



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 7, 2019 – 04:39 PM EST

PDB ID : 6M9T
Title : Crystal structure of EP3 receptor bound to misoprostol-FA
Authors : Audet, M.; White, K.L.; Breton, B.; Zarzycka, B.; Han, G.W.; Lu, Y.; Gati, C.; Batyuk, A.; Popov, P.; Velasquez, J.; Manahan, D.; Hu, H.; Weierstall, U.; Liu, W.; Shui, W.; Katrich, V.; Cherezov, V.; Hanson, M.A.; Stevens, R.C.
Deposited on : 2018-08-24
Resolution : 2.50 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

<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 : 1.7.3 (157068), CSD as539be (2018)
Xtrriage (Phenix) : 1.13
EDS : rb-20031633
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)
Refmac : 5.8.0158
CCP4 : 7.0 (Gargrove)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20031633

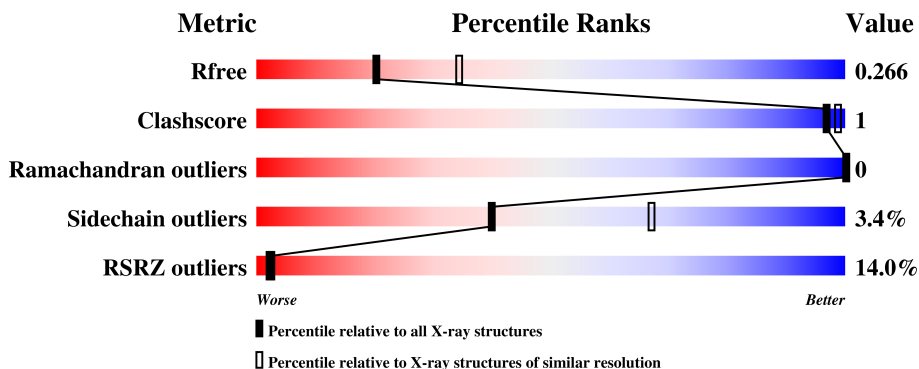
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.50 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}	111664	4155 (2.50-2.50)
Clashscore	122126	4827 (2.50-2.50)
Ramachandran outliers	120053	4735 (2.50-2.50)
Sidechain outliers	120020	4737 (2.50-2.50)
RSRZ outliers	108989	4058 (2.50-2.50)

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. 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	537	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 3570 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 Prostaglandin E2 receptor EP3 subtype, Endolysin chimera.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	444	3438	2237	585	595	21	0	2	0

There are 44 discrepancies between the modelled and reference sequences:

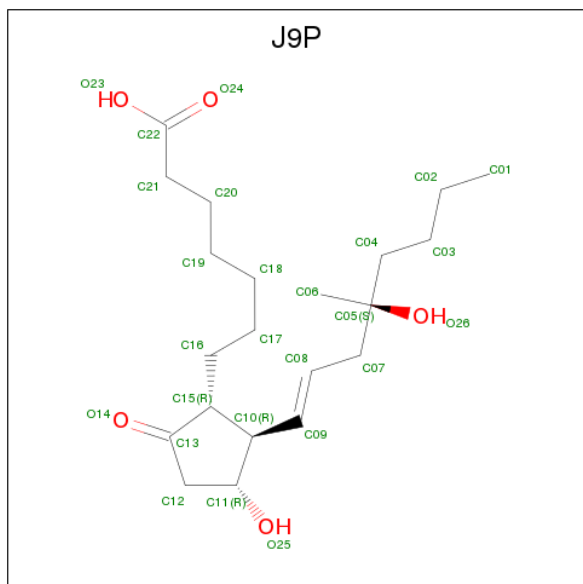
Chain	Residue	Modelled	Actual	Comment	Reference
A	-11	VAL	-	expression tag	UNP P43115
A	-10	PHE	-	expression tag	UNP P43115
A	-9	ALA	-	expression tag	UNP P43115
A	-8	ASP	-	expression tag	UNP P43115
A	-7	TYR	-	expression tag	UNP P43115
A	-6	LYS	-	expression tag	UNP P43115
A	-5	ASP	-	expression tag	UNP P43115
A	-4	ASP	-	expression tag	UNP P43115
A	-3	ASP	-	expression tag	UNP P43115
A	-2	ASP	-	expression tag	UNP P43115
A	-1	GLY	-	expression tag	UNP P43115
A	0	ALA	-	expression tag	UNP P43115
A	1	PRO	-	expression tag	UNP P43115
A	1000	GLY	-	linker	UNP P43115
A	1001	SER	-	linker	UNP P43115
A	1012	GLY	ARG	variant	UNP P00720
A	1054	THR	CYS	engineered mutation	UNP P00720
A	1074	THR	ALA	engineered mutation	UNP P00720
A	1097	ALA	CYS	engineered mutation	UNP P00720
A	1137	ARG	ILE	variant	UNP P00720
A	1162	GLY	-	linker	UNP P00720
A	1163	SER	-	linker	UNP P00720
A	286	ALA	GLY	engineered mutation	UNP P43115
A	2000	GLY	-	expression tag	UNP P43115
A	2001	ARG	-	expression tag	UNP P43115
A	2002	PRO	-	expression tag	UNP P43115
A	2003	LEU	-	expression tag	UNP P43115

