



wwPDB X-ray Structure Validation Summary Report ⓘ

May 21, 2020 – 01:31 am BST

PDB ID : 6OW3
Title : X-ray crystal structure of a bacterial reiterative transcription complex of pyrG promoter variant -1T
Authors : Shin, Y.; Murakami, K.S.
Deposited on : 2019-05-09
Resolution : 2.77 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix) : 1.13
EDS : 2.11
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac : 5.8.0158
CCP4 : 7.0.044 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : 2.11

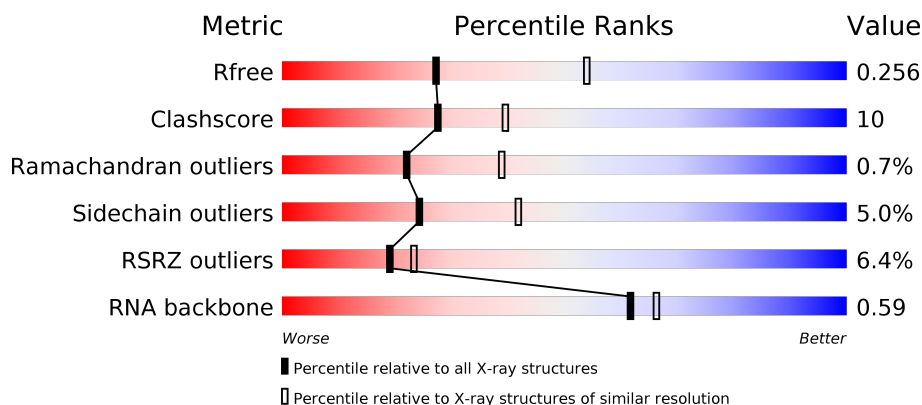
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.77 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}	130704	1235 (2.78-2.74)
Clashscore	141614	1277 (2.78-2.74)
Ramachandran outliers	138981	1257 (2.78-2.74)
Sidechain outliers	138945	1257 (2.78-2.74)
RSRZ outliers	127900	1207 (2.78-2.74)
RNA backbone	3102	1060 (3.02-2.50)

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 respectively. 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	315	<div> <div>6%</div> <div>54% 16% • 28%</div> </div>
1	B	315	<div> <div>52% 18% • 29%</div> </div>
2	C	1119	<div> <div>6%</div> <div>72% 25% ••</div> </div>
3	D	1524	<div> <div>6%</div> <div>73% 22% ••</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	E	99	<div><div></div><div>3%</div><div>77%</div><div>18%</div><div>5%</div></div>
5	F	423	<div><div></div><div>13%</div><div>61%</div><div>19%</div><div>18%</div></div>
6	G	22	<div><div></div><div>23%</div><div>41%</div><div>9%</div><div>27%</div></div>
7	H	27	<div><div></div><div>7%</div><div>37%</div><div>33%</div><div>30%</div></div>
8	I	4	<div><div></div><div>25%</div><div>50%</div><div>25%</div></div>

2 Entry composition

There are 11 unique types of molecules in this entry. The entry contains 28430 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-directed RNA polymerase subunit alpha.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	226	Total	C	N	O	S	0	0	0
			1782	1138	310	332	2			
1	B	224	Total	C	N	O	S	0	0	0
			1767	1129	307	329	2			

- Molecule 2 is a protein called DNA-directed RNA polymerase subunit beta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	1111	Total	C	N	O	S	0	0	0
			8758	5542	1558	1634	24			

- Molecule 3 is a protein called DNA-directed RNA polymerase subunit beta'.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	1484	Total	C	N	O	S	0	0	0
			11722	7431	2065	2191	35			

- Molecule 4 is a protein called DNA-directed RNA polymerase subunit omega.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	E	94	Total	C	N	O	S	0	0	0
			761	486	132	139	4			

- Molecule 5 is a protein called RNA polymerase sigma factor SigA.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	F	346	Total	C	N	O	S	0	0	0
			2807	1770	509	524	4			

There is a discrepancy between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
F	46	THR	ALA	conflict	UNP Q72L95

