



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 06:58 pm GMT

PDB ID : 4ZDK  
Title : Crystal structure of the M. tuberculosis CTP synthase PyrG in complex with UTP, AMP-PCP and oxonorleucine  
Authors : Bellinzoni, M.; Barilone, N.; Alzari, P.M.  
Deposited on : 2015-04-17  
Resolution : 3.49 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](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 : 1.7.2 (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

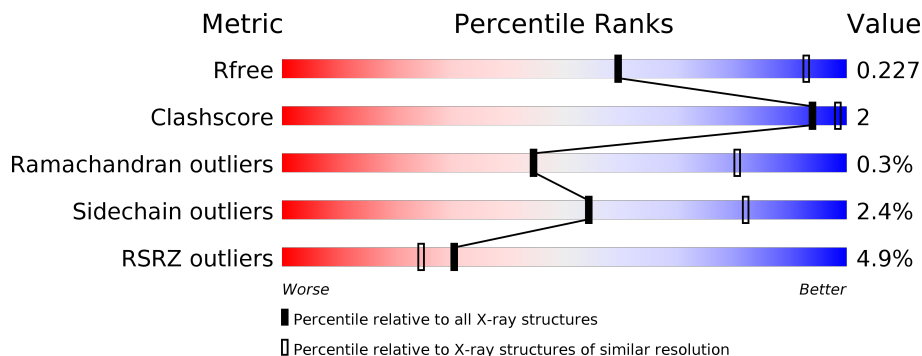
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.49 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))
$R_{free}$	100719	1195 (3.60-3.40)
Clashscore	112137	1322 (3.60-3.40)
Ramachandran outliers	110173	1283 (3.60-3.40)
Sidechain outliers	110143	1284 (3.60-3.40)
RSRZ outliers	101464	1226 (3.60-3.40)

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	A	586	<div> <div style="width: 100%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">%</span> </div> <div style="width: 86%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">86%</span> </div> <div style="width: 5%; height: 10px; background-color: yellow; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">5%</span> </div> <div style="width: 9%; height: 10px; background-color: grey; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">9%</span> </div> </div>
1	B	586	<div> <div style="width: 8%; height: 10px; background-color: red; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">8%</span> </div> <div style="width: 87%; height: 10px; background-color: green; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">87%</span> </div> <div style="width: 5%; height: 10px; background-color: yellow; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">•</span> </div> <div style="width: 9%; height: 10px; background-color: grey; position: relative;"> <span style="position: absolute; top: -5px; left: 0;">9%</span> </div> </div>

## 2 Entry composition [i](#)

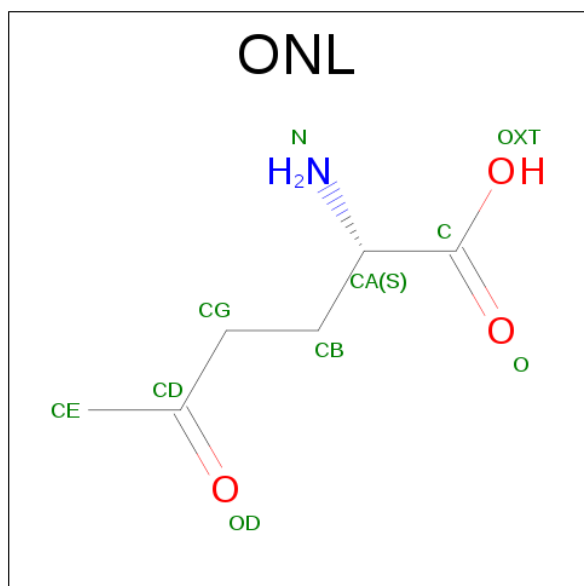
There are 5 unique types of molecules in this entry. The entry contains 7986 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 CTP synthase.