Continued on next page...

Continued from previous page...

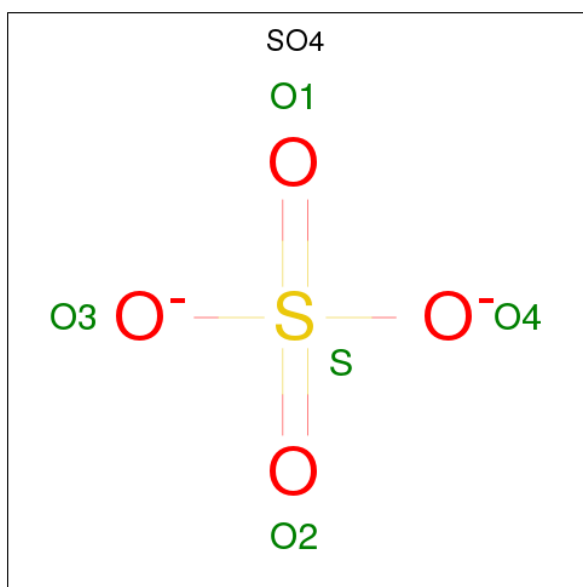
Chain	Residue	Modelled	Actual	Comment	Reference
A	2004	GLU	-	expression tag	UNP P43115
A	2005	VAL	-	expression tag	UNP P43115
A	2006	LEU	-	expression tag	UNP P43115
A	2007	PHE	-	expression tag	UNP P43115
A	2008	GLN	-	expression tag	UNP P43115
A	2009	GLY	-	expression tag	UNP P43115
A	2010	PRO	-	expression tag	UNP P43115
A	2011	HIS	-	expression tag	UNP P43115
A	2012	HIS	-	expression tag	UNP P43115
A	2013	HIS	-	expression tag	UNP P43115
A	2014	HIS	-	expression tag	UNP P43115
A	2015	HIS	-	expression tag	UNP P43115
A	2016	HIS	-	expression tag	UNP P43115
A	2017	HIS	-	expression tag	UNP P43115
A	2018	HIS	-	expression tag	UNP P43115
A	2019	HIS	-	expression tag	UNP P43115
A	2020	HIS	-	expression tag	UNP P43115

- Molecule 2 is (11alpha,12alpha,13E,16S)-11,16-dihydroxy-16-methyl-9-oxoprost-13-en-1-oic acid (three-letter code: J9P) (formula: C₂₁H₃₆O₅).



