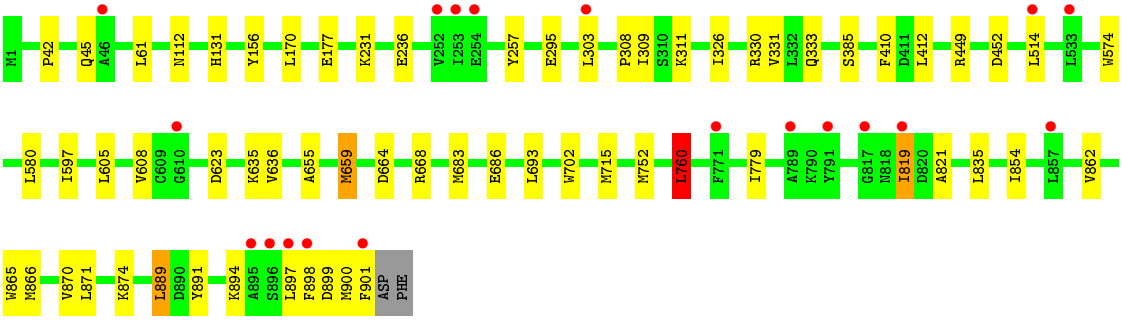
3 Residue-property plots (i)

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

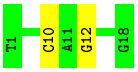
Chain A:





- Molecule 2: DNA template

Chain T:



- Molecule 3: DNA primer

Chain P:

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.20Å 119.82Å 130.64Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.31 – 1.99 49.31 – 1.99	Depositor EDS
% Data completeness (in resolution range)	98.1 (49.31-1.99) 98.2 (49.31-1.99)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.03 (at 1.98Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.184 , 0.225 0.183 , 0.225	Depositor DCC
R_{free} test set	4089 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	25.0	Xtriage
Anisotropy	0.077	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 39.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Outliers	1 of 80463 reflections (0.001%)	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8615	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.18% 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, 0.375 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: DOC, CA, TTP

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	0.48	3/7569 (0.0%)	0.55	2/10226 (0.0%)
2	T	0.34	0/410	0.82	1/631 (0.2%)
3	P	0.33	0/273	0.88	2/420 (0.5%)
All	All	0.47	3/8252 (0.0%)	0.58	5/11277 (0.0%)

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	702	TRP	CD2-CE2	5.60	1.48	1.41
1	A	865	TRP	CD2-CE2	5.02	1.47	1.41
1	A	574	TRP	CD2-CE2	5.01	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	105	DG	P-O3'-C3'	6.68	127.71	119.70
1	A	889	LEU	CA-CB-CG	5.82	128.69	115.30
1	A	760	LEU	CA-CB-CG	5.26	127.40	115.30
2	T	10	DC	P-O3'-C3'	5.25	126.00	119.70
3	P	108	DC	C1'-O4'-C4'	-5.01	105.09	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	A	7374	0	7281	32	0
2	T	366	0	203	1	0
3	P	262	0	148	4	0
4	A	29	0	13	0	0
5	A	5	0	0	0	0
6	A	525	0	0	3	0
6	P	18	0	0	1	0
6	T	36	0	0	0	0
All	All	8615	0	7645	36	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 2.

All (36) 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:330[B]:ARG:HH11	1:A:333:GLN:HE22	1.40	0.67
1:A:112[A]:ASN:ND2	6:A:1256:HOH:O	2.28	0.66
1:A:899:ASP:O	1:A:901:PHE:N	2.30	0.64
1:A:898:PHE:HB2	1:A:899:ASP:O	1.98	0.64
3:P:103:DG:H5"	6:P:201:HOH:O	1.99	0.63
1:A:170:LEU:HA	1:A:177:GLU:HG3	1.81	0.61
1:A:231:LYS:HD2	1:A:236:GLU:HG2	1.83	0.60
1:A:131:HIS:HD2	1:A:156:TYR:OH	1.87	0.58
1:A:870:VAL:HG13	1:A:874:LYS:HD3	1.85	0.57
1:A:605:LEU:HA	1:A:608:VAL:HG22	1.87	0.57
3:P:103:DG:H2"	3:P:104:DC:C6	2.41	0.56
1:A:760:LEU:HD13	1:A:891:TYR:HA	1.87	0.56
1:A:112[B]:ASN:HD21	1:A:331:VAL:CG1	2.19	0.54
1:A:874:LYS:HE3	2:T:12:DG:OP1	2.07	0.54
1:A:42:PRO:HD2	1:A:45:GLN:HE21	1.73	0.54
1:A:897:LEU:HD12	1:A:897:LEU:H	1.76	0.51
1:A:449:ARG:NH1	1:A:452:ASP:OD1	2.43	0.51
1:A:752:MET:HG3	1:A:760:LEU:HG	1.95	0.48
3:P:105:DG:H2"	3:P:106:DG:O5'	2.12	0.48
3:P:105:DG:C2'	3:P:106:DG:C8	2.96	0.48
1:A:231:LYS:HE3	6:A:1140:HOH:O	2.14	0.48
1:A:819:ILE:H	1:A:819:ILE:HD13	1.79	0.47
1:A:862:VAL:O	1:A:866:MET:HG3	2.14	0.47
1:A:779:ILE:HB	1:A:871:LEU:HD21	1.97	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:898:PHE:HA	1:A:899:ASP:C	2.36	0.46
1:A:410:PHE:HB3	1:A:683:MET:HG2	1.98	0.46
1:A:854:ILE:HD13	1:A:862:VAL:HG21	1.98	0.45
1:A:655:ALA:HA	1:A:659:MET:HG3	1.99	0.45
1:A:898:PHE:HB2	1:A:901:PHE:HD2	1.82	0.44
1:A:686:GLU:HG2	1:A:715:MET:SD	2.59	0.43
1:A:664:ASP:O	1:A:668:ARG:HG3	2.18	0.43
1:A:821:ALA:HA	6:A:1162:HOH:O	2.19	0.42
1:A:597:ILE:HA	1:A:597:ILE:HD12	1.95	0.42
1:A:412:LEU:HD12	1:A:623:ASP:HB2	2.02	0.41
1:A:326:ILE:O	1:A:330[B]:ARG:HG2	2.21	0.40
1:A:308:PRO:HG2	1:A:311:LYS:HG3	2.03	0.40

