



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:51 PM GMT

PDB ID : 1VFG  
Title : Crystal structure of tRNA nucleotidyltransferase complexed with a primer tRNA and an incoming ATP analog  
Authors : Tomita, K.; Fukai, S.; Ishitani, R.; Ueda, T.; Takeuchi, N.; Vassylyev, D.G.; Nureki, O.; RIKEN Structural Genomics/Proteomics Initiative (RSGI)  
Deposited on : 2004-04-13  
Resolution : 2.80 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](mailto: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

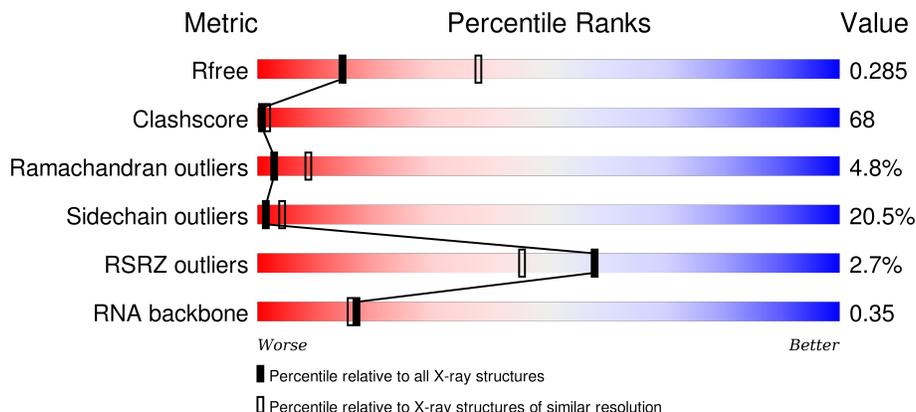
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)
RNA backbone	2183	1091 (3.20-2.40)

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  $\geq 3$ , 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	C	75	
1	D	75	
2	A	390	
2	B	390	

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 7319 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 RNA chain called RNA (75-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	P			
1	C	31	Total 661	C 293	N 116	O 221	P 31	0	0	0
1	D	34	Total 723	C 321	N 127	O 241	P 34	0	0	0

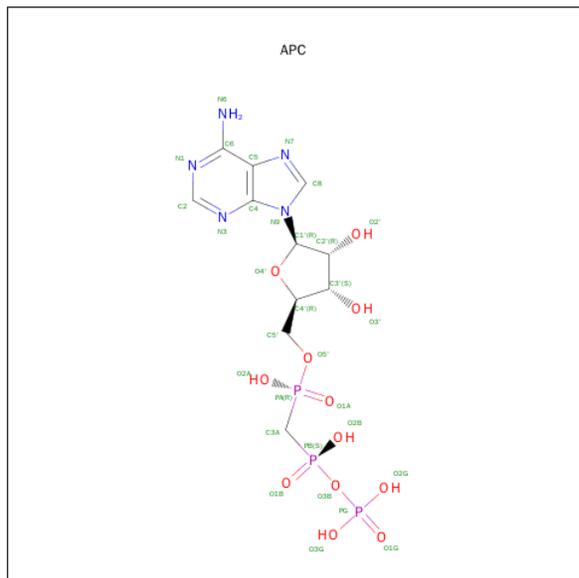
- Molecule 2 is a protein called poly A polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	A	342	Total 2833	C 1846	N 489	O 492	S 6	0	0	0
2	B	358	Total 2961	C 1928	N 512	O 514	S 7	0	0	0

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	INITIATING METHIONINE	UNP 066728
A	384	LYS	-	CLONING ARTIFACT	UNP 066728
A	385	LEU	-	CLONING ARTIFACT	UNP 066728
A	386	ALA	-	CLONING ARTIFACT	UNP 066728
A	387	ALA	-	CLONING ARTIFACT	UNP 066728
A	388	ALA	-	CLONING ARTIFACT	UNP 066728
A	389	LEU	-	CLONING ARTIFACT	UNP 066728
A	390	GLU	-	CLONING ARTIFACT	UNP 066728
B	1	MET	-	INITIATING METHIONINE	UNP 066728
B	384	LYS	-	CLONING ARTIFACT	UNP 066728
B	385	LEU	-	CLONING ARTIFACT	UNP 066728
B	386	ALA	-	CLONING ARTIFACT	UNP 066728
B	387	ALA	-	CLONING ARTIFACT	UNP 066728
B	388	ALA	-	CLONING ARTIFACT	UNP 066728
B	389	LEU	-	CLONING ARTIFACT	UNP 066728
B	390	GLU	-	CLONING ARTIFACT	UNP 066728

