



wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 20, 2020 – 03:18 PM BST

PDB ID : 6D9K
Title : Ternary RsAgo Complex with Guide RNA and Target DNA Containing A-G Non-canonical Pair
Authors : Liu, Y.; Esyunina, D.; Olovnikov, I.; Teplova, M.; Patel, D.J.
Deposited on : 2018-04-30
Resolution : 2.00 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 : 2.13.1
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) : 2.13.1

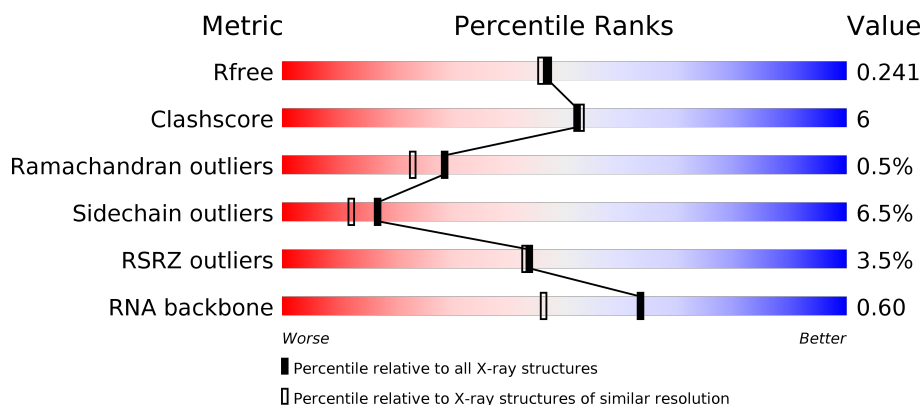
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.00 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	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)
RNA backbone	3102	1079 (2.50-1.50)

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	791	<div> <div>2%</div> <div> <div></div> <div>82%</div> <div>12%</div> <div>• •</div> </div> </div>
1	F	791	<div> <div>5%</div> <div> <div></div> <div>81%</div> <div>13%</div> <div>• 5%</div> </div> </div>
2	C	18	<div> <div></div> <div> <div>78%</div> <div>22%</div> </div> </div>
2	H	18	<div> <div></div> <div> <div>39%</div> <div>56%</div> <div>6%</div> </div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	G	24	<div><div><div>4%</div><div>50%</div><div>29%</div><div>21%</div></div></div>
3	J	24	<div><div><div>4%</div><div>79%</div><div>8%</div><div>13%</div></div></div>

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 13456 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 Uncharacterized protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	758	Total	C	N	O	S	0	0	0
			5757	3671	1022	1048	16			
1	F	755	Total	C	N	O	S	0	0	0
			5725	3649	1013	1047	16			

There are 30 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-13	MET	-	initiating methionine	UNP A4WYU7
A	-12	HIS	-	expression tag	UNP A4WYU7
A	-11	HIS	-	expression tag	UNP A4WYU7
A	-10	HIS	-	expression tag	UNP A4WYU7
A	-9	HIS	-	expression tag	UNP A4WYU7
A	-8	HIS	-	expression tag	UNP A4WYU7
A	-7	HIS	-	expression tag	UNP A4WYU7
A	-6	ASP	-	expression tag	UNP A4WYU7
A	-5	TYR	-	expression tag	UNP A4WYU7
A	-4	LYS	-	expression tag	UNP A4WYU7
A	-3	ASP	-	expression tag	UNP A4WYU7
A	-2	ASP	-	expression tag	UNP A4WYU7
A	-1	ASP	-	expression tag	UNP A4WYU7
A	0	ASP	-	expression tag	UNP A4WYU7
A	1	LYS	-	expression tag	UNP A4WYU7
F	-13	MET	-	initiating methionine	UNP A4WYU7
F	-12	HIS	-	expression tag	UNP A4WYU7
F	-11	HIS	-	expression tag	UNP A4WYU7
F	-10	HIS	-	expression tag	UNP A4WYU7
F	-9	HIS	-	expression tag	UNP A4WYU7
F	-8	HIS	-	expression tag	UNP A4WYU7
F	-7	HIS	-	expression tag	UNP A4WYU7
F	-6	ASP	-	expression tag	UNP A4WYU7
F	-5	TYR	-	expression tag	UNP A4WYU7
F	-4	LYS	-	expression tag	UNP A4WYU7