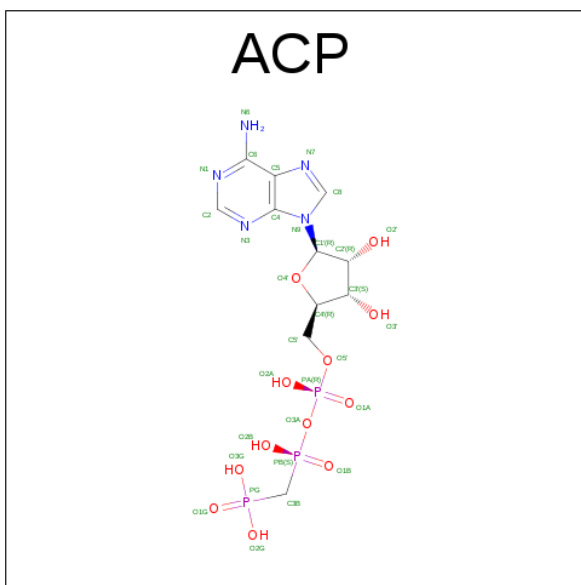
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	534	Total	C	N	O	S	0	0	0
			3969	2502	701	756	10			
1	B	534	Total	C	N	O	S	0	0	0
			3875	2431	688	746	10			

- Molecule 2 is 5-OXO-L-NORLEUCINE (three-letter code: ONL) (formula:  $C_6H_{11}NO_3$ ).



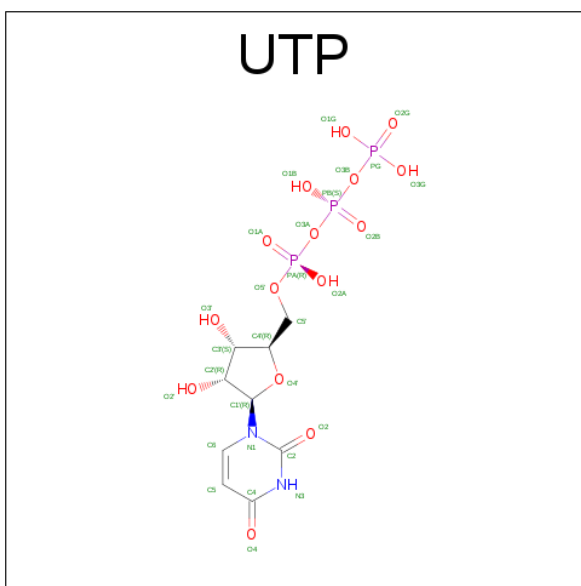
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			10	6	1	3		
2	B	1	Total	C	N	O	0	0
			10	6	1	3		

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



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

- Molecule 4 is URIDINE 5'-TRIPHOSPHATE (three-letter code: UTP) (formula:  $\text{C}_9\text{H}_{15}\text{N}_2\text{O}_{15}\text{P}_3$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total 29	C 9	N 2	O 15	P 3	0	0
4	B	1	Total 29	C 9	N 2	O 15	P 3	0	0

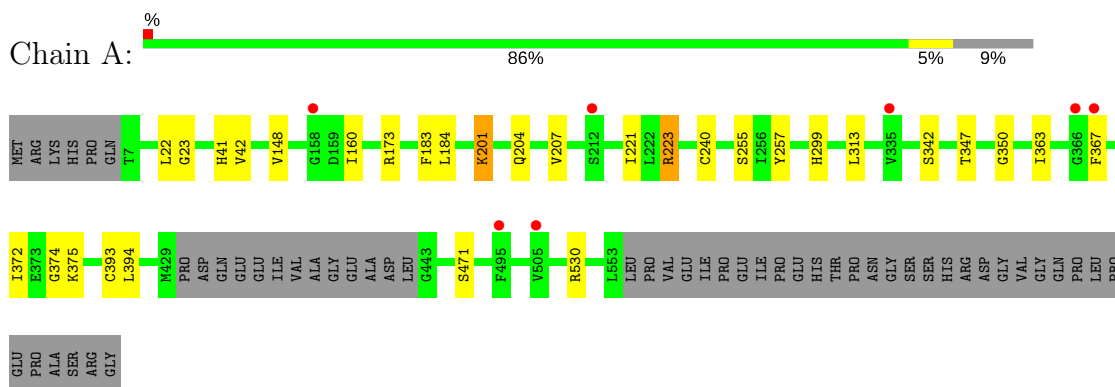
- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	B	1	Total 1	Mg 1	0	0
5	A	1	Total 1	Mg 1	0	0