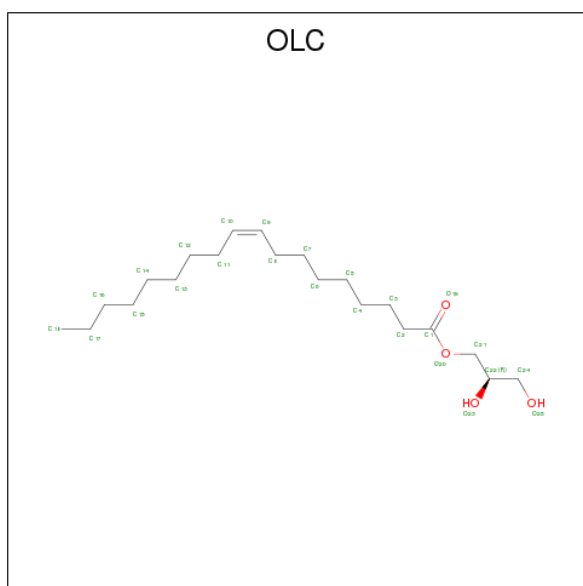
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	C O	0	0
			26	21 5		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O₄S).



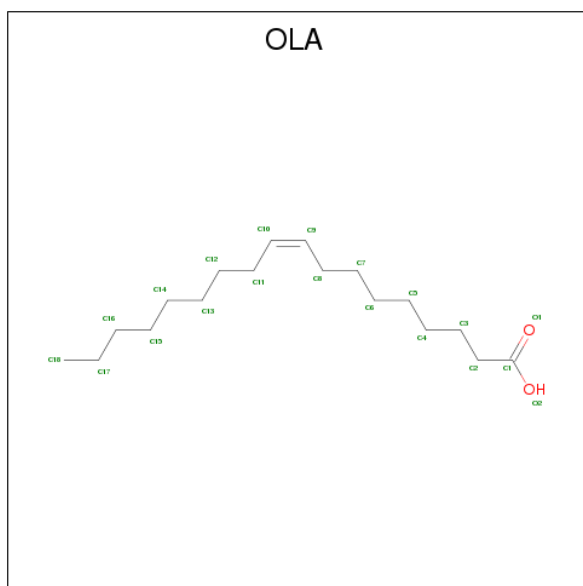
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0

- Molecule 4 is (2R)-2,3-dihydroxypropyl (9Z)-octadec-9-enoate (three-letter code: OLC) (formula: C₂₁H₄₀O₄).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			21	17	4		

- Molecule 5 is OLEIC ACID (three-letter code: OLA) (formula: $C_{18}H_{34}O_2$).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	A	1	Total	C	O	0	0
			15	13	2		
5	A	1	Total	C	O	0	0
			9	7	2		
5	A	1	Total	C	O	0	0
			16	14	2		
5	A	1	Total	C	O	0	0
			20	18	2		

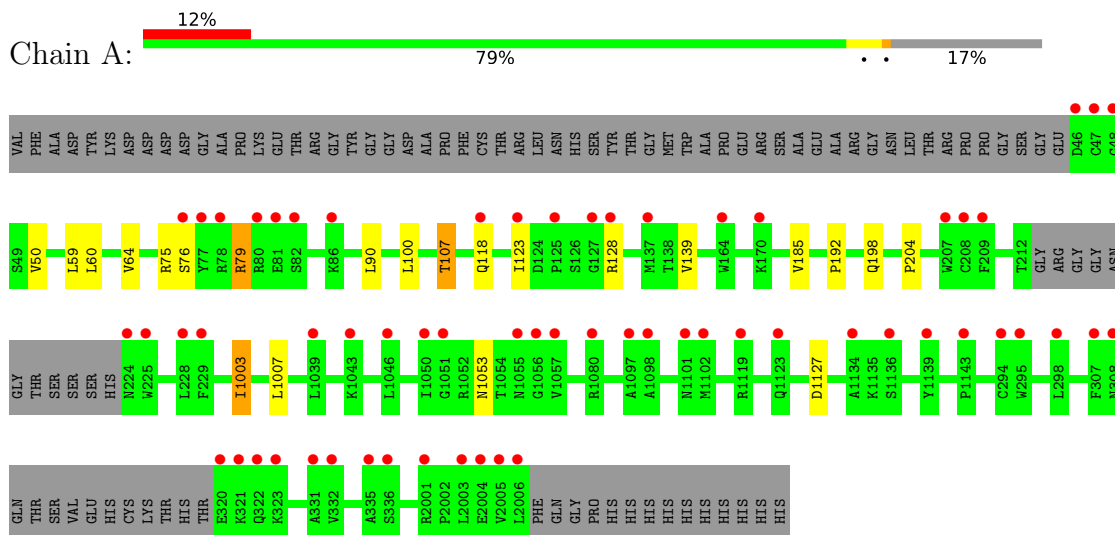
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	5	Total	O	0	0
			5	5		