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	-3	ASP	-	expression tag	UNP A4WYU7
F	-2	ASP	-	expression tag	UNP A4WYU7
F	-1	ASP	-	expression tag	UNP A4WYU7
F	0	ASP	-	expression tag	UNP A4WYU7
F	1	LYS	-	expression tag	UNP A4WYU7

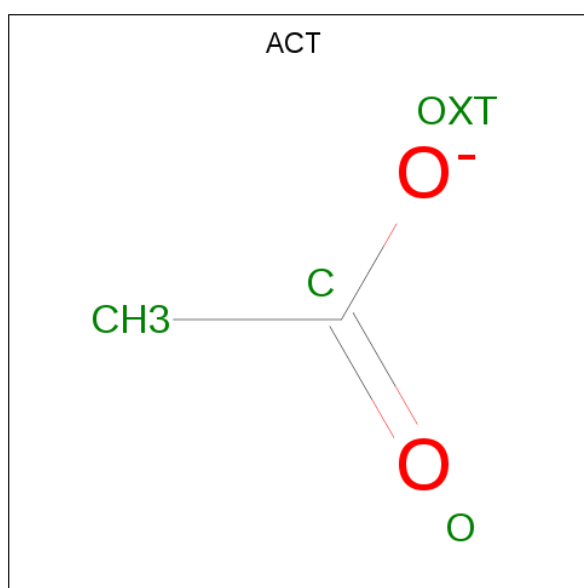
- Molecule 2 is a RNA chain called RNA (5'-R(P*UP*UP*AP*CP*UP*GP*CP*AP*CP*AP*GP*GP*UP*GP*AP*CP*GP*A)-3').

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	18	Total	C	N	O	P	0	0	0
			386	172	70	126	18			
2	H	18	Total	C	N	O	P	0	0	0
			386	172	70	126	18			

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

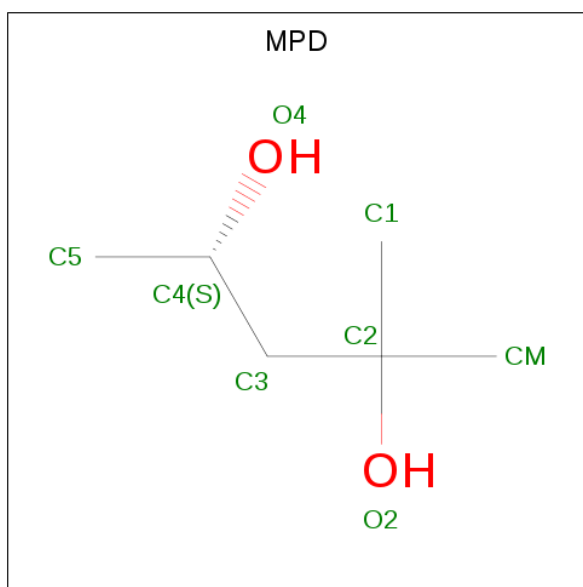
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	19	Total	C	N	O	P	0	0	0
			388	184	71	114	19			
3	J	21	Total	C	N	O	P	0	0	0
			430	204	78	127	21			

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0
4	F	1	Total C O 4 2 2	0	0

- Molecule 5 is (4S)-2-METHYL-2,4-PENTANEDIOL (three-letter code: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 8 6 2	0	0
5	H	1	Total C O 8 6 2	0	0

- Molecule 6 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	H	1	Total Mg 1 1	0	0
6	C	1	Total Mg 1 1	0	0