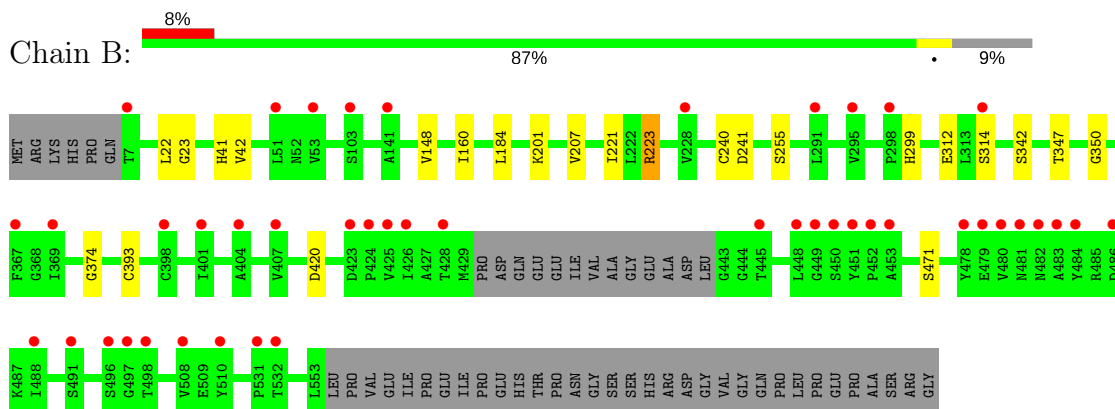
### 3 Residue-property plots [i](#)

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



#### • Molecule 1: CTP synthase



## 4 Data and refinement statistics

Property	Value	Source
Space group	I 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.73Å 194.72Å 207.46Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.68 – 3.49 48.68 – 3.49	Depositor EDS
% Data completeness (in resolution range)	99.9 (48.68-3.49) 99.6 (48.68-3.49)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.90 (at 3.48Å)	Xtriage
Refinement program	BUSTER 2.10.1	Depositor
R, $R_{free}$	0.203 , 0.214 0.216 , 0.227	Depositor DCC
$R_{free}$ test set	1599 reflections (5.31%)	DCC
Wilson B-factor (Å <sup>2</sup> )	133.4	Xtriage
Anisotropy	0.375	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 108.1	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7986	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	143.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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, 0.333 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: UTP, ACP, MG, ONL

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	0.41	0/4057	0.62	0/5539
1	B	0.41	0/3958	0.61	0/5410
All	All	0.41	0/8015	0.61	0/10949

There are no bond length outliers.

There are no bond angle outliers.

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	3969	0	3811	15	0
1	B	3875	0	3634	9	0
2	A	10	0	9	1	0
2	B	10	0	8	0	0
3	A	31	0	14	1	0
3	B	31	0	14	1	0
4	A	29	0	11	1	0
4	B	29	0	11	1	0
5	A	1	0	0	0	0
5	B	1	0	0	0	0
All	All	7986	0	7512	24	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 (24) 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:22:LEU:O	1:A:223:ARG:HD3	1.78	0.82
1:B:22:LEU:O	1:B:223:ARG:HD3	1.82	0.78
1:B:23:GLY:HA2	3:B:602:ACP:O2A	1.90	0.72
1:B:42:VAL:HG22	1:B:148:VAL:CG1	2.26	0.65
1:A:42:VAL:HG22	1:A:148:VAL:CG1	2.27	0.63
1:A:22:LEU:O	1:A:223:ARG:CD	2.50	0.60
1:B:160:ILE:H	4:B:603:UTP:H3	1.51	0.59
1:A:394:LEU:HB2	2:A:601:ONL:HG2	1.91	0.53
1:A:160:ILE:H	4:A:603:UTP:H3	1.56	0.52
1:A:173:ARG:HE	1:A:183:PHE:HE2	1.59	0.51
1:B:42:VAL:HG22	1:B:148:VAL:HG13	1.94	0.49
1:A:42:VAL:HG22	1:A:148:VAL:HG13	1.94	0.49
1:B:347:THR:HG23	1:B:350:GLY:H	1.79	0.47
1:A:23:GLY:HA3	3:A:602:ACP:H8	1.98	0.45
1:A:530:ARG:HH11	1:A:530:ARG:HD3	1.65	0.45
1:A:201:LYS:HE3	1:A:204:GLN:OE1	2.18	0.43
1:B:342:SER:O	1:B:374:GLY:HA3	2.18	0.43
1:A:347:THR:HG23	1:A:350:GLY:H	1.84	0.42
1:A:184:LEU:HD11	1:A:221:ILE:HG12	2.01	0.42
1:A:342:SER:O	1:A:374:GLY:HA3	2.18	0.42
1:A:207:VAL:HG11	1:A:240:CYS:HB3	2.01	0.42
1:B:184:LEU:HD11	1:B:221:ILE:HG12	2.01	0.42
1:A:363:ILE:CG2	1:A:375:LYS:HE3	2.50	0.41
1:B:207:VAL:HG11	1:B:240:CYS:HB3	2.02	0.41