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	904/903 (100%)	885 (98%)	18 (2%)	1 (0%)	56 53

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	900	MET

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
1	A	799/796 (100%)	782 (98%)	17 (2%)	61 63

All (17) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	61	LEU
1	A	257	TYR
1	A	295	GLU
1	A	303	LEU
1	A	309	ILE
1	A	385	SER
1	A	514	LEU
1	A	580	LEU
1	A	635	LYS
1	A	636	VAL
1	A	659	MET
1	A	693	LEU
1	A	760	LEU
1	A	819	ILE
1	A	835	LEU
1	A	889	LEU
1	A	894	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	45	GLN
1	A	131	HIS
1	A	203	ASN
1	A	333	GLN
1	A	339	GLN
1	A	354	GLN
1	A	546	GLN
1	A	761	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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	$\#\|Z\| > 2$
3	DOC	P	115	3,2	11,19,20	0.77	0	14,26,29	1.35	2 (14%)

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
3	DOC	P	115	3,2	-	0/3/18/19	0/2/2/2

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	P	115	DOC	O4'-C4'-C5'	3.10	114.12	109.54
3	P	115	DOC	C2-N3-C4	3.13	120.03	115.61

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 (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

Of 6 ligands modelled in this entry, 5 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	$\#\|Z\| > 2$
4	TTP	A	1001	5	21,30,30	0.54	0	31,47,47	1.89	4 (12%)

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	TTP	A	1001	5	-	0/18/34/34	0/2/2/2

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	1001	TTP	C5-C4-N3	-5.31	119.22	125.14
4	A	1001	TTP	C2'-C1'-N1	-2.50	108.07	114.16
4	A	1001	TTP	O4'-C1'-N1	2.61	112.24	107.72
4	A	1001	TTP	C4-N3-C2	6.40	120.78	115.25

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 (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	901/903 (99%)	-0.01	19 (2%) 67 67	16, 26, 49, 101	0
2	T	18/18 (100%)	-0.43	0 100 100	21, 28, 43, 46	0
3	P	12/13 (92%)	-0.27	0 100 100	22, 38, 57, 61	0
All	All	931/934 (99%)	-0.02	19 (2%) 68 69	16, 27, 50, 101	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	898	PHE	5.4
1	A	897	LEU	4.0
1	A	895	ALA	4.0
1	A	896	SER	3.8
1	A	610	GLY	3.7
1	A	819	ILE	3.6
1	A	303	LEU	3.2
1	A	514	LEU	3.1
1	A	46	ALA	3.0
1	A	857	LEU	2.9
1	A	817	GLY	2.8
1	A	901	PHE	2.8
1	A	253	ILE	2.4
1	A	771	PHE	2.4
1	A	252	VAL	2.2
1	A	789	ALA	2.2
1	A	791	TYR	2.2
1	A	533	LEU	2.2
1	A	254	GLU	2.1

6.2 Non-standard residues in protein, DNA, RNA chains (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(Å^2)	Q<0.9
3	DOC	P	115	18/19	0.98	0.12	-	18,23,27,28	0

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(Å^2)	Q<0.9
4	TTP	A	1001	29/29	0.98	0.12	0.70	14,20,23,26	0
5	CA	A	1002	1/1	1.00	0.10	0.00	19,19,19,19	0
5	CA	A	1004	1/1	0.99	0.03	-1.92	36,36,36,36	0
5	CA	A	1003	1/1	0.97	0.11	-	49,49,49,49	0
5	CA	A	1006	1/1	0.87	0.07	-	63,63,63,63	0
5	CA	A	1005	1/1	0.99	0.07	-	40,40,40,40	0

6.5 Other polymers (i)

There are no such residues in this entry.