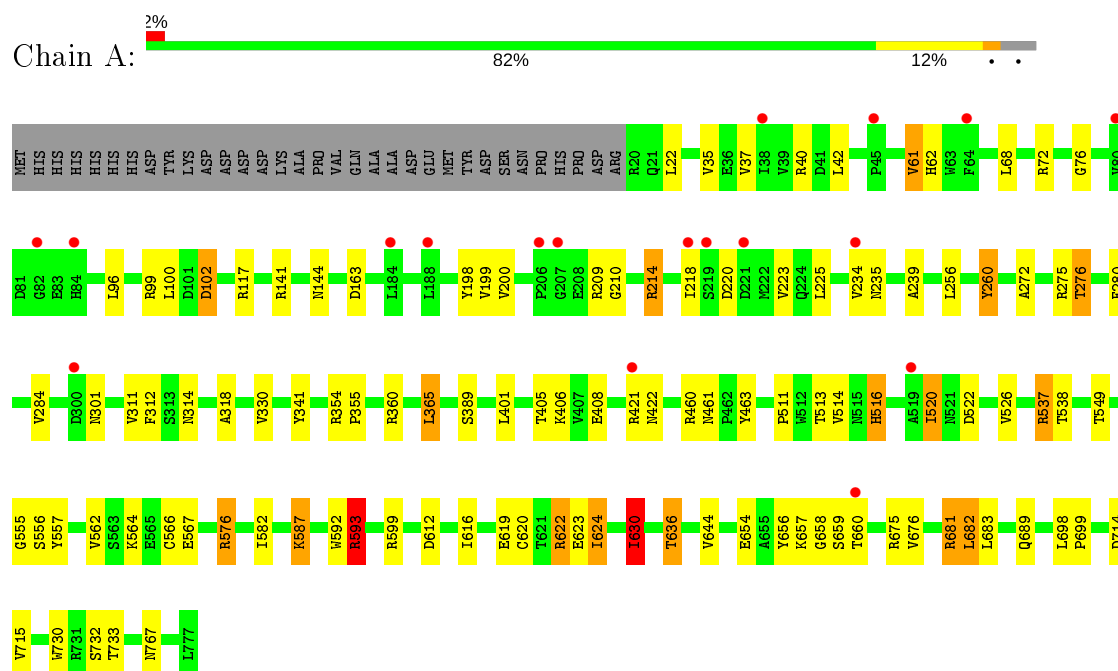
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	160	Total 160	O 160	0	0
7	C	20	Total 20	O 20	0	0
7	G	22	Total 22	O 22	0	0
7	F	110	Total 110	O 110	0	0
7	H	21	Total 21	O 21	0	0
7	J	21	Total 21	O 21	0	0

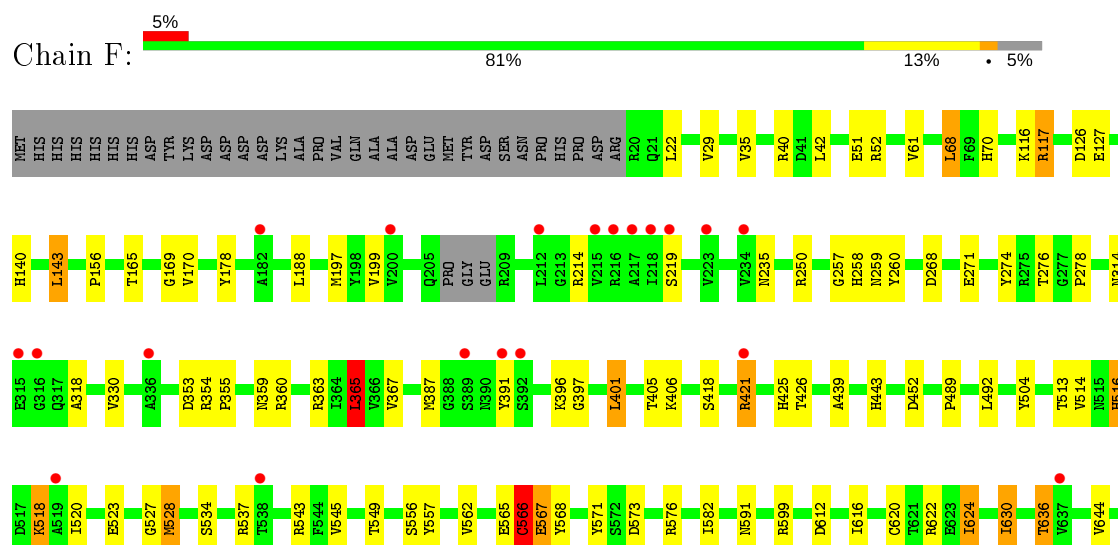
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Uncharacterized protein