There are no symmetry-related clashes.

## 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	530/586 (90%)	517 (98%)	11 (2%)	2 (0%)	38	77
1	B	530/586 (90%)	521 (98%)	8 (2%)	1 (0%)	51	85
All	All	1060/1172 (90%)	1038 (98%)	19 (2%)	3 (0%)	44	80

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	313	LEU
1	B	393	CYS
1	A	393	CYS

### 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	404/480 (84%)	395 (98%)	9 (2%)	57	83
1	B	380/480 (79%)	370 (97%)	10 (3%)	51	80
All	All	784/960 (82%)	765 (98%)	19 (2%)	54	82

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	41	HIS
1	A	201	LYS
1	A	223	ARG
1	A	255	SER
1	A	257	TYR
1	A	299	HIS
1	A	367	PHE
1	A	372	ILE
1	A	471	SER
1	B	41	HIS
1	B	201	LYS
1	B	223	ARG
1	B	241	ASP

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	255	SER
1	B	299	HIS
1	B	312	GLU
1	B	314	SER
1	B	420	ASP
1	B	471	SER

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

### 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 8 ligands modelled in this entry, 2 are monoatomic - leaving 6 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	$\# Z  > 2$
2	ONL	A	601	1	4,9,9	0.13	0	5,11,11	0.38	0
3	ACP	A	602	5	27,33,33	0.85	2 (7%)	30,52,52	1.19	2 (6%)
4	UTP	A	603	-	26,30,30	2.39	4 (15%)	29,47,47	2.23	1 (3%)
2	ONL	B	601	1	4,9,9	0.14	0	5,11,11	0.85	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
3	ACP	B	602	5	27,33,33	1.09	3 (11%)	30,52,52	1.24	3 (10%)
4	UTP	B	603	-	26,30,30	2.41	4 (15%)	29,47,47	2.23	1 (3%)

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
2	ONL	A	601	1	-	0/5/9/9	0/0/0/0
3	ACP	A	602	5	-	0/15/38/38	0/3/3/3
4	UTP	A	603	-	-	0/22/38/38	0/2/2/2
2	ONL	B	601	1	-	0/5/9/9	0/0/0/0
3	ACP	B	602	5	-	0/15/38/38	0/3/3/3
4	UTP	B	603	-	-	0/22/38/38	0/2/2/2

All (13) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	603	UTP	C6-C5	-9.41	1.35	1.52
4	B	603	UTP	C6-C5	-9.37	1.35	1.52
4	B	603	UTP	C6-N1	-6.40	1.39	1.47
4	A	603	UTP	C6-N1	-6.32	1.39	1.47
4	A	603	UTP	C5-C4	-2.59	1.43	1.50
4	B	603	UTP	C5-C4	-2.57	1.44	1.50
3	A	602	ACP	PG-O3G	-2.11	1.49	1.54
3	B	602	ACP	PG-O3G	-2.08	1.50	1.54
4	B	603	UTP	C2-N1	2.06	1.38	1.35
4	A	603	UTP	C2-N1	2.16	1.38	1.35
3	B	602	ACP	PG-O1G	2.25	1.55	1.50
3	A	602	ACP	PG-O1G	2.37	1.55	1.50
3	B	602	ACP	PB-O3A	3.69	1.62	1.58