- Molecule 3 is DIPHOSPHOMETHYLPHOSPHONIC ACID ADENOSYL ESTER (three-letter code: APC) (formula:  $C_{11}H_{18}N_5O_{12}P_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	C	N	O	P		
3	A	1	Total	C	N	O	P	0	0
			31	11	5	12	3		
3	B	1	Total	C	N	O	P	0	0
			31	11	5	12	3		

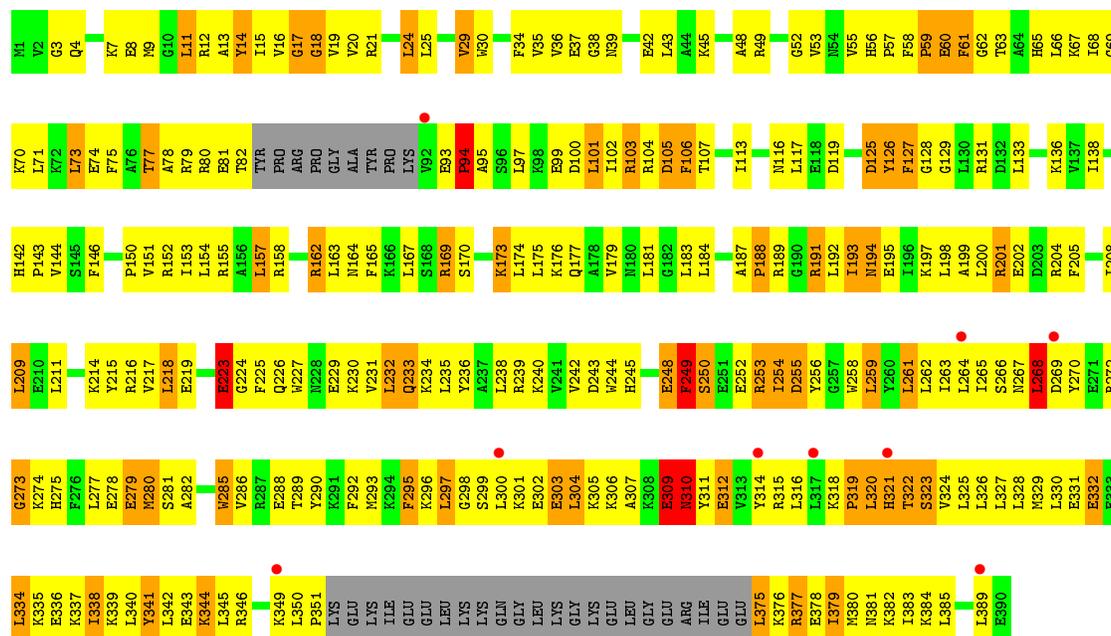
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	31	Total	O	0	0
			31	31		
4	B	38	Total	O	0	0
			38	38		
4	C	2	Total	O	0	0
			2	2		
4	D	8	Total	O	0	0
			8	8		





• Molecule 2: poly A polymerase



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	109.48Å 125.90Å 58.92Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	40.00 – 2.80 39.19 – 2.80	Depositor EDS
% Data completeness (in resolution range)	(Not available) (40.00-2.80) 99.1 (39.19-2.80)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.92 (at 2.81Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.230 , 0.286 0.228 , 0.285	Depositor DCC
$R_{free}$ test set	1945 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	71.5	Xtriage
Anisotropy	0.332	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 81.7	EDS
Estimated twinning fraction	0.499 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	2 of 39038 reflections (0.005%)	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	7319	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	87.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.49% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical 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:  
APC

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	C	0.62	1/735 (0.1%)	0.93	0/1140
1	D	0.64	1/804 (0.1%)	0.93	0/1247
2	A	0.61	1/2885 (0.0%)	0.78	3/3876 (0.1%)
2	B	0.63	0/3012	0.79	2/4041 (0.0%)
All	All	0.62	3/7436 (0.0%)	0.82	5/10304 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1
1	D	0	1
All	All	0	2

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	1	G	OP3-P	-7.11	1.52	1.61
1	D	1	G	OP3-P	-7.07	1.52	1.61
2	A	1	MET	CG-SD	5.60	1.95	1.81

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1	MET	CB-CG-SD	5.53	128.99	112.40
2	B	307	ALA	N-CA-C	5.49	125.83	111.00
2	A	81	GLU	N-CA-C	5.46	125.73	111.00
2	A	307	ALA	N-CA-C	5.40	125.59	111.00

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
2	B	259	LEU	CA-CB-CG	5.14	127.13	115.30

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	59	U	Sidechain
1	D	59	U	Sidechain

## 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	C	661	0	338	37	0
1	D	723	0	371	57	0
2	A	2833	0	2976	399	0
2	B	2961	0	3124	468	0
3	A	31	0	14	5	0
3	B	31	0	14	5	0
4	A	31	0	0	9	0
4	B	38	0	0	24	0
4	C	2	0	0	0	0
4	D	8	0	0	1	0
All	All	7319	0	6837	952	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 68.