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: Prostaglandin E2 receptor EP3 subtype, Endolysin chimera



4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α , β , γ	120.80Å 54.85Å 96.80Å 90.00° 95.83° 90.00°	Depositor
Resolution (Å)	27.27 – 2.50 27.19 – 2.50	Depositor EDS
% Data completeness (in resolution range)	100.0 (27.27-2.50) 100.0 (27.19-2.50)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	1.11 (at 2.51Å)	Xtrriage
Refinement program	BUSTER 2.10.2	Depositor
R, R_{free}	0.200 , 0.242 0.228 , 0.266	Depositor DCC
R_{free} test set	1050 reflections (4.75%)	wwPDB-VP
Wilson B-factor (Å ²)	80.7	Xtrriage
Anisotropy	0.220	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.29 , 119.6	EDS
L-test for twinning ²	$\langle L \rangle = 0.48$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	3570	wwPDB-VP
Average B, all atoms (Å ²)	128.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.12% 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 [i](#)

5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: OLA, OLC, SO4, J9P

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	0.43	0/3509	0.58	0/4766

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 [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	3438	0	3497	7	0
2	A	26	0	0	0	0
3	A	20	0	0	0	0
4	A	21	0	29	0	0
5	A	60	0	85	0	0
6	A	5	0	0	0	0
All	All	3570	0	3611	7	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 1.

All (7) 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:59:LEU:HB2	1:A:107:THR:HG21	1.83	0.59
1:A:192:PRO:HB3	1:A:198:GLN:HA	1.90	0.52
1:A:1003:ILE:H	1:A:1003:ILE:HG13	1.55	0.47
1:A:50:VAL:HG22	1:A:204:PRO:HG3	2.01	0.42
1:A:139:VAL:HG23	1:A:185:VAL:HG12	2.01	0.42
1:A:76:SER:O	1:A:79:ARG:HD2	2.20	0.41
1:A:60:LEU:O	1:A:64:VAL:HG23	2.21	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	440/537 (82%)	418 (95%)	22 (5%)	0	100 100

There are no Ramachandran outliers to report.

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	356/448 (80%)	344 (97%)	12 (3%)	40 67

All (12) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	75	ARG
1	A	79	ARG
1	A	90	LEU
1	A	100	LEU
1	A	107	THR
1	A	118	GLN
1	A	123	ILE
1	A	128	ARG
1	A	1003	ILE
1	A	1007	LEU
1	A	1053	ASN
1	A	1127	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1122	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](#)

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](#)

10 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
2	J9P	A	2101	-	21,26,26	3.53	7 (33%)	22,34,34	0.84	2 (9%)
3	SO4	A	2102	-	4,4,4	0.20	0	6,6,6	0.19	0
3	SO4	A	2103	-	4,4,4	0.19	0	6,6,6	0.07	0
3	SO4	A	2104	-	4,4,4	0.19	0	6,6,6	0.06	0
3	SO4	A	2105	-	4,4,4	0.20	0	6,6,6	0.07	0
4	OLC	A	2106	-	20,20,24	0.51	0	21,21,25	0.90	1 (4%)
5	OLA	A	2107	-	11,14,19	0.38	0	10,14,19	0.67	0
5	OLA	A	2108	-	5,8,19	0.28	0	4,8,19	0.82	0
5	OLA	A	2109	-	12,15,19	0.39	0	11,15,19	0.69	0
5	OLA	A	2110	-	16,19,19	0.36	0	15,19,19	0.76	0

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. '2' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	J9P	A	2101	-	-	0/19/37/37	0/1/1/1
3	SO4	A	2102	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2103	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2104	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2105	-	-	0/0/0/0	0/0/0/0
4	OLC	A	2106	-	-	0/20/20/24	0/0/0/0
5	OLA	A	2107	-	-	0/10/12/17	0/0/0/0
5	OLA	A	2108	-	-	0/4/6/17	0/0/0/0
5	OLA	A	2109	-	-	0/11/13/17	0/0/0/0
5	OLA	A	2110	-	-	0/15/17/17	0/0/0/0