- Molecule 6 is a DNA chain called DNA (5'-D(P*GP*CP*AP*TP*CP*AP*GP*AP*GP*CP*CP*CP*TP*AP*AP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
6	G	16	Total	C	N	O	P	0	0	0
			327	155	64	92	16			

- Molecule 7 is a DNA chain called DNA (5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*GP*AP*TP*CP*TP*GP*AP*TP*GP*C)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
7	H	19	Total	C	N	O	P	0	0	0
			392	188	73	113	18			

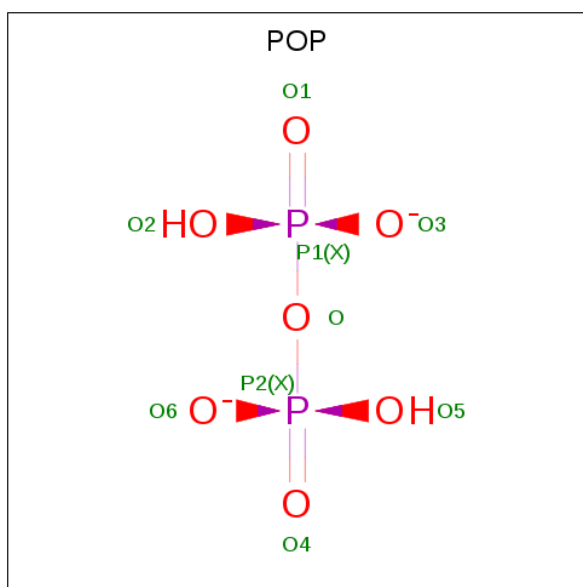
- Molecule 8 is a RNA chain called RNA (5'-D(*(GTP))-R(P*GP*GP*G)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
8	I	4	Total	C	N	O	P	0	0	0
			101	40	20	35	6			

- Molecule 9 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	D	2	Total	Mg	0	0
			2	2		

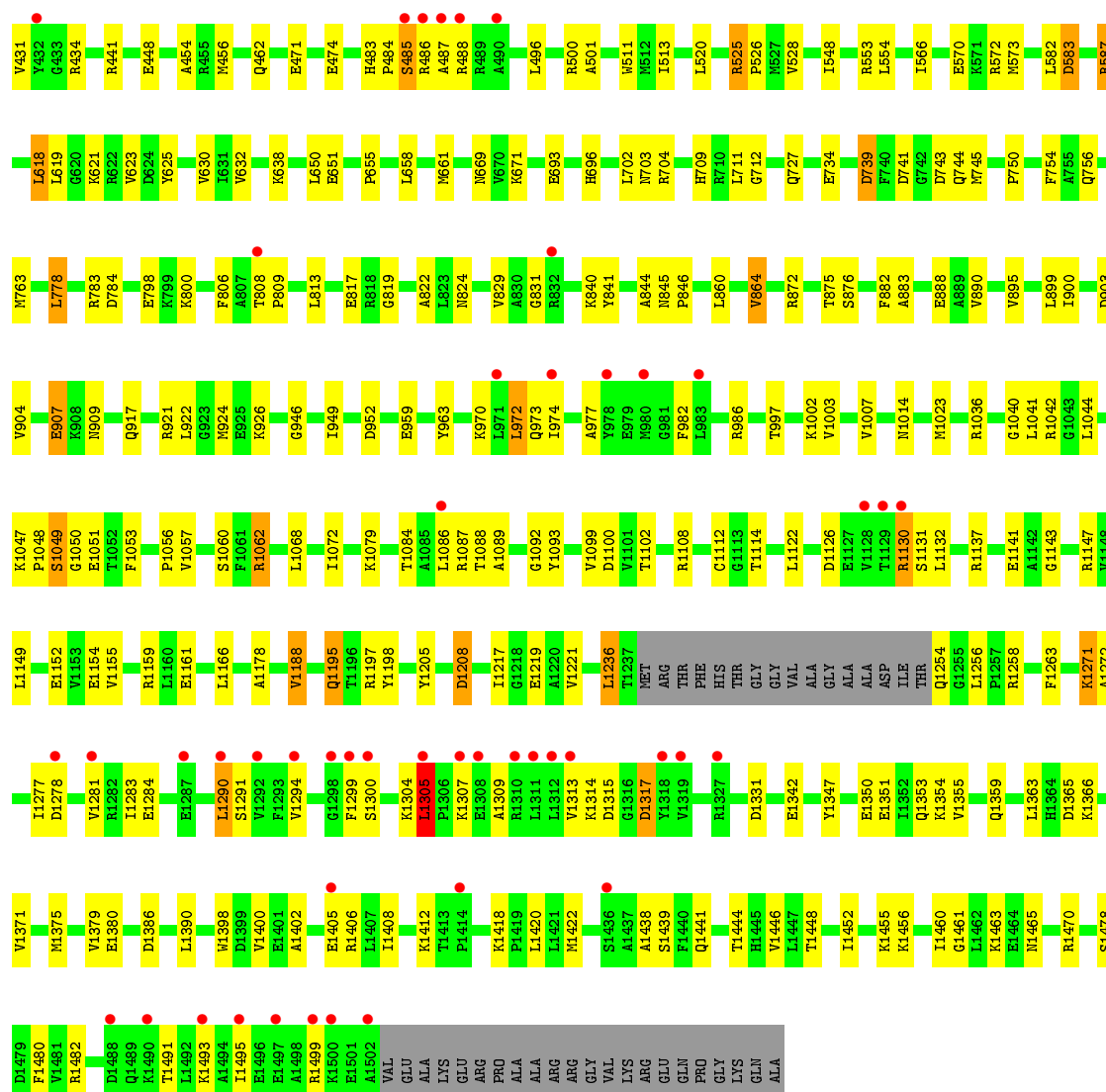
- Molecule 10 is PYROPHOSPHATE 2- (three-letter code: POP) (formula: H₂O₇P₂).