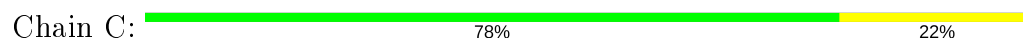


• Molecule 1: Uncharacterized protein

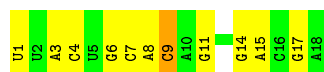




- Molecule 2: RNA (5'-R(P*UP*UP*AP*CP*UP*GP*CP*AP*CP*AP*GP*GP*UP*GP*AP*CP*GP*A)-3')



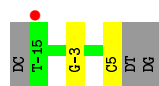
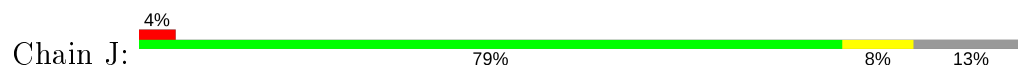
- Molecule 2: RNA (5'-R(P*UP*UP*AP*CP*UP*GP*CP*AP*CP*AP*GP*GP*UP*GP*AP*CP*GP*A)-3')



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



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



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	68.12Å 119.13Å 118.13Å 90.00° 95.60° 90.00°	Depositor
Resolution (Å)	43.54 – 2.00 43.50 – 2.00	Depositor EDS
% Data completeness (in resolution range)	97.3 (43.54-2.00) 97.3 (43.50-2.00)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.88 (at 2.00Å)	Xtriage
Refinement program	REFMAC 5.7.0032	Depositor
R, R_{free}	0.195 , 0.242 0.199 , 0.241	Depositor DCC
R_{free} test set	6206 reflections (5.04%)	wwPDB-VP
Wilson B-factor (Å ²)	47.0	Xtriage
Anisotropy	0.389	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.31 , 45.8	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	13456	wwPDB-VP
Average B, all atoms (Å ²)	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.83% 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: MPD, MG, ACT

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	0.70	0/5886	0.88	14/8011 (0.2%)
1	F	0.64	0/5852	0.80	7/7966 (0.1%)
2	C	0.93	1/431 (0.2%)	0.85	0/668
2	H	0.82	1/431 (0.2%)	0.94	3/668 (0.4%)
3	G	0.67	1/434 (0.2%)	0.90	1/667 (0.1%)
3	J	0.57	0/481	0.84	0/740
All	All	0.68	3/13515 (0.0%)	0.85	25/18720 (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
1	A	0	1
1	F	0	2
All	All	0	3

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	1	U	OP3-P	-11.40	1.47	1.61
2	H	1	U	OP3-P	-9.71	1.49	1.61
3	G	-7	DC	O3'-P	-5.68	1.54	1.61

The worst 5 of 25 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	537	ARG	NE-CZ-NH2	-9.30	115.65	120.30
1	A	537	ARG	NE-CZ-NH1	9.20	124.90	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	9	C	O5'-P-OP1	-7.92	98.57	105.70
1	A	630	ILE	CB-CA-C	-6.76	98.07	111.60
1	A	567	GLU	N-CA-C	6.76	129.26	111.00

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	566	CYS	Peptide
1	F	565	GLU	Peptide
1	F	566	CYS	Peptide

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	5757	0	5586	65	0
1	F	5725	0	5533	62	0
2	C	386	0	195	2	0
2	H	386	0	195	6	0
3	G	388	0	214	6	0
3	J	430	0	237	4	0
4	A	4	0	3	0	0
4	F	8	0	6	0	0
5	A	8	0	14	0	0
5	H	8	0	14	0	0
6	C	1	0	0	0	0
6	H	1	0	0	0	0
7	A	160	0	0	5	0
7	C	20	0	0	3	0
7	F	110	0	0	6	0
7	G	22	0	0	1	0
7	H	21	0	0	0	0
7	J	21	0	0	0	0
All	All	13456	0	11997	139	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 6.