All (7) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	A	2101	J9P	C12-C11	-11.32	1.39	1.52
2	A	2101	J9P	C10-C15	-7.23	1.41	1.55
2	A	2101	J9P	O25-C11	-2.74	1.37	1.43
2	A	2101	J9P	O26-C05	-2.58	1.40	1.44
2	A	2101	J9P	C15-C13	2.23	1.55	1.52
2	A	2101	J9P	C12-C13	4.52	1.57	1.51
2	A	2101	J9P	C10-C11	5.74	1.67	1.55

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	2101	J9P	C06-C05-C07	-2.15	106.77	110.66
2	A	2101	J9P	C07-C08-C09	-2.06	118.92	125.40
4	A	2106	OLC	O20-C1-C2	2.49	119.12	111.92

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

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, 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	444/537 (82%)	0.59	62 (13%) 2 2	82, 126, 179, 206	0

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	82	SER	10.0
1	A	1056	GLY	5.7
1	A	81	GLU	4.9
1	A	322	GLN	4.8
1	A	307	PHE	4.8
1	A	320	GLU	4.4
1	A	1057	VAL	4.4
1	A	80	ARG	4.1
1	A	76	SER	4.1
1	A	323	LYS	4.0
1	A	335	ALA	4.0
1	A	47	CYS	4.0
1	A	2005	VAL	3.9
1	A	170	LYS	3.8
1	A	1098	ALA	3.8
1	A	225	TRP	3.7
1	A	2004	GLU	3.7
1	A	336	SER	3.6
1	A	228	LEU	3.6
1	A	2003	LEU	3.6
1	A	295	TRP	3.5
1	A	224	ASN	3.4
1	A	1039	LEU	3.4
1	A	332	VAL	3.4
1	A	78	ARG	3.4
1	A	1139	TYR	3.3
1	A	77	TYR	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	1097	ALA	3.1
1	A	298	LEU	3.0
1	A	118	GLN	3.0
1	A	125	PRO	2.9
1	A	127	GLY	2.9
1	A	229	PHE	2.8
1	A	2006	LEU	2.7
1	A	2001	ARG	2.6
1	A	48	GLY	2.6
1	A	208	CYS	2.5
1	A	207	TRP	2.5
1	A	1143	PRO	2.5
1	A	46	ASP	2.4
1	A	123	ILE	2.4
1	A	1046	LEU	2.4
1	A	1043	LYS	2.4
1	A	137	MET	2.3
1	A	86	LYS	2.3
1	A	1119	ARG	2.3
1	A	1055	ASN	2.2
1	A	1051	GLY	2.2
1	A	321	LYS	2.2
1	A	1134	ALA	2.2
1	A	331	ALA	2.2
1	A	294	CYS	2.2
1	A	308	ASN	2.2
1	A	1080	ARG	2.2
1	A	164	TRP	2.1
1	A	1101	ASN	2.1
1	A	1136	SER	2.1
1	A	209	PHE	2.1
1	A	1102	MET	2.1
1	A	128	ARG	2.1
1	A	1123	GLN	2.1
1	A	1050	ILE	2.1

6.2 Non-standard residues in protein, DNA, RNA chains

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
4	OLC	A	2106	21/25	0.64	0.30	128,140,148,148	0
5	OLA	A	2108	9/20	0.71	0.31	165,167,171,172	0
5	OLA	A	2107	15/20	0.72	0.35	122,123,129,130	0
3	SO4	A	2105	5/5	0.72	0.26	200,200,200,200	0
5	OLA	A	2109	16/20	0.80	0.37	109,113,119,121	0
5	OLA	A	2110	20/20	0.85	0.21	110,113,116,117	0
3	SO4	A	2104	5/5	0.86	0.15	219,219,219,219	0
3	SO4	A	2103	5/5	0.89	0.17	175,175,175,175	0
2	J9P	A	2101	26/26	0.93	0.37	79,93,105,107	0
3	SO4	A	2102	5/5	0.93	0.10	129,130,131,132	0

6.5 Other polymers [i](#)

There are no such residues in this entry.