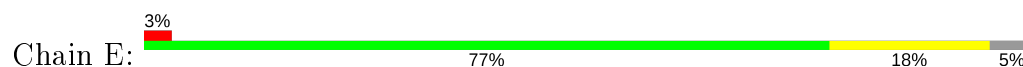
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
10	D	1	Total	O	P	0	0
			9	7	2		

- Molecule 11 is ZINC ION (three-letter code: ZN) (formula: Zn).

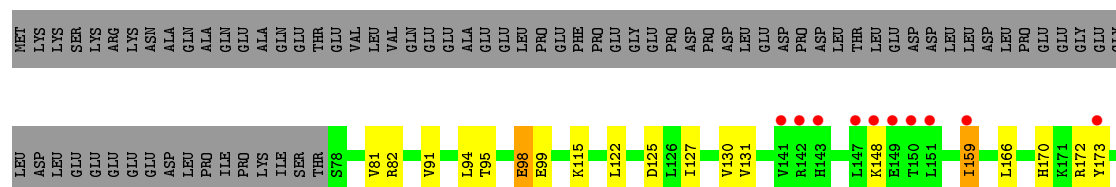
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
11	D	2	Total	Zn	0	0
			2	2		

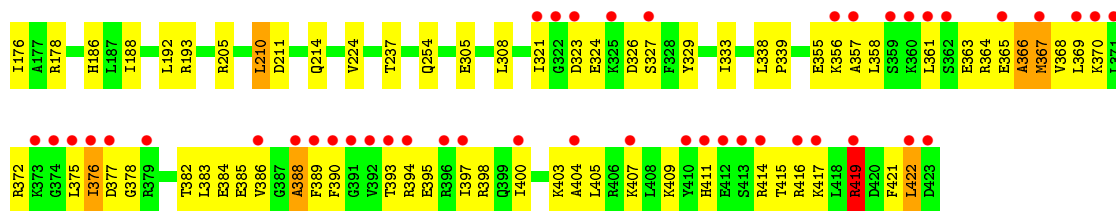


• Molecule 4: DNA-directed RNA polymerase subunit omega

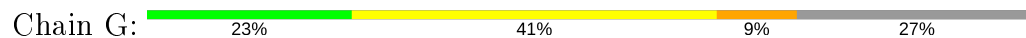


• Molecule 5: RNA polymerase sigma factor SigA

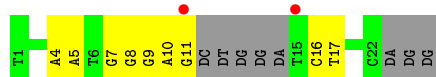




- Molecule 6: DNA (5'-D(P*GP*CP*AP*TP*CP*AP*GP*AP*GP*CP*CP*TP*AP*AP*A)-3')