The worst 5 of 139 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:256:LEU:CB	7:A:1041:HOH:O	2.08	1.01
1:F:513:THR:HG23	1:F:557:TYR:O	1.63	0.99
1:F:117:ARG:CB	7:F:1005:HOH:O	2.21	0.86
1:A:593:ARG:HG3	1:A:593:ARG:HH11	1.40	0.85
3:G:1:DG:OP2	7:G:101:HOH:O	1.93	0.85

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	756/791 (96%)	732 (97%)	23 (3%)	1 (0%)	51	49
1	F	751/791 (95%)	710 (94%)	35 (5%)	6 (1%)	19	13
All	All	1507/1582 (95%)	1442 (96%)	58 (4%)	7 (0%)	29	23

5 of 7 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	259	ASN
1	F	219	SER
1	F	518	LYS
1	A	220	ASP
1	F	439	ALA

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	579/672 (86%)	539 (93%)	40 (7%)	15	11
1	F	575/672 (86%)	540 (94%)	35 (6%)	18	14
All	All	1154/1344 (86%)	1079 (94%)	75 (6%)	17	12

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	636	THR
1	F	35	VAL
1	F	644	VAL
1	A	644	VAL
1	A	683	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	767	ASN
1	F	689	GLN
1	F	92	HIS
1	A	144	ASN
1	F	359	ASN

5.3.3 RNA [i](#)

Mol	Chain	Analysed	Backbone Outliers	Pucker Outliers
2	C	17/18 (94%)	1 (5%)	0
2	H	17/18 (94%)	2 (11%)	0
All	All	34/36 (94%)	3 (8%)	0

All (3) RNA backbone outliers are listed below:

Mol	Chain	Res	Type
2	C	12	G
2	H	11	G
2	H	17	G

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 monosaccharides in this entry.

5.6 Ligand geometry [i](#)

Of 7 ligands modelled in this entry, 2 are monoatomic - leaving 5 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	ACT	F	802	-	1,3,3	0.95	0	0,3,3	0.00	-
5	MPD	H	102	-	7,7,7	0.33	0	9,10,10	0.92	0
5	MPD	A	802	-	7,7,7	0.59	0	9,10,10	0.94	0
4	ACT	F	801	-	1,3,3	2.06	1 (100%)	0,3,3	0.00	-
4	ACT	A	801	-	1,3,3	1.66	0	0,3,3	0.00	-

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	MPD	A	802	-	-	2/5/5/5	-

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
5	MPD	H	102	-	-	0/5/5/5	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	F	801	ACT	CH3-C	2.06	1.51	1.48

There are no bond angle outliers.

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	802	MPD	C2-C3-C4-O4
5	A	802	MPD	C2-C3-C4-C5

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	758/791 (95%)	-0.05	18 (2%) 59 57	32, 51, 84, 110	2 (0%)
1	F	755/791 (95%)	0.22	36 (4%) 30 29	33, 57, 88, 128	1 (0%)
2	C	18/18 (100%)	-0.82	0 100 100	35, 41, 65, 104	0
2	H	18/18 (100%)	-0.85	0 100 100	39, 46, 57, 68	0
3	G	19/24 (79%)	-0.35	1 (5%) 26 25	40, 48, 109, 133	0
3	J	21/24 (87%)	-0.10	1 (4%) 30 29	41, 56, 100, 134	0
All	All	1589/1666 (95%)	0.06	56 (3%) 44 43	32, 54, 86, 134	3 (0%)

The worst 5 of 56 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	660	THR	5.0
1	A	218	ILE	4.8
1	F	219	SER	4.3
1	F	656	TYR	4.0
1	F	218	ILE	4.0

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 monosaccharides 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(Å ²)	Q<0.9
4	ACT	F	801	4/4	0.84	0.24	60,68,68,71	0
4	ACT	A	801	4/4	0.85	0.15	58,62,66,66	0
5	MPD	H	102	8/8	0.86	0.26	55,59,63,73	0
4	ACT	F	802	4/4	0.89	0.14	60,60,60,64	0
5	MPD	A	802	8/8	0.91	0.32	62,66,71,71	0
6	MG	H	101	1/1	0.99	0.13	37,37,37,37	0
6	MG	C	101	1/1	0.99	0.10	35,35,35,35	0

6.5 Other polymers [i](#)

There are no such residues in this entry.