The worst 5 of 952 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance ( $\text{\AA}$ )	Clash overlap ( $\text{\AA}$ )
2:B:253:ARG:HH11	2:B:253:ARG:HB3	1.04	1.15
2:B:302:GLU:HA	2:B:305:LYS:HE3	1.26	1.13
2:A:15:ILE:HD11	2:A:113:ILE:HB	1.22	1.10
2:A:5:ILE:HD12	2:A:47:LEU:HD22	1.32	1.08
2:A:39:ASN:HD21	2:A:41:ILE:HG12	1.12	1.06

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	
2	A	338/390 (87%)	274 (81%)	48 (14%)	16 (5%)	3	9
2	B	352/390 (90%)	285 (81%)	50 (14%)	17 (5%)	3	9
All	All	690/780 (88%)	559 (81%)	98 (14%)	33 (5%)	3	9

5 of 33 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	A	105	ASP
2	A	268	LEU
2	A	273	GLY
2	A	310	ASN
2	A	323	SER

### 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	
2	A	303/343 (88%)	241 (80%)	62 (20%)	1	4
2	B	316/343 (92%)	251 (79%)	65 (21%)	1	4
All	All	619/686 (90%)	492 (80%)	127 (20%)	1	4

5 of 127 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	A	312	GLU
2	B	60	GLU
2	B	321	HIS
2	A	325	LEU
2	B	14	TYR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 10 such sidechains are listed below:

Mol	Chain	Res	Type
2	B	51	HIS
2	B	116	ASN
2	B	194	ASN
2	A	228	ASN
2	B	177	GLN

### 5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
1	C	29/75 (38%)	8 (27%)	0
1	D	32/75 (42%)	8 (25%)	0
All	All	61/150 (40%)	16 (26%)	0

5 of 16 RNA backbone outliers are listed below:

Mol	Chain	Res	Type
1	C	5	A
1	C	55	U
1	C	56	C
1	C	57	G
1	C	58	A

There are no RNA pucker outliers to report.

## 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

2 ligands are 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	APC	A	500	-	25,33,33	1.90	5 (20%)	30,52,52	2.52	6 (20%)
3	APC	B	1500	-	25,33,33	1.45	2 (8%)	30,52,52	2.57	6 (20%)

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	APC	A	500	-	-	0/15/38/38	0/3/3/3
3	APC	B	1500	-	-	0/15/38/38	0/3/3/3

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	500	APC	O4'-C4'	-2.37	1.39	1.45
3	A	500	APC	C2-N1	2.27	1.38	1.33
3	A	500	APC	PB-O3B	2.69	1.61	1.58
3	B	1500	APC	PA-O5'	3.26	1.61	1.57
3	A	500	APC	C2-N3	3.51	1.38	1.32

The worst 5 of 12 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	A	500	APC	N3-C2-N1	-11.00	120.47	128.89
3	B	1500	APC	N3-C2-N1	-10.55	120.82	128.89
3	B	1500	APC	C4-C5-N7	-4.72	105.14	109.48
3	A	500	APC	C4-C5-N7	-4.08	105.73	109.48
3	B	1500	APC	O4'-C1'-N9	-3.15	101.50	108.10

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 10 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	500	APC	5	0
3	B	1500	APC	5	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	C	31/75 (41%)	-0.31	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	131, 142, 152, 153	0
1	D	34/75 (45%)	-0.34	0 <span style="border: 1px solid blue; padding: 2px;">100</span> <span style="border: 1px solid blue; padding: 2px;">100</span>	122, 136, 148, 154	0
2	A	342/390 (87%)	0.08	12 (3%) <span style="border: 1px solid gray; padding: 2px;">48</span> <span style="border: 1px solid red; padding: 2px;">35</span>	33, 65, 126, 143	0
2	B	358/390 (91%)	0.00	9 (2%) <span style="border: 1px solid gray; padding: 2px;">61</span> <span style="border: 1px solid gray; padding: 2px;">48</span>	33, 63, 119, 132	0
All	All	765/930 (82%)	0.01	21 (2%) <span style="border: 1px solid gray; padding: 2px;">58</span> <span style="border: 1px solid gray; padding: 2px;">45</span>	33, 68, 139, 154	0

The worst 5 of 21 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	A	351	PRO	7.1
2	B	92	VAL	4.5
2	A	348	VAL	4.3
2	B	314	TYR	3.3
2	A	330	LEU	3.1

### 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, 95<sup>th</sup> 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( $\text{\AA}^2$ )	Q<0.9
3	APC	B	1500	31/31	0.93	0.16	-0.33	75,80,89,91	0
3	APC	A	500	31/31	0.93	0.15	-1.08	65,71,84,84	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.