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	602	ACP	O2G-PG-O1G	-2.31	106.13	112.32
3	A	602	ACP	O3G-PG-C3B	2.06	111.39	106.40
3	B	602	ACP	O3G-PG-C3B	2.22	111.78	106.40
3	B	602	ACP	O1B-PB-C3B	4.03	118.94	108.97
3	A	602	ACP	O1B-PB-C3B	4.57	120.27	108.97
4	B	603	UTP	C5-C6-N1	11.03	122.16	110.70

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	603	UTP	C5-C6-N1	11.11	122.25	110.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	ONL	1	0
3	A	602	ACP	1	0
4	A	603	UTP	1	0
3	B	602	ACP	1	0
4	B	603	UTP	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

### 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, 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	A	534/586 (91%)	0.13	7 (1%) 77 69	94, 125, 152, 180	0
1	B	534/586 (91%)	0.47	45 (8%) 12 12	115, 156, 193, 211	0
All	All	1068/1172 (91%)	0.30	52 (4%) 30 24	94, 140, 186, 211	0

All (52) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	481	ASN	6.2
1	B	480	VAL	4.1
1	B	445	THR	3.8
1	B	482	ASN	3.6
1	B	488	ILE	3.6
1	B	452	PRO	3.3
1	A	367	PHE	3.2
1	B	451	TYR	3.2
1	B	532	THR	3.1
1	B	398	CYS	3.1
1	A	366	GLY	3.1
1	B	367	PHE	3.0
1	B	103	SER	3.0
1	B	497	GLY	2.9
1	B	401	ILE	2.9
1	B	298	PRO	2.8
1	B	496	SER	2.8
1	B	479	GLU	2.7
1	B	423	ASP	2.7
1	B	407	VAL	2.7
1	B	51	LEU	2.7
1	B	426	ILE	2.7
1	B	531	PRO	2.6
1	A	335	VAL	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	510	TYR	2.6
1	A	158	GLY	2.5
1	B	425	VAL	2.5
1	B	228	VAL	2.5
1	B	424	PRO	2.5
1	B	450	SER	2.5
1	B	498	THR	2.4
1	B	484	TYR	2.4
1	B	291	LEU	2.4
1	B	448	LEU	2.4
1	B	7	THR	2.3
1	B	53	VAL	2.3
1	B	449	GLY	2.3
1	B	141	ALA	2.3
1	B	508	VAL	2.3
1	B	491	SER	2.3
1	B	428	THR	2.3
1	B	478	TYR	2.2
1	B	295	VAL	2.2
1	A	495	PHE	2.2
1	B	486	ASP	2.2
1	B	483	ALA	2.2
1	A	505	VAL	2.2
1	B	453	ALA	2.1
1	B	314	SER	2.1
1	B	369	ILE	2.1
1	A	212	SER	2.0
1	B	404	ALA	2.0

## 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(Å <sup>2</sup> )	Q<0.9
3	ACP	B	602	31/31	0.81	0.33	1.49	177,183,185,187	0
2	ONL	A	601	10/10	0.90	0.30	-0.06	119,122,131,132	0
3	ACP	A	602	31/31	0.82	0.22	-0.10	134,136,146,147	0
4	UTP	A	603	29/29	0.88	0.26	-0.24	210,213,216,217	0
5	MG	B	604	1/1	0.96	0.17	-0.71	113,113,113,113	0
2	ONL	B	601	10/10	0.74	0.26	-0.79	150,157,163,166	0
4	UTP	B	603	29/29	0.90	0.18	-1.30	162,164,167,168	0
5	MG	A	604	1/1	0.98	0.15	-	94,94,94,94	0

## 6.5 Other polymers

There are no such residues in this entry.