- Molecule 7: DNA (5'-D(*TP*AP*TP*AP*AP*TP*GP*GP*GP*GP*AP*TP*CP*TP*GP*AP*TP*GP*C)-3')



- Molecule 8: RNA (5'-D(*(GTP))-R(P*GP*GP*G)-3')



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	186.19Å 101.83Å 295.75Å 90.00° 98.64° 90.00°	Depositor
Resolution (Å)	36.87 – 2.77 36.87 – 2.77	Depositor EDS
% Data completeness (in resolution range)	97.4 (36.87-2.77) 97.4 (36.87-2.77)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.61 (at 2.77Å)	Xtriage
Refinement program	PHENIX (1.14_3260: ???)	Depositor
R, R_{free}	0.217 , 0.255 0.217 , 0.256	Depositor DCC
R_{free} test set	2006 reflections (1.47%)	wwPDB-VP
Wilson B-factor (Å ²)	76.1	Xtriage
Anisotropy	0.668	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 53.2	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	28430	wwPDB-VP
Average B, all atoms (Å ²)	94.0	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.*

¹Intensities estimated from amplitudes.

²Theoretical 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: GTP, ZN, MG, POP

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	0.55	0/1814	0.82	1/2466 (0.0%)
1	B	0.51	0/1799	0.78	0/2447
2	C	0.52	0/8925	0.80	3/12073 (0.0%)
3	D	0.57	1/11928 (0.0%)	0.83	7/16127 (0.0%)
4	E	0.48	0/775	0.76	0/1045
5	F	0.51	0/2852	0.89	5/3837 (0.1%)
6	G	1.41	4/367 (1.1%)	1.45	11/563 (2.0%)
7	H	1.01	0/439	1.10	1/675 (0.1%)
8	I	1.75	2/77 (2.6%)	1.40	0/119
All	All	0.57	7/28976 (0.0%)	0.84	28/39352 (0.1%)

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
2	C	0	1
3	D	0	3
5	F	0	3
All	All	0	7

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
8	I	3	G	O3'-P	-9.75	1.49	1.61
6	G	14	DG	N9-C4	-9.26	1.30	1.38
3	D	734	GLU	CG-CD	5.93	1.60	1.51
6	G	14	DG	N3-C4	-5.80	1.31	1.35
8	I	4	G	O3'-P	-5.49	1.54	1.61

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	G	13	DA	O5'-P-OP2	-10.32	96.42	105.70
6	G	16	DC	O5'-P-OP2	-9.40	97.24	105.70
5	F	422	LEU	CB-CG-CD2	8.66	125.73	111.00
6	G	13	DA	O5'-P-OP1	8.04	120.35	110.70
3	D	739	ASP	CB-CG-OD1	-7.68	111.39	118.30

There are no chirality outliers.

5 of 7 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	C	766	GLU	Peptide
3	D	1195	GLN	Sidechain
3	D	583	ASP	Mainchain
3	D	903	ASP	Mainchain
5	F	363	GLU	Mainchain

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	1782	0	1833	40	0
1	B	1767	0	1816	38	0
2	C	8758	0	8852	219	0
3	D	11722	0	11950	234	1
4	E	761	0	778	15	0
5	F	2807	0	2882	71	0
6	G	327	0	179	11	0
7	H	392	0	218	9	0
8	I	101	0	44	7	0
9	D	2	0	0	0	0
10	D	9	0	0	0	0
11	D	2	0	0	0	0
All	All	28430	0	28552	567	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 10.

The worst 5 of 567 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)
5:F:382:THR:HG22	5:F:383:LEU:H	1.25	0.98
3:D:218:LYS:HG2	3:D:338:GLU:HG2	1.49	0.95
3:D:899:LEU:HD21	3:D:921:ARG:HD3	1.54	0.88
5:F:358:LEU:HB3	5:F:366:ALA:HB1	1.55	0.88
5:F:365:GLU:OE1	5:F:403:LYS:NZ	2.09	0.85

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)
3:D:34:TYR:OH	3:D:327:GLU:OE1[4_1359]	2.08	0.12

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	224/315 (71%)	220 (98%)	4 (2%)	0	100	100
1	B	222/315 (70%)	210 (95%)	11 (5%)	1 (0%)	29	47
2	C	1107/1119 (99%)	1062 (96%)	36 (3%)	9 (1%)	19	34
3	D	1480/1524 (97%)	1434 (97%)	39 (3%)	7 (0%)	29	47
4	E	92/99 (93%)	89 (97%)	2 (2%)	1 (1%)	14	25
5	F	344/423 (81%)	319 (93%)	18 (5%)	7 (2%)	7	13
All	All	3469/3795 (91%)	3334 (96%)	110 (3%)	25 (1%)	22	39

5 of 25 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	C	212	GLY

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	C	363	SER
3	D	486	ARG
5	F	324	GLU
5	F	388	ALA

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	199/273 (73%)	189 (95%)	10 (5%)	24	42
1	B	197/273 (72%)	187 (95%)	10 (5%)	24	41
2	C	934/941 (99%)	889 (95%)	45 (5%)	25	44
3	D	1250/1279 (98%)	1183 (95%)	67 (5%)	22	38
4	E	83/88 (94%)	80 (96%)	3 (4%)	35	55
5	F	301/371 (81%)	288 (96%)	13 (4%)	29	48
All	All	2964/3225 (92%)	2816 (95%)	148 (5%)	24	42

5 of 148 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	D	67	ARG
3	D	312	ARG
5	F	125	ASP
3	D	106	LYS
3	D	191	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 11 such sidechains are listed below:

Mol	Chain	Res	Type
3	D	388	HIS
3	D	714	GLN
3	D	744	GLN
2	C	728	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
3	D	737	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
8	I	3/4 (75%)	0	1 (33%)

There are no RNA backbone outliers to report.

All (1) RNA pucker outliers are listed below:

Mol	Chain	Res	Type
8	I	2	GTP

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, 4 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	$\# Z > 2$
10	POP	D	1602	-	6,8,8	0.86	0	13,13,13	0.94	1 (7%)

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
10	POP	D	1602	-	-	0/6/6/6	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed($^{\circ}$)	Ideal($^{\circ}$)
10	D	1602	POP	P2-O-P1	-2.32	124.86	132.83

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	226/315 (71%)	0.03	2 (0%) 84 89	65, 88, 111, 123	0
1	B	224/315 (71%)	0.11	1 (0%) 92 95	63, 90, 120, 134	0
2	C	1111/1119 (99%)	0.38	69 (6%) 20 25	50, 87, 154, 180	0
3	D	1484/1524 (97%)	0.35	93 (6%) 20 24	45, 81, 145, 175	0
4	E	94/99 (94%)	0.42	3 (3%) 47 56	60, 98, 137, 147	0
5	F	346/423 (81%)	0.90	55 (15%) 1 2	56, 97, 180, 196	0
6	G	16/22 (72%)	-0.04	0 100 100	56, 92, 179, 180	0
7	H	19/27 (70%)	-0.00	2 (10%) 6 7	82, 125, 162, 166	0
8	I	3/4 (75%)	-1.02	0 100 100	58, 58, 64, 73	0
All	All	3523/3848 (91%)	0.38	225 (6%) 19 23	45, 87, 154, 196	0

The worst 5 of 225 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
5	F	375	LEU	11.9
5	F	392	VAL	11.7
5	F	376	ILE	11.6
5	F	390	PHE	10.0
5	F	393	THR	10.0

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. 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	B-factors(\AA^2)	Q<0.9
9	MG	D	1601	1/1	0.89	0.51	71,71,71,71	0
10	POP	D	1602	9/9	0.94	0.25	69,80,94,96	0
9	MG	D	1603	1/1	0.96	0.21	50,50,50,50	0
11	ZN	D	1604	1/1	0.96	0.11	116,116,116,116	0
11	ZN	D	1605	1/1	0.99	0.24	89,89,89,89	0

6.5 Other polymers [i](#)

There are no such residues